

1. INTRODUCCIÓN

Para que un estudio sobre la demanda en el transporte aéreo resulte completo, debemos vincularlo de manera obligatoria con un amplio análisis territorial y demográfico de la zona objeto de investigación, y es precisamente ese vínculo el que trata de comprender y analizar este proyecto de fin de carrera. Es lógico por tanto establecer la comprensión de esta simbiosis como objetivo principal del estudio.

Para una mayor comprensión del tema de estudio es fundamental conocer la maraña de factores que intervienen en él, aunque su análisis detallado no está dentro de los márgenes de esta investigación. El estudio geográfico del transporte aéreo ha adquirido hoy en día una complejidad considerable al intervenir variables de tipo geográfico, geoeconómico, empresarial, social, urbanístico y medioambiental. La bibliografía existente en castellano sobre demanda de transporte aéreo y demografía no es muy extensa, aunque existen magníficos textos de cada una de los temas por separado. Este proyecto se propone contribuir de forma humilde, aportando más luz sobre los datos de los que ya disponemos, haciendo un análisis detallado de muchas de las variables que intervienen en la demanda del transporte aéreo, así como proporcionar unas conclusiones y posibles propuestas para mejorar en la medida de lo posible la situación actual.

Por otra parte el desarrollo del transporte aéreo adquirió un enorme impulso en los veinticinco años posteriores a la Segunda Guerra Mundial. En las dos décadas transcurridas entre 1950 y 1970 el número de pasajeros se multiplicó por un factor de diez.

Posteriormente, la tasa de crecimiento se ha moderado manteniéndose en cifras próximas al 5% anual. Hoy en día más del 90% del tráfico mundial se desarrolla en y entre América del Norte, Europa y Asia/Pacífico. Estados Unidos lidera de forma abismal el tráfico aéreo desde 1929, momento en que el transporte aéreo sufrió una enorme expansión, seguido de China, Reino Unido o Alemania.

Por lo que se refiere a la evolución de la demanda de transporte aéreo en España, ha pasado por tres fases diferenciadas después de la Guerra Civil. En primer lugar entre los años 1940 y 1959 el crecimiento fue muy bajo ya que los intercambios comerciales y turísticos con el extranjero estaban muy limitados, a consecuencia de la política

económica autárquica mantenida por el Gobierno español; y en cuanto al tráfico nacional, también era bajo por la falta de poder adquisitivo de la población. Más adelante, durante las décadas de los 60 y 70 se produjo una fuerte expansión de la demanda gracias a la apertura de la economía al exterior y la llegada de un gran número de turistas procedentes del norte y centro de Europa. En los años posteriores el ritmo de crecimiento se redujo hasta que la evolución del sector se acompasó con el de la economía.

La evolución de la demanda de transporte aéreo nacional e internacional es muy distinta. La demanda nacional está íntimamente relacionada con la evolución económica empujada por volumen de negocio de las empresas y el aumento de la renta disponible por la población. Si tenemos en cuenta el fuerte carácter vacacional del transporte aéreo, la demanda en España ha dependido también de la situación económica de los principales países emisores (Alemania y Reino Unido), como del atractivo turístico de España comparado con respecto a otros destinos similares como Grecia, Italia, Portugal, Turquía o la antigua Yugoslavia.

Por otro lado, algunos autores sostienen que el concepto tradicional de aeropuerto-ciudad ha entrado en crisis y está siendo sustituido por el de aeropuerto-región, ampliando su radio de influencia más allá de los límites de la ciudad a la que directamente sirve. En algunas regiones se desarrollan *sistemas multiaeroportuarios* en los que varios aeropuertos comparten funciones. Se convierten además en un *nodo plurimodal* de transporte conectado hacia el exterior por redes de alta velocidad y autopistas, así como hacia el interior del sistema metropolitano por modos de transporte público, urbano y regional.

Respecto al radio de acción del sistema aeropuerto-región, renuncian a establecer una distancia física ya que ésta puede modificarse en función de la densidad de población y de la presencia o ausencia de los enlaces de conexión terrestres. Delimitan por lo tanto el radio de acción del aeropuerto-región a todos los puntos de la geografía circundante al sistema aeroportuario desde los que se invierte un tiempo igual o inferior a 90 minutos para acceder al aeropuerto. Anotan también que en el caso de grandes *hubs* éste área de influencia podría ampliarse.

De entre varios estudios, uno de los más interesantes es el de Walter Christaller, que examina cómo diversos productos y servicios se articulan en el territorio dando lugar a

una jerarquía urbana. El modelo ordena los centros urbanos de una región para establecer una jerarquía de tamaños, que a su vez refleja que el tamaño del mercado de cualquier bien o servicio tiene un límite máximo y otro mínimo. El límite mínimo se denomina *umbral* y se define como la escala mínima que permite que se oferte el servicio. El límite máximo se denomina *amplitud* y su determinante fundamental es el coste de transporte. Conjugando el umbral y la amplitud se determina el número y tamaño de los lugares centrales que ofrece cada servicio, y a partir de ellos la jerarquía de ciudades. Ésta es la esencia del *paradigma del lugar central* de Christaller, también conocido como *paradigma chistalleriano*. En términos conceptuales, la estructura urbana derivada del paradigma del lugar central es asimilable con un gráfico en forma de árbol. Es muy interesante, también, cómo el modelo de Christaller da lugar a áreas de mercado hexagonales. Esta geometría, como se intuye, es la más eficiente dado que las áreas no se solapan ni se dejan zonas sin cubrir, exactamente igual que en un panel. En esta clasificación se utilizan las enseñanzas de Christaller para determinar el rango de cada ciudad haciendo cuatro divisiones: *grandes metrópolis nacionales* o *nodos*, *metrópolis nacionales de primer orden*, *metrópolis nacionales de segundo orden* y *ciudades intermedias*.

Uno de los aspectos interesantes en un estudio de este tipo es el de la competencia modal. El 80% del transporte de pasajeros hoy en día se realiza por carretera. El modo aéreo tiene una cuota de mercado algo superior al 10%, mientras que el ferrocarril se ocupa del resto con muy pequeñas participaciones del transporte marítimo y fluvial. En el caso de la Unión Europea, el coche abarca el 71% del tráfico, el autobús el 10%, el tren el 7% y el avión lo hace en un 12%. Son datos similares a los de Estados Unidos, aunque en América del Norte la utilización del autobús y el tren es prácticamente nula (3% y 1% respectivamente), en comparación con el coche (80%) o el avión (16%).

Las ventajas de las carreteras son importantes, en términos de flexibilidad, comodidad y extensión de infraestructuras.

Dentro de los diferentes elementos que pueden considerarse a la hora de establecer un ámbito modal, es decir, aquellos factores que califican a un modo de transporte como alternativa válida para un cierto viaje, destacan cinco: distancia efectiva del viaje, velocidad real del viaje, sujeto del transporte, competencia intermodal y coste generalizado.

Centrando la atención en la competencia modal, ésta se da cuando un viaje puede ser realizado en varios modos de transporte y su intensidad varía según ciertos condicionantes como son los tecnológicos, económicos, políticos y sociales.

Hay un tema particularmente interesante que no puede pasarse por alto en un estudio como este. Se trata de cómo el hombre, a lo largo de la historia, ha conseguido reducir el efecto de la fricción de la distancia de forma considerable. El transporte aéreo (entre otros modos de transporte) ha conseguido deformar el espacio de una u otra manera.

El aumento de la velocidad real del viaje hace que la distancia medida en tiempo sea mucho mejor.

La *convergencia espaciotiempo* es por lo tanto una medida de cómo dos lugares se acercan si la distancia que les separa se calcula en tiempo. Hay que tener en cuenta que este proceso no es ni mucho menos uniformador, de hecho favorece particularmente las relaciones entre las principales ciudades que reciben mayor inversión.

Pero no sólo se acortan distancias desde la óptica del tiempo; el transporte aéreo en su proceso de contracción del espacio también lo hace si las distancias se miden en costes.

Los costes de las telecomunicaciones y los transportes se han abaratado de forma espectacular en el último siglo. Por lo tanto se está produciendo también una *convergencia espacio-coste*.

Esta *convergencia espacio-tiempo* y *espacio-coste* en los transportes y las telecomunicaciones tiene importantes consecuencias económicas y sociales. Debido a la reducción del tiempo y de los costes de la interacción, se pueden alcanzar mercados más lejanos, con lo que se incrementa la competitividad general del sistema y se favorece la aparición de las economías de escala y de los beneficios de la especialización.

El transporte aéreo responde a la creciente demanda de movilidad con una fuerte inversión en: innovación tecnológica; seguridad y mejoras en la seguridad; renovación de la flota con aviones más silenciosos y combustibles más eficientes; aeropuertos, rutas aéreas y mejoras de gestión del tráfico aéreo.

Además, los esfuerzos de la industria deben ir acompañados de la acción gubernamental para liberalizar más los mercados de aviación sin microgestión; conseguir mejoras en la infraestructura mediante el aumento de la capacidad del aeropuerto, así como nuevas rutas y más cortas; y proporcionar un marco global al transporte aéreo, que no se base en reglas nacionales.

Hay que tener en cuenta que la competencia dentro del transporte aéreo ha aumentado de forma muy significativa en los mercados liberados, como el nacional norteamericano, el Atlántico Norte o el intercomunitario. El sector del transporte aéreo tiene un gran número de empresas y las barreras de acceso al mercado son muy pequeñas, por lo que la creación de nuevos servicios y su desaparición es constante. Actualmente coexisten varios modelos de mercado: *compañías de red*, con sus afiliadas subsidiarias y regionales y *compañías punto a punto*, bien cubriendo la necesidad del mercado de viajes vacacionales (*Chárter*), o bien centrándose en vuelos de corto y medio radio, atendiendo a mercados tanto vacacionales como de negocios (*low cost*). Las compañías de red ofrecen una amplia gama de servicios comercializados a través de canales de ventas diferentes, y concentran su actividad en uno o varios aeropuertos nodales (*hub*). La complejidad de la operación a través de un hub (*hub&spoke*) hace difícil que un mismo aeropuerto albergue a más de una compañía de red. El desarrollo de destinos de largo radio hace necesaria la concentración del tráfico en un punto nodal. Por otra parte, hay una tendencia a la globalización de la industria; a medida que se reducen las barreras regulatorias las compañías de red se agrupan en alianzas con una cobertura de mercado mundial.

De un tiempo a esta parte, se ha implantado otro tipo de mercado punto a punto.

Los operadores ofrecen servicios a muy bajo precio gracias a la eliminación de todos los elementos no esenciales en el viaje. Son los conocidos como *low cost* o *no-frills* que surgieron en los años 60 en Estados Unidos y alcanzaron importancia mundial después de los atentados del 11 de Septiembre. Sus vuelos son exclusivamente punto a punto y se comercializan a través del teléfono y en mucha mayor medida gracias a internet. Estos vuelos buscan aeropuertos secundarios, en los que las tasas sean más baratas. Además, las tasas de ocupación de los aviones son las más altas, reduciendo al mínimo los tiempos de escala. No ofrecen conexiones y no tienen ningún interés por pertenecer a ninguna alianza.

Las tarifas que ofrecen las compañías de bajo coste son tan bajas que han creado mercados en lugares donde antes no había demanda o era muy limitada. Por otra parte, están ganando cada vez más mercado de los chárter ofreciendo tarifas muy reducidas sin la obligación de adquirir el paquete turístico completo.

El éxito del *low cost* está más que comprobado en el medio radio pero no ha funcionado con el mismo éxito en los vuelos de largo recorrido, donde el menor número de

frecuencias, los tiempos de escala más largos, la necesidad de un mejor servicio a bordo y la necesidad de combinar este servicio con conexiones de corto radio no favorecen a las compañías de este tipo. Aun así, algunas compañías como Ryanair han mostrado su interés por participar en el mercado del Atlántico Norte.

El mercado español es uno de los más competitivos del mundo y, probablemente, el que tiene mayores niveles de competencia de Europa. España es el único país de la Unión Europea en la que existen tres compañías nacionales (Air Europa, Iberia y Spanair) compitiendo por muchas rutas de medio recorrido. Debido a esta situación es probable que la demanda induzca un proceso de contracción que racionalice la oferta del sector, como ocurrió en 2008 con Vueling y Clickair.

2. AEROPUERTOS

2.1 RED AENA

AENA (Aeropuertos Españoles y Navegación Aérea) gestiona 46 aeropuertos y 2 helipuertos en el territorio español, de los cuales los tres aeropuertos principales de la red son Adolfo Suárez Madrid-Barajas, Barcelona-El Prat y Palma de Mallorca, y el resto de aeropuertos quedan integrados en alguno de los siguientes grupos:

Grupo Canarias: integrado por los 8 aeropuertos de la Comunidad Autónoma de Canarias. Debido a su lejanía al territorio continental y por la importancia del tráfico interinsular, estos aeropuertos tienen unas características que los hacen diferentes al resto de la red.

Grupo I: formado por aeropuertos con más de 2 millones de pasajeros al año. Este grupo lo componen 8 aeropuertos: Málaga-Costa del Sol, Alicante-Elche, Ibiza, Valencia, Bilbao, Sevilla, Girona-Costa Brava y Menorca.

Grupo II: formado por aeropuertos que procesan entre 0,5 y 2 millones de pasajeros al año. Este grupo lo componen 11 aeropuertos: Almería, Asturias, FGL Granada-Jaén, Jerez, A Coruña, Murcia-San Javier, Reus, Santander, Santiago, Vigo y Zaragoza.

Grupo III: formado por aeropuertos con menos de 0,5 millones de pasajeros al año. Es un grupo heterogéneo compuesto por:

- Bases aéreas abiertas al tráfico civil: Valladolid, León, Badajoz, Salamanca, Madrid-Torrejón (desde febrero de 2013 es una base militar sin tráfico civil) y Albacete.
- Aeropuertos civiles con tráfico comercial: Melilla, San Sebastián, Pamplona, Burgos-Villafría y Logroño-Agoncillo.
- Aeropuerto de carga: Vitoria.
- Helipuertos: Ceuta y Algeciras.
- Aeropuertos de aviación general: Córdoba, Sabadell, Son Bonet, Madrid-Cuatro Vientos y Huesca-Pirineos.

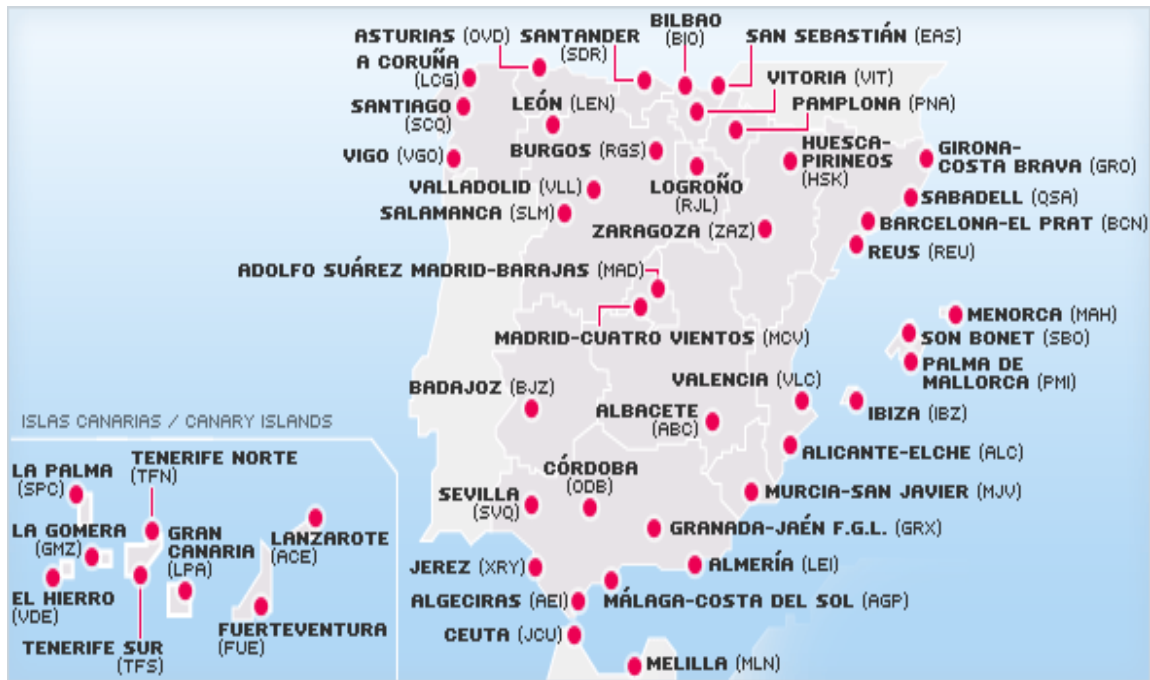


Figura 1. Aeropuertos de la red AENA

Todos estos aeropuertos gestionaron en 2013 un total de 187,4 millones de pasajeros, de los cuales 129,5 millones fueron pasajeros internacionales.

Además, tanto los aeropuertos Madrid-Barajas como Barcelona-El Prat están en el Top Ten de la Unión Europea por tráfico de pasajeros.

2.2 AEROPUERTOS ESCOGIDOS PARA LOS MODELOS

Para el análisis estadístico del presente proyecto así como de todos los datos recopilados se tienen en cuenta 39 de estos aeropuertos. Se han elegido aquellos que hayan tenido más de 10.000 pasajeros y algún destino regular en el año 2013.

Como se puede ver se han cogido la práctica totalidad de los aeropuertos españoles, dejando a un lado los de menor importancia, pues a lo largo del proyecto se explicará la difícil obtención de datos para algunos de estos; además con 39 aeropuertos se tiene un número importante de observaciones y por tanto de grados de libertad para el modelo estadístico.

Los 39 aeropuertos estudiados con más de 10.000 pasajeros y algún destino regular en 2013 son (por orden alfabético):

Aeropuerto	Aeropuerto
A Coruña	Madrid Barajas
Alicante	Málaga
Almería	Melilla
Asturias	Menorca
Badajoz	Murcia
Barcelona	Palma de Mallorca
Bilbao	Pamplona
Burgos	Reus
El Hierro	Salamanca
Fuerteventura	San Sebastián
Girona	Santander
Gran Canaria	Santiago
Granada-Jaén	Sevilla
Ibiza	Tenerife Norte
Jerez	Tenerife Sur
La Gomera	Valencia
La Palma	Valladolid
Lanzarote	Vigo
León	Zaragoza
Logroño	

Tabla 1. Aeropuertos del modelo

3. VARIABLES DE LOS MODELOS ESTADÍSTICOS

3.1 INTRODUCCIÓN

En este bloque se explica cada una de las variables introducidas en el modelo estadístico y la fuente de obtención de datos para cada variable. También se expone cómo han sido aplicados esos datos según el aeropuerto, y se adjunta para cada variable una tabla resumen según los años analizados.

En cuanto al desarrollo de la investigación del presente proyecto de fin de carrera, es fundamental aclarar que todo el análisis que se encuentra en estas páginas está centrado y referido a los datos de los ejercicios del año 2009, 2010, 2011, 2012 y 2013. No obstante, si hubiera alguna referencia fuera de este rango, se indicará de forma expresa. Es obligación señalar, también, que alguno de los datos en los que se basa este trabajo pueden ser de carácter provisional e incluso en algún caso puntual los datos utilizados son previsiones.

No se debe olvidar que este estudio está sujeto a cierta subjetividad y que se ha desarrollado con toda la cautela posible. La elección de las variables que se introducirán en el modelo ha sido desde el primer momento el mayor quebradero de cabeza del presente proyecto; pues al principio se escogió una larga lista de variables en función de la bibliografía consultada y de las diversas ideas que iban surgiendo, pero más tarde, algunas de estas se fueron descartando por varios motivos como la falta de datos o la imposible adecuación de esos datos en función de cada aeropuerto.

Además, con criterio propio y estudiando la bibliografía quedaba claro que alguna de ellas no ayudaría en la mejora del modelo e incluso lo podría empeorar, y otras se descartaron por tener demasiada relación con las ya existentes. En cualquier caso, podrían ser incluidas y analizadas con los datos precisos en futuros proyectos, así como otras variables, según el enfoque que se pretenda dar.

Como ejemplo de alguna de las variables excluidas tenemos la estacionalidad, el factor de utilización, la intermodalidad, la ocupación media de las aeronaves, las tarifas aéreas, etc...

3.2 FACTOR DE CORRECCIÓN PARA LAS VARIABLES SOCIOECONÓMICAS

En las variables socioeconómicas que se analizarán a continuación, los datos que se han obtenido vienen dados por comunidades autónomas y provincias (fuentes del Instituto Nacional de Estadística, Anuario Económico de la Caixa, Ministerio de Economía y Competitividad, etc.), por lo que si se quiere cometer el menor error posible en el modelo estadístico habrá que tratar de manera diferente los aeropuertos insulares.

Para ello a dichos aeropuertos se les aplica un factor de corrección en función de la población de la isla a la que pertenecen (tomando como referencia la población del año 2013, pues la variación es mínima entre los años 2009 y 2013). Este factor no es otra cosa que el porcentaje entre la población de una isla determinada y la población total de la provincia a la que pertenece.

El factor de corrección servirá más tarde para ser multiplicado por el dato de la provincia en la variable correspondiente para los aeropuertos insulares. Las variables donde se corrigen los datos de los aeropuertos insulares son:

- Población residente de la Unión Europea
- Índice Comercial
- Índice de Restauración y Bares
- Índice Turístico
- Índice de Actividad Económica
- Índice Industrial
- Pernotaciones
- Exportaciones
- Importaciones
- Mercado Potencial
- Cuota de Mercado

A continuación se muestra la tabla con los factores de corrección para cada aeropuerto insular (en tanto por uno):

Provincia	Aeropuerto	Factor de Corrección
Islas Baleares	Palma de Mallorca	0,7872
	Ibiza	0,1279
	Menorca	0,0849
Las Palmas	Gran Canaria	0,7744
	Lanzarote	0,1276
	Fuerteventura	0,098
Santa Cruz de Tenerife	Tenerife Norte	0,8858
	Tenerife Sur	0,8858
	La Palma	0,0828

	La Gomera	0,0209
	El Hierro	0,0105

Tabla 2. Factores de corrección

3.3 VARIABLES SOCIOECONÓMICAS

3.3.1 PASAJEROS (Variable Dependiente)

El año 2013 se caracterizó por un escenario macroeconómico negativo que mejoró en los últimos meses y que posibilitó el inicio de un cambio de tendencia en los principales datos de actividad.

La vinculación del crecimiento del tráfico al ciclo económico mundial ha condicionado la evolución del tráfico nacional e internacional en los aeropuertos españoles.

El tráfico internacional ha tenido un comportamiento mejor que el tráfico doméstico durante el periodo de crisis económica, alcanzando una cifra récord de pasajeros internacionales de 129,5 millones en 2013.

El mejor comportamiento del tráfico internacional se ha visto impulsado por el crecimiento del número de turistas que visitaron España y que alcanzó en 2013 la cifra récord de 60,7 millones. A continuación se puede ver la distribución del tráfico de pasajeros de los aeropuertos españoles por áreas geográficas en 2012 y 2013.

Zona	Pasajeros 2012	Pasajeros 2013	% Variación	Variación
Europa (excluye España)	113.302.245	116.006.871	2,4%	2.704.626
España	67.196.423	57.823.418	-13,9%	-9.373.005
América Latina y Caribe	5.741.783	5.428.774	-5,5%	-313.009
América del Norte (incluye EEUU y Canadá)	3.822.670	3.596.278	-5,9%	-226.392
África	2.315.348	2.418.357	4,4%	103.009
Oriente Medio	1.489.341	1.726.718	15,9%	237.377
Asia y Pacífico	362.889	360.931	-0,5%	-1.958
Total	194.230.699	187.361.347	-3,5%	-6.869.352

Tabla 3. Distribución del tráfico de pasajeros

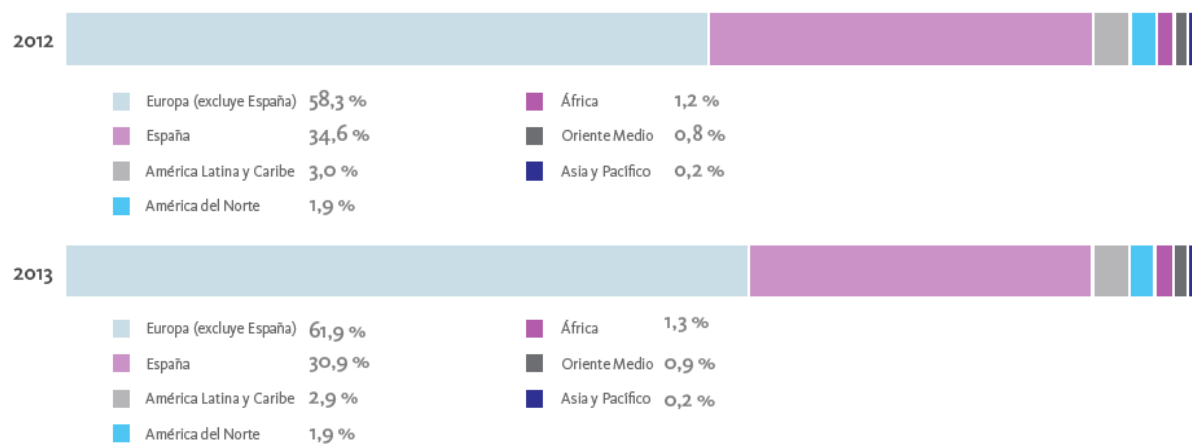


Gráfico 1. Distribución del tráfico de pasajeros

Respecto a esta distribución del tráfico por áreas geográficas destaca, un incremento de la cuota de tráfico con Europa, que pasa del 58,3% en 2012 al 61,9% en 2013, lo que ha supuesto un crecimiento del 2,4%, y en términos cuantitativos, 2,7 millones de pasajeros más que en 2012. Y además un descenso del tráfico doméstico, con una cuota del 30,9% en 2013 frente al 34,6% en 2012, consecuencia de una disminución de cerca del 14%, es decir, 9,4 millones de pasajeros menos que en 2012. El siguiente gráfico representa la evolución de la tipología del tráfico total de pasajeros (cuota nacional/internacional), en millones de pasajeros, desde el año 2000.



Gráfico 2. Cuota nacional e internacional

En la tabla adjunta podemos ver la evolución del tráfico total de pasajeros para los años de análisis según el mes, así como la variación interanual y el total.

Mes del Año	2013		2012		2011		2010		2009	
	Pasajeros	%	Pasajeros	%	Pasajeros	%	Pasajeros	%	Pasajeros	%
Enero	10.621.873	-9,20%	11.692.830	-3,00%	12.055.079	6,10%	11.366.511	3,30%	11.004.344	-16,90%
Febrero	10.386.312	-10,30%	11.577.607	-5,60%	12.269.605	4,50%	11.737.617	1,40%	11.575.715	-18,20%
Marzo	13.344.253	-5,70%	14.152.811	-7,00%	15.220.018	4,50%	14.569.431	5,40%	13.822.006	-18,70%
Abril	14.815.368	-7,70%	16.053.262	-7,30%	17.316.226	20,30%	14.391.126	-8,40%	15.702.563	-5,00%
Mayo	16.842.528	-1,80%	17.153.399	-3,80%	17.827.328	6,40%	16.752.168	3,00%	16.265.873	-11,70%
Junio	18.509.931	-3,30%	19.133.547	-1,30%	19.392.724	6,80%	18.151.704	3,20%	17.586.135	-8,60%
Julio	21.014.936	-4,50%	21.996.576	-2,60%	22.574.985	6,70%	21.155.098	3,60%	20.422.575	-4,70%
Agosto	21.523.628	-2,50%	22.083.607	-3,10%	22.785.470	4,70%	21.772.441	3,60%	21.013.670	-5,30%
Septiembre	18.998.253	-2,80%	19.546.781	-3,50%	20.257.156	7,50%	18.837.671	5,60%	17.842.298	-5,00%
Octubre	16.847.697	-0,50%	16.940.142	-7,00%	18.221.194	2,30%	17.815.496	8,60%	16.409.555	-2,50%
Noviembre	12.167.576	3,30%	11.773.508	-9,80%	13.055.573	-2,10%	13.335.564	5,20%	12.677.670	-0,60%
Diciembre	11.682.394	3,00%	11.345.067	-10,30%	12.651.925	4,00%	12.161.568	-2,30%	12.441.599	2,30%
Total	186.754.749	-3,50%	193.449.137	-5,00%	203.627.283	6,00%	192.046.395	2,80%	186.764.003	-7,90%

Tabla 4. Evolución del tráfico total de pasajeros

Todos los datos referentes a los pasajeros han sido sacados de la página web de AENA (Aeropuertos Españoles y Navegación Aérea), en la sección “Estadísticas”, donde se pueden descargar archivos tanto del tráfico por año como de la evolución del tráfico por aeropuerto, así como los informes anuales y mensuales.

A continuación, la tabla que sigue muestra los pasajeros por aeropuerto y año. Estos datos serán los introducidos en el modelo estadístico como variable dependiente.

Aeropuerto	2013	2012	2011	2010	2009
A Coruña	839.837	845.451	1.012.800	1.101.208	1.068.823
Alicante	9.638.835	8.855.444	9.913.731	9.382.931	9.139.479
Almería	705.514	749.720	780.853	786.877	791.837
Asturias	1.039.406	1.309.770	1.339.010	1.355.364	1.316.212
Badajoz	29.112	65.642	56.981	61.179	75.351
Barcelona	35.216.828	35.144.503	34.398.226	29.209.536	27.421.682
Bilbao	3.800.774	4.171.065	4.046.172	3.888.955	3.654.957
Burgos	18.905	21.057	35.447	33.595	27.716
El Hierro	139.153	152.726	170.225	170.968	183.891
Fuerteventura	4.258.069	4.399.023	4.948.018	4.173.590	3.738.492
Girona	2.736.867	2.844.571	3.007.977	4.863.954	5.286.970
Gran Canaria	9.770.039	9.892.067	10.538.829	9.486.035	9.155.665
Granada-Jaén	638.288	728.428	872.752	978.254	1.187.813
Ibiza	5.726.579	5.555.048	5.643.180	5.040.800	4.572.819
Jerez	811.457	913.394	1.032.493	1.043.163	1.079.616
La Gomera	24.446	19.707	32.713	32.488	34.605

La Palma	809.521	965.779	1.067.431	992.363	1.043.274
Lanzarote	5.334.599	5.168.775	5.543.744	4.938.343	4.701.669
León	30.890	51.061	85.725	93.373	95.189
Logroño	10.598	19.263	17.877	24.527	35.663
Madrid Barajas	39.735.618	45.190.528	49.671.270	49.866.113	48.437.147
Málaga	12.925.186	12.581.944	12.823.117	12.064.521	11.622.429
Melilla	289.551	315.850	286.701	292.608	293.695
Menorca	2.565.462	2.545.942	2.576.200	2.511.629	2.433.666
Murcia	1.140.813	1.181.782	1.262.597	1.349.579	1.630.684
Palma de Mallorca	22.768.032	22.666.858	22.726.707	21.117.417	21.203.041
Pamplona	159.090	190.329	238.511	291.553	335.612
Reus	971.020	937.341	1.362.683	1.419.851	1.706.615
Salamanca	15.830	22.718	37.257	43.179	53.088
San Sebastián	245.003	262.783	248.050	286.077	315.294
Santander	974.043	1.117.630	1.116.398	919.871	958.157
Santiago	2.073.147	2.194.611	2.464.330	2.172.869	1.944.068
Sevilla	3.687.714	4.292.020	4.959.359	4.224.718	4.051.392
Tenerife Norte	3.524.470	3.717.944	4.095.103	4.051.356	4.054.147
Tenerife Sur	8.701.728	8.530.729	8.656.487	7.358.986	7.108.055
Valencia	4.618.072	4.752.020	4.979.511	4.934.268	4.748.997
Valladolid	260.285	378.418	462.504	392.689	365.720
Vigo	678.720	828.725	976.152	1.093.576	1.103.285
Zaragoza	457.595	551.406	751.097	605.912	528.313

Tabla 5. Variable dependiente Pasajeros

3.3.2 POBLACIÓN

La primera pregunta que surge cuando se empieza a desarrollar un estudio como este es cuánta gente reside alrededor de un aeropuerto. Más tarde nos preguntaremos qué capacidad económica tienen, de qué medios alternativos de transporte disponen o cómo de importante es la ciudad en la que residen, entre otras cosas.

Resulta básico conocer el número bruto de personas que residen en las cercanías de un aeropuerto determinado y que, por lo tanto, son susceptibles de utilizarlo; dicho en otras palabras: potenciales clientes.

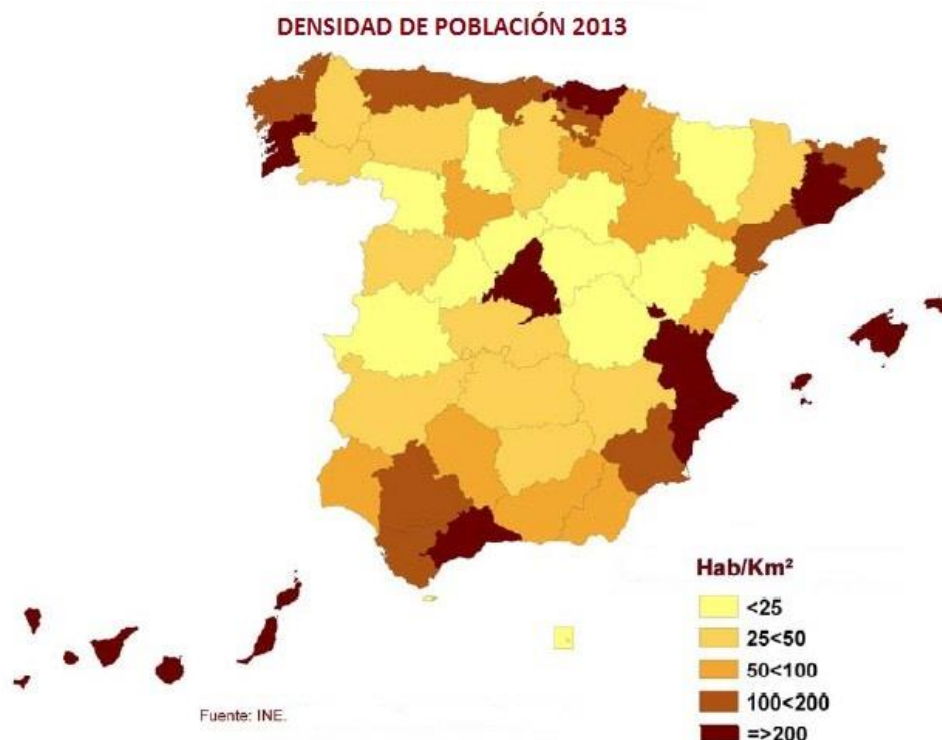


Figura 2. Densidad de población de España en 2013

En el mapa se muestra la densidad de población en habitantes por kilómetro cuadrado para las provincias españolas según los datos del Instituto Nacional de Estadística para el 2013.

El Instituto Nacional de Estadística (INE) hace públicas las cifras de población referidas al 1 de enero de cada año. En este caso se han utilizado las cifras a 1 de enero de 2014 para el análisis del año 2013, y así sucesivamente hasta el año 2009. El INE las ofrece de forma online o en forma de fichero *Excel* como detalle

Municipal en el que aparecen los datos de población totales y por sexo de cada uno de los municipios de España.

Para poder introducir en el modelo los datos de población referidos a cada aeropuerto, se ha supuesto el dato de población de cada provincia para su aeropuerto correspondiente. En el caso de aeropuertos insulares se ha tomado la población de la isla donde se ubica dicho aeropuerto. A continuación se ofrecen los datos para cada aeropuerto desde 2009 hasta 2013.

Aeropuerto	2013	2012	2011	2010	2009
A Coruña	1.133.471	1.138.494	1.123.734	1.125.354	1.123.514
Alicante	1.855.270	1.854.244	1.907.987	1.895.859	1.890.328

Almería	688.920	691.680	691.643	691.462	687.903
Asturias	1.058.976	1.067.802	1.052.707	1.055.557	1.058.116
Badajoz	689.179	690.894	679.107	677.981	676.672
Barcelona	5.440.235	5.493.078	5.357.427	5.375.778	5.352.037
Bilbao	1.139.795	1.148.871	1.130.234	1.136.355	1.136.859
Burgos	365.382	368.701	359.846	361.609	363.890
El Hierro	10.641	10.669	10.706	10.623	10.600
Fuerteventura	107.884	106.363	103.423	100.160	98.534
Girona	744.126	749.191	729.556	728.340	725.580
Gran Canaria	852.951	852.354	847.263	839.911	831.532
Granada-Jaén	921.275	922.138	913.396	911.668	907.563
Ibiza	141.174	137.929	133.702	130.705	128.718
Jerez	1.247.374	1.247.578	1.229.927	1.225.164	1.218.436
La Gomera	21.155	21.223	21.565	22.346	22.591
La Palma	83.915	84.131	83.847	83.739	83.622
Lanzarote	140.565	139.518	138.364	137.016	136.426
León	484.770	489.239	477.391	480.309	481.819
Logroño	314.829	318.639	312.206	313.149	314.009
Madrid Barajas	6.368.706	6.414.709	6.387.824	6.369.167	6.335.780
Málaga	1.620.181	1.611.983	1.614.058	1.600.000	1.580.007
Melilla	83.669	83.619	76.399	74.078	72.514
Menorca	93.650	94.280	94.230	93.651	93.067
Murcia	1.461.876	1.461.987	1.476.344	1.468.127	1.460.671
Palma de Mallorca	868.570	866.573	862.179	857.164	852.299
Pamplona	636.629	638.949	624.606	622.125	619.013
Reus	795.054	802.806	799.918	797.625	793.278
Salamanca	344.448	347.249	342.163	344.483	345.709
San Sebastián	706.774	708.207	688.079	691.221	691.169
Santander	587.686	590.037	578.901	578.632	577.995
Santiago	1.133.471	1.138.494	1.123.734	1.125.354	1.123.514
Sevilla	1.937.385	1.936.703	1.887.465	1.877.058	1.864.884
Tenerife Norte	897.879	890.975	880.760	871.081	861.864
Tenerife Sur	897.879	890.975	880.760	871.081	861.864
Valencia	2.527.615	2.547.044	2.512.926	2.516.812	2.512.609
Valladolid	528.647	531.141	526.770	525.660	524.874
Vigo	950.004	953.241	946.686	947.435	946.431
Zaragoza	969.091	972.508	954.818	951.423	948.062

Tabla 6. Variable Población

3.3.3 POBLACIÓN RESIDENTE DE LA UNIÓN EUROPEA

Además de la población autóctona referida a cada región aeroportuaria, es importante tener en cuenta los habitantes extranjeros que residen en cada una de ellas. En el modelo

estadístico se ha creído oportuno solo tener en cuenta los extranjeros residentes de la Unión Europea, ya que como se aprecia en el siguiente gráfico, más del 65 % del tráfico total de pasajeros proviene o viaja hacia países de la UE. Evidentemente muchos de esos pasajeros son españoles que viajan fuera del país por diversos motivos, pero a su vez, nos hace pensar que un amplio porcentaje del total serán extranjeros residentes; por lo que esta variable puede ser interesante para la evaluación de la demanda.

DISTRIBUCIÓN DEL TRÁFICO AÉREO POR MERCADOS

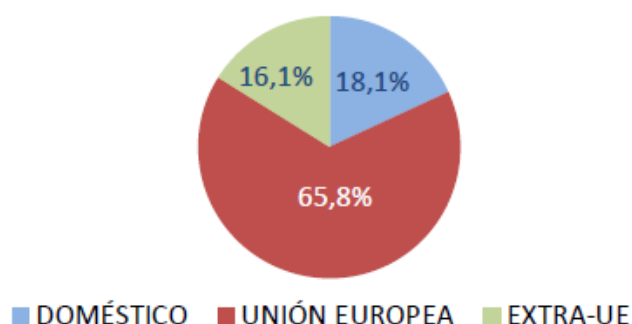


Gráfico 3. Tráfico aéreo por mercados

Los datos para el modelo han sido sacados del Instituto Nacional de Estadística, para el período del 2009 al 2013, procediendo de la misma forma que en el apartado anterior para la población total. Con la salvedad de que al no haber datos para los aeropuertos insulares, puesto que la población residente viene dada por comunidades y provincias, se ha aplicado un factor de corrección a cada isla en función de su población total, como ya se explicó con anterioridad. En la siguiente tabla podemos observar los datos para cada aeropuerto desde 2009 a 2013.

Aeropuerto	2013	2012	2011	2010	2009
A Coruña	11.304	19.671	19.658	19.038	18.532
Alicante	216.083	216.816	217.973	218.765	222.766
Almería	61.969	65.065	64.439	63.225	62.072
Asturias	18.674	25.245	25.185	24.357	23.592
Badajoz	13.014	14.879	15.039	14.617	14.354
Barcelona	152.342	148.993	151.447	150.382	154.514
Bilbao	14.395	17.791	18.763	18.635	18.532
Burgos	15.836	16.711	18.069	18.328	18.800
El Hierro	955	858	806	782	759
Fuerteventura	7.027	6.550	6.247	5.956	5.753
Girona	42.433	42.539	43.411	43.471	44.099

Gran Canaria	55.527	51.761	49.366	47.068	45.460
Granada-Jaén	29.471	33.725	33.806	33.388	32.786
Ibiza	15.169	13.792	13.426	13.115	12.952
Jerez	21.973	26.473	26.263	25.904	25.396
La Gomera	1.900	1.707	1.604	1.556	1.511
La Palma	7.529	6.764	6.357	6.164	5.985
Lanzarote	9.149	8.529	8.134	7.756	7.491
León	10.269	14.355	14.440	13.928	13.504
Logroño	14.560	15.711	16.799	16.610	16.530
Madrid Barajas	293.632	308.782	323.455	320.521	327.117
Málaga	143.958	140.789	137.487	133.771	131.151
Melilla	1.246	1.053	994	964	871
Menorca	10.069	9.155	8.912	8.706	8.598
Murcia	56.990	64.944	65.592	65.365	65.801
Palma de Mallorca	93.363	84.888	82.633	80.722	79.717
Pamplona	20.541	22.358	23.334	23.151	23.257
Reus	44.887	45.570	47.803	48.384	49.420
Salamanca	5.892	8.679	8.883	8.755	8.589
San Sebastián	10.585	13.845	14.844	15.060	15.314
Santander	11.592	13.739	14.279	13.943	13.744
Santiago	11.304	19.671	19.658	19.038	18.532
Sevilla	25.631	31.731	31.302	29.613	28.985
Tenerife Norte	80.548	72.357	68.003	65.942	64.032
Tenerife Sur	80.548	72.357	68.003	65.942	64.032
Valencia	104.956	121.978	129.658	129.426	130.338
Valladolid	15.593	17.935	18.388	18.224	18.043
Vigo	15.249	24.047	24.007	23.689	23.269
Zaragoza	57.484	56.654	58.036	57.690	57.249

Tabla 7. Variable Población residente de la UE

3.3.4 P.I.B. PER CÁPITA POR ENTORNO AEROPORTUARIO

El Producto interior bruto (PIB) per cápita, también conocido como renta per cápita o ingreso per cápita, relaciona el PIB de un país con su número de habitantes (datos en € en esta variable). Se trata de un indicador de bienestar social o calidad de vida y está estrechamente relacionado con la riqueza económica de un país. Dada su naturaleza debe ser utilizado con cautela. El propio creador del sistema estadounidense de contabilidad nacional e inventor del PIB, Simon Kuznets (1934), se mostraba muy

crítico ante la intención de medir la calidad de vida de un país exclusivamente a través del ingreso per cápita derivado del PIB.

Esta variable hace una estimación de la capacidad económica de las regiones circundantes a los aeropuertos que nos ocupan, pero no valora la calidad de vida o el bienestar social que ofrecen estas zonas.

Por tanto, siguiendo los consejos de Kuznets, además de esta variable, más adelante se explicarán otras que complementen a la renta per cápita para mejorar el modelo.

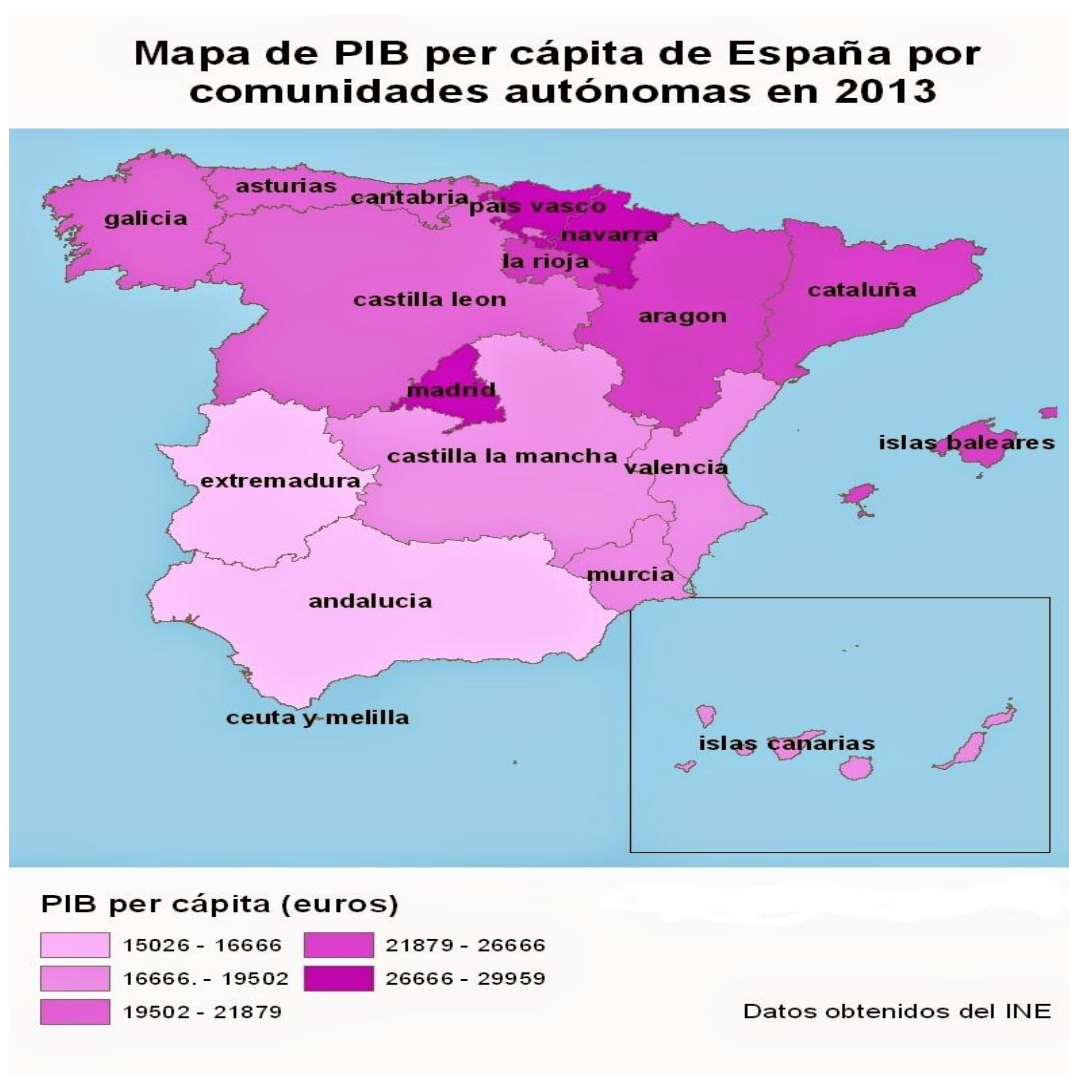


Figura 3. P.I.B. per cápita por comunidades en 2013

De nuevo se vuelven a utilizar los datos ofrecidos por el INE (Instituto Nacional de Estadística) en su fichero de contabilidad regional de España para los años 2009 (de carácter definitivo), 2010 y 2011 (estimación provisional). Para los años 2012 y 2013 (primera estimación), solo hay datos referidos a las comunidades autónomas, por lo que

para obtener los datos provinciales se hace un regla de tres respecto al año 2011, viendo su evolución y sacando los valores para cada provincia, que serán los que se asignen a cada aeropuerto. En el caso de las islas como es lógico, se aplicará el P.I.B. per cápita de la provincia a la que pertenecen.

Es especialmente interesante observar la tasa de variación interanual de los últimos años; a partir del año 2002 el PIB per cápita total en España ha ido aumentando entorno al 5,6% hasta el año 2008 que lo hizo sólo en un 1,7%. En el año 2009 la tasa de variación es del -4,5%, es decir, la renta per cápita en España se reduce en casi un 4%. Del año 2010 al 2013, y ya inmersos en plena crisis económica la tasa de variación es de aproximadamente del -2,2% (estimaciones).

Para la estimación del PIB per cápita el INE procede según el reglamento europeo; asignando el PIB correspondiente a la Extra-Regio de manera proporcional al resto de provincias y comunidades autónomas.

Este dato de renta per cápita se tiene en cuenta debido a que un poder adquisitivo alto se traduce, normalmente, en una mayor utilización del transporte aéreo. Es lógico, por lo tanto, preguntarse cómo de alta es la capacidad económica de los habitantes de las zonas que vamos a estudiar.

Quizás sería más conveniente extraer los datos de las regiones aeroportuarias o su área de influencia, en vez de la provincia, pero no se dispone de ellos; además, el error que se comete al utilizar estos datos es totalmente asumible ya que no importa no utilizar matemáticamente el número exacto, en euros, asignado a cada potencial cliente de un aeropuerto determinado.

El P.I.B per cápita medio en España es de 22.279€ (estimación 2013), y la desviación estándar (lo que suelen alejarse los valores del promedio) se sitúa en 4.233€. El valor mínimo corresponde al aeropuerto de Badajoz con 16.009€, algo por debajo de los aeropuertos de Granada y Jerez (16.391€ y 16.824€); y el valor máximo a Madrid con 30.182€, muy seguido por *San Sebastián* con 29.824€.

En la siguiente tabla podemos apreciar el P.I.B. per cápita desde el 2009 al 2013, y ver su evolución por aeropuerto.

Aeropuerto	2013	2012	2011	2010	2009
A Coruña	21.816	21.748	21.898	22.199	21.929
Alicante	17.084	17.069	17.405	18.078	18.259
Almería	16.406	16.483	16.855	17.905	18.743
Asturias	20.591	20.867	21.310	21.247	21.140
Badajoz	14.991	15.098	15.617	15.780	16.009
Barcelona	26.594	26.347	26.531	26.478	26.555
Bilbao	28.129	28.215	28.618	29.004	28.609
Burgos	26.643	26.791	27.128	26.317	25.827
El Hierro	18.756	18.822	19.205	19.340	18.891
Fuerteventura	18.983	19.051	19.438	19.350	19.558
Girona	26.785	26.537	26.722	26.694	26.425
Gran Canaria	18.983	19.051	19.438	19.350	19.558
Granada-Jaén	15.703	15.777	16.133	16.076	16.391
Ibiza	23.446	23.596	23.769	23.829	24.169
Jerez	16.465	16.543	16.916	16.938	16.824
La Gomera	18.756	18.822	19.205	19.340	18.891
La Palma	18.756	18.822	19.205	19.340	18.891
Lanzarote	18.983	19.051	19.438	19.350	19.558
León	20.318	20.431	20.688	20.658	20.791
Logroño	25.277	25.191	25.537	25.335	25.008
Madrid Barajas	28.915	28.914	29.576	29.471	30.182
Málaga	16.807	16.886	17.267	17.666	17.923
Melilla	16.426	16.708	17.824	18.275	19.115
Menorca	23.446	23.596	23.769	23.829	24.169
Murcia	17.901	18.032	18.470	18.952	19.035
Palma de Mallorca	23.446	23.596	23.769	23.829	24.169
Pamplona	28.358	28.499	29.134	28.846	28.682
Reus	26.855	26.607	26.792	26.559	25.898
Salamanca	18.920	19.024	19.264	19.115	19.956
San Sebastián	30.905	30.999	31.442	30.651	29.824
Santander	21.550	21.698	22.055	22.064	22.016
Santiago	21.816	21.748	21.898	22.199	21.929
Sevilla	17.738	17.821	18.223	18.155	18.590
Tenerife Norte	18.756	18.822	19.205	19.340	18.891
Tenerife Sur	18.756	18.822	19.205	19.340	18.891
Valencia	20.701	20.683	21.091	21.246	21.268
Valladolid	23.744	23.875	24.176	24.633	23.682
Vigo	19.474	19.414	19.548	19.847	20.142
Zaragoza	24.568	24.647	25.150	25.352	25.242

Tabla 8. Variable P.I.B. per cápita

3.3.5 TASA DE PARO

La Encuesta de Población Activa (EPA) es una investigación llevada a cabo por el INE de forma trimestral para obtener datos del mercado de trabajo (ambos sexos). De dicha encuesta se deducen el porcentaje de personas activas, ocupadas y paradas (tasa de paro) que hay en España y en sus distintas Comunidades Autónomas y provincias.

Se define como parados a aquellos que tienen 16 años o más, que están sin trabajo, disponibles para trabajar y buscando empleo activamente.

Debido a la crisis económica, la tasa de paro en todo el país ha aumentado considerablemente en los últimos años, aunque como se observa en la imagen de abajo, hay una fuerte disparidad norte-sur.

La tasa de paro media en España, para los años de interés del presente estudio es: 13,91% en 2009, 20,05% en 2010, 21,29% en 2011, 24,63% en 2012 y 26,03 para 2013.

De nuevo, para esta variable, a cada aeropuerto se le ha asignado la tasa de paro de la provincia a la que pertenece para dicha época.

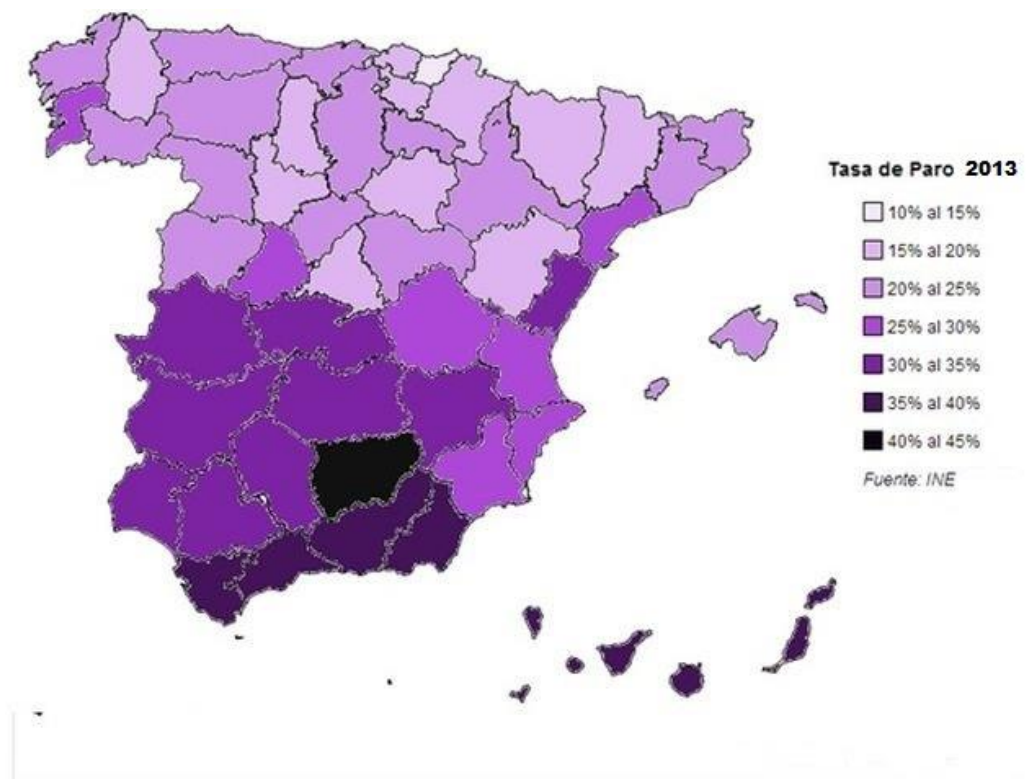


Figura 4. Tasa de paro por provincias en 2013

Aeropuerto	2013	2012	2011	2010	2009
A Coruña	20,06	18,64	15,61	13,92	9,45
Alicante	30,1	27,52	23,79	23,46	16,04
Almería	36,25	35,32	32,04	27,79	24,98
Asturias	22,25	21,03	18,2	16,56	9,95
Badajoz	32,72	33,73	27,22	24,36	19,22
Barcelona	21,74	21,73	18,95	17,74	11,45
Bilbao	17,03	15,77	13,3	11,44	9,57
Burgos	18,73	17,91	16,09	13,61	10,46
El Hierro	33,01	29,79	25,91	26,13	19,65
Fuerteventura	33,32	36,33	31,01	29,12	22,51
Girona	24,08	25,78	22,33	20,47	14,14
Gran Canaria	33,32	36,33	31,01	29,12	22,51
Granada-Jaén	35,99	36,49	29,24	26,89	23,14
Ibiza	22,91	21,27	25,36	22,41	12,32
Jerez	40,77	34,51	32,68	31,87	21,96
La Gomera	33,01	29,79	25,91	26,13	19,65
La Palma	33,01	29,79	25,91	26,13	19,65
Lanzarote	33,32	36,33	31,01	29,12	22,51
León	24,15	20,8	17,8	17,47	10,52
Logroño	20,05	22,71	15,44	15,23	9,75
Madrid Barajas	21,03	18,86	15,43	16,17	10,15
Málaga	36,2	34,45	32,73	30,02	23,6
Melilla	35,36	30,27	22,48	21,02	16,87
Menorca	22,91	21,27	25,36	22,41	12,32
Murcia	28,98	26,19	26,16	23,17	15,53
Palma de Mallorca	22,91	21,27	25,36	22,41	12,32
Pamplona	16,83	16,41	13,4	12,32	8,12
Reus	27,03	21,53	19,98	18,38	13,73
Salamanca	23,08	19,06	16,18	15,8	16,49
San Sebastián	12,75	12,5	8,98	8,95	6,68
Santander	20,06	17,38	16,37	14,49	8,9
Santiago	20,06	18,64	15,61	13,92	9,45
Sevilla	34,9	30,94	26,52	25,81	19,72
Tenerife Norte	33,01	29,79	25,91	26,13	19,65
Tenerife Sur	33,01	29,79	25,91	26,13	19,65
Valencia	26,06	26,58	23,68	22,28	14,03
Valladolid	22,38	18,9	15,77	15,36	9,65
Vigo	24,77	25,58	20,71	18,39	12,04
Zaragoza	20,83	19,74	19,55	16,19	10,54

Tabla 9. Variable Tasa de paro

3.3.6 ÍNDICE COMERCIAL

Índice comparativo de la importancia del comercio (comercio mayorista y minorista). Este índice se elabora en función del impuesto de actividades económicas (IAE) correspondiente a las actividades del comercio mayorista y comercio minorista conjuntamente. El valor del índice refleja el peso relativo (en tanto por cien mil) de la actividad comercial de un municipio, provincia o comunidad autónoma respecto al total de España, con base: total euros de recaudación del impuesto en España = 100.000 unidades.

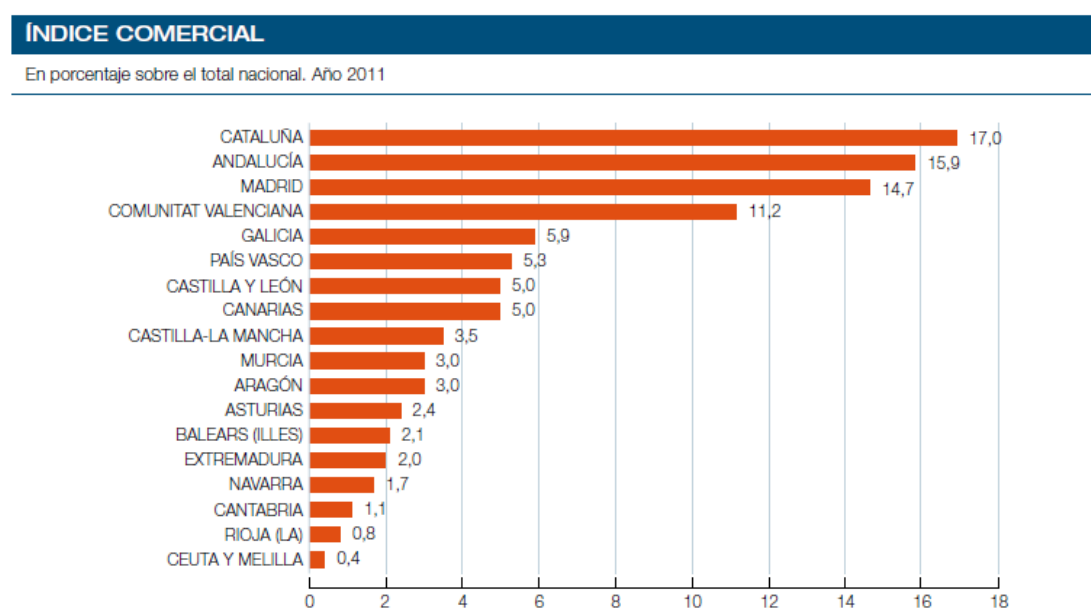


Gráfico 4. Porcentaje del índice comercial por comunidades autónomas

3.3.7 ÍNDICE DE RESTAURACIÓN Y BARES

Índice comparativo de la importancia de la restauración y bares (restaurantes, cafeterías, bares, etc.). Este índice se elabora en función del impuesto de actividades económicas (IAE) correspondiente a las actividades de restauración y bares. El valor del índice refleja el peso relativo (en tanto por cien mil) de la actividad de restauración y bares de un municipio, provincia o comunidad autónoma respecto al total de España, con base: total euros de recaudación del impuesto en España = 100.000 unidades.

ÍNDICE DE RESTAURACIÓN Y BARES

En porcentaje sobre el total nacional. Año 2011

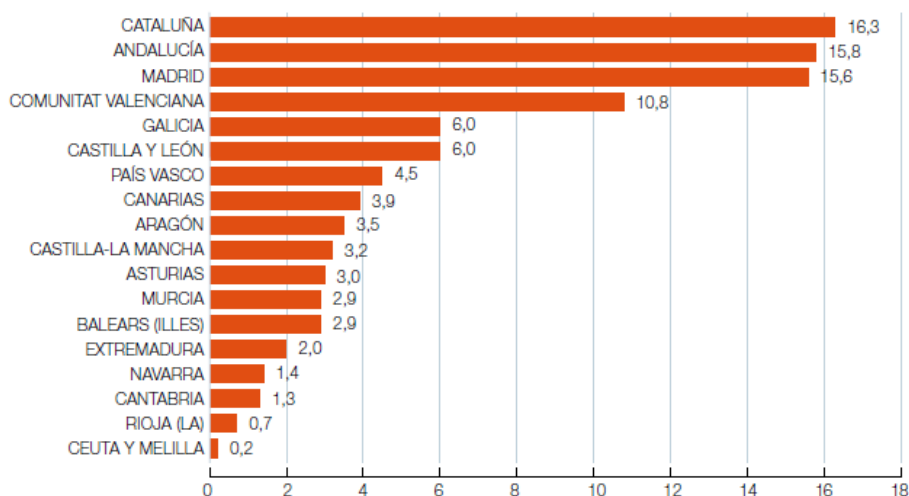


Gráfico 5. Porcentaje del índice de restauración y bares por comunidades autónomas

3.3.8 ÍNDICE TURÍSTICO

Índice comparativo de la importancia turística. Se obtiene en función de la cuota o impuesto de actividades económicas correspondiente a las actividades turísticas, el cual se basa a su vez en la categoría de los establecimientos turísticos (hoteles y moteles, hoteles-apartamentos, hostales y pensiones, fondas y casas de huéspedes, campings y apartamentos gestionados por empresas), número de habitaciones y ocupación anual (todo el año o parte del año); por lo que constituye prácticamente un indicador de la oferta turística.

El valor del índice indica la participación (en tanto por 100.000) que corresponde a cada municipio, provincia o comunidad autónoma sobre una base nacional de 100.000 unidades (total euros recaudación del impuesto = 100.000).

ÍNDICE TURÍSTICO

En porcentaje sobre el total nacional. Año 2011

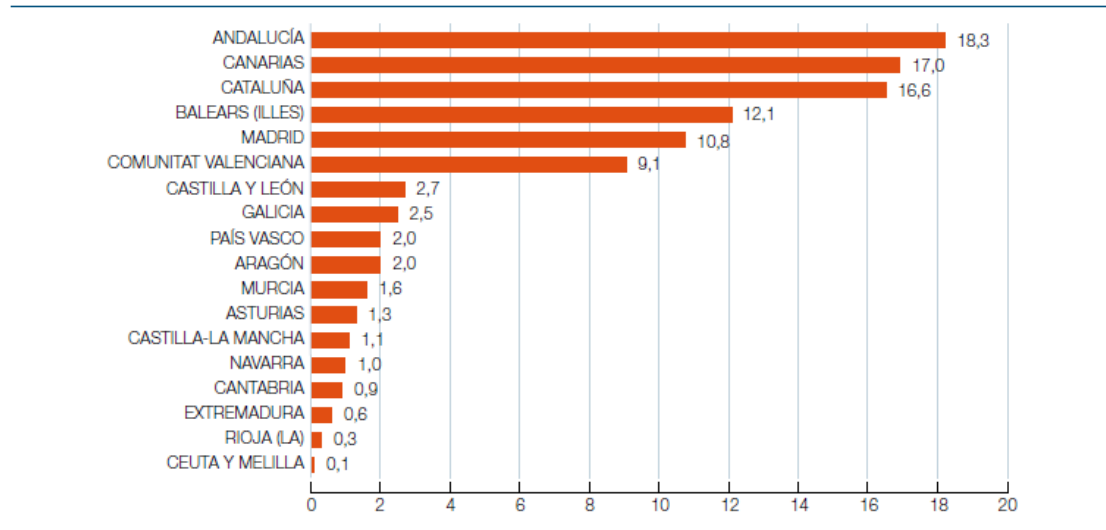


Gráfico 6. Porcentaje del índice turístico por comunidades autónomas

3.3.9 ÍNDICE DE ACTIVIDAD ECONÓMICA

Índice comparativo del conjunto de la actividad económica. Se obtiene en función del impuesto correspondiente al total de actividades económicas empresariales (industriales, comerciales y de servicios) y profesionales. Es decir, incluye todas las actividades económicas excepto las agrarias (que no están sujetas al impuesto de actividades económicas).

El valor del índice expresa la participación de la actividad económica (en tanto por 100.000) de cada municipio, provincia o comunidad autónoma sobre una base nacional de 100.000 unidades equivalente a la recaudación del impuesto de actividades económicas empresariales y profesionales.

El índice de actividad económica está en general correlacionado con la cuota de mercado (variable que se estudiará más adelante). Es decir, suele existir bastante correlación entre la capacidad de compra y los impuestos correspondientes a las actividades económicas empresariales y profesionales. Sin embargo, se observan a veces diferencias, algunas importantes, entre ambos indicadores.

ÍNDICE DE ACTIVIDAD ECONÓMICA

En porcentaje sobre el total nacional. Año 2011

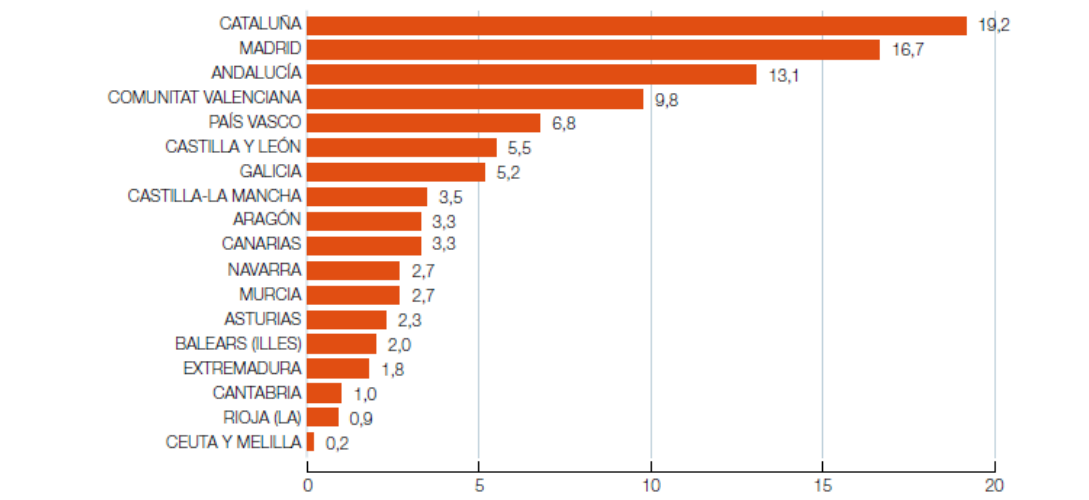


Gráfico 7. Porcentaje del índice de actividad económica

3.3.10 ÍNDICE INDUSTRIAL

Índice comparativo de la importancia de la industria (incluida la construcción).

Este índice se elabora en función del impuesto de actividades económicas (IAE) correspondiente a las actividades industriales. El valor del índice refleja el peso relativo (en tanto por cien mil) de la industria de un municipio, provincia o comunidad autónoma respecto al total de España, con base: total euros de recaudación del impuesto en España = 100.000 unidades.

ÍNDICE INDUSTRIAL

En porcentaje sobre el total nacional. Año 2011

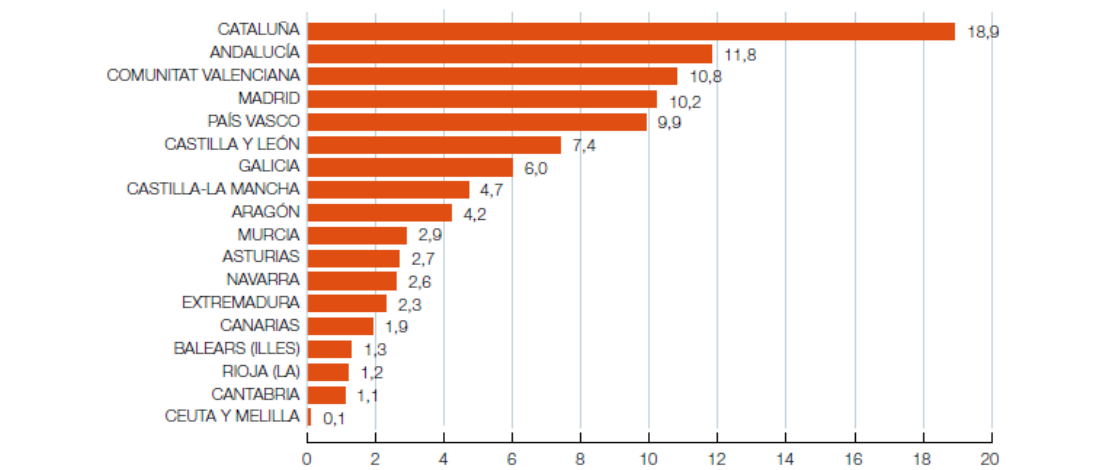


Gráfico 8. Porcentaje del índice industrial por comunidades autónomas

3.3.11 DELIMITACIÓN DE LAS ÁREAS Y SUBÁREAS COMERCIALES EN EL ANUARIO ECONÓMICO DE LA CAIXA, Y DATOS TABULADOS DE CADA ÍNDICE SEGÚN EL AÑO

La metodología aplicada a la determinación de las áreas y subáreas comerciales se basa principalmente en modelos de gravitación comercial y encuestas a los municipios.

Modelización de las gravitaciones comerciales

Se han aplicado modelos de gravitación comercial en numerosos municipios, principalmente en los periféricos o limítrofes con otras áreas y subáreas comerciales, para delimitar las cabeceras de área o subárea por las que son atraídos. Los modelos empleados dependen de la disponibilidad de información y de las características o grado de dificultad existente, siendo los más utilizados los de Reilly y Huff.

Modelo de Reilly

Se expone a continuación el modelo de gravitación comercial de Converse, derivado del modelo de Reilly, que se basa a su vez, en la ley de gravitación universal de Newton.

$$D_{ab} = \frac{d}{1 + \sqrt{\frac{P_b}{P_a}}}$$

donde:

D_{ab} = Límite del área o subárea comercial de la localidad o municipio "a" (cabecera de área o subárea), medido en kilómetros a lo largo de la carretera o vía de comunicación que conduce a la localidad "b" (cabecera de otra área o subárea). Dicha distancia es, por tanto, el "punto de indiferencia" entre dos municipios (a y b) en el cual los consumidores serían indiferentes a comprar o gravitar comercialmente en cualquiera de dichos dos municipios.

Los municipios situados entre "a" y el "punto de indiferencia" gravitan comercialmente sobre "a", mientras que los situados entre el "punto de indiferencia" y "b" gravitan sobre "b".

d = Distancia en kilómetros a lo largo de la carretera más importante entre "a" y "b". Se sustituye por el "tiempo de viaje" cuando un municipio situado entre "a" y "b" está

comunicado con una y otra cabecera de área por carreteras de distinta categoría (por ejemplo, una carretera normal y una autovía).

Pa = Población del municipio "a".

Pb = Población del municipio "b".

El modelo introduce la variable "equipamiento comercial de grandes superficies y cadenas" (grandes almacenes, hipermercados, centros comerciales y cadenas comerciales importantes), en los casos donde no existe una gran correlación entre las variables "población" y "equipamiento comercial de grandes superficies y cadenas"; es decir, cuando la población de una localidad "a" puede ser menor que la de la localidad "b" pero ser mayor su equipamiento comercial.

Modelo de Huff

Se aplica el modelo de Huff, principalmente en aquellos casos en que existe duda sobre la gravitación de un municipio sobre más de dos cabeceras (de área o subárea).

El modelo de Huff se puede expresar de la forma siguiente:

$$P_{ij} = \frac{S_j^\alpha \cdot T_{ij}^{-\beta}}{\sum_{k=1}^n S_k^\alpha \cdot T_{ik}^{-\beta}}$$

donde:

Pij = Probabilidad de que un consumidor de un municipio *i* (punto de origen) se desplace a un municipio *j* determinado (punto de destino).

Sj = Metros cuadrados de superficie de venta de las grandes superficies (centros comerciales, grandes almacenes, hipermercados, grandes superficies especializadas) y cadenas de tiendas especializadas ubicadas en el municipio *j*.

Sk = Metros cuadrados de superficie de venta de las grandes superficies y cadenas ubicadas en el municipio *k*.

Tij = Tiempo empleado en el desplazamiento para llegar de la base *i* del consumidor al municipio *j*.

Tik = Tiempo empleado en el desplazamiento para llegar de la base *i* del consumidor al municipio *k*.

α = Parámetro que refleja la sensibilidad del cliente al equipamiento comercial existente (metros cuadrados de grandes superficies y cadenas).

β = Parámetro que refleja la sensibilidad del cliente al tiempo de desplazamiento.

n = Número de municipios considerados por el consumidor como otros posibles centros de atracción, además del municipio j .

Los parámetros α y β reflejan la sensibilidad del cliente a la atracción de la oferta comercial y al tiempo de viaje, respectivamente. Estos parámetros se calculan en base a datos empíricos, resultando generalmente: $\alpha = 1$ y $\beta = 2$.

Otros modelos de gravitación comercial

Los modelos gravitatorios son el fundamento de los llamados Modelos de Interacción Espacial, cuyo objetivo es modelizar todo movimiento o comunicación sobre el espacio resultante de un proceso de decisión; esto implica un origen, un destino y el movimiento resultante de la elección que hace el origen del destino. Todos estos modelos tienen muchos campos de aplicación: migraciones, desplazamientos para comprar o por causa de servicios administrativos, sanitarios, financieros, ocio, etc.

Se trata de modelos, todos ellos, muy útiles para explicar el comportamiento de los individuos ante situaciones de elección. Tal es el caso del Modelo Multiplicativo de Interacción Competitiva, de Nakanishi y Cooper (1974), del Modelo Logit Multinomial de McFadden (1974) (que es una de las generalizaciones posibles del Modelo gravitatorio de Huff) y del Modelo Logit de Destinos Competitivos de Fotheringham y O'Kelly (1989), entre otros.

Sin embargo, los modelos más empleados han sido los de Reilly y Huff. Estos modelos sólo tienen en cuenta dos variables: la variable "masa" de atracción comercial (equivalente en el Anuario a la variable "superficie de venta") y la variable "fricción", que suele ser la distancia entre el punto de origen y el de destino (o también, el tiempo de viaje) y no consideran otras variables que pueden ser importantes en la determinación de los flujos comerciales, como son las relativas al punto de origen (nivel económico, edades de la población, etc.), que sí recogen los Modelos Logit.

En definitiva, se quiere destacar que no todos los modelos son igualmente aplicables en cada caso. Según nuestro análisis y para situaciones de grandes agregados, como es el caso de las áreas comerciales (donde el punto de origen está constituido por el conjunto de consumidores de un municipio y el punto de destino es el conjunto de todos los centros comerciales del municipio cabecera de área), los modelos de gravitación comercial de Reilly y Huff resultan más adecuados que los Modelos Logit; más eficaces estos últimos, en cambio, en situaciones de menor agregación (donde, por ejemplo, se

consideran gravitaciones de consumidores o familias sobre un centro comercial individual).

Aplicación del método de encuesta

La modelización de las gravitaciones comerciales que se acaba de explicar se complementa con la realización de una encuesta telefónica dirigida a los ayuntamientos de los municipios, principalmente a sus secretarios, y a los directores de las OCU's municipales (Oficinas de Consumidores y Usuarios).

Dicha encuesta se basa en un cuestionario de pequeña extensión y debidamente estructurado con objeto de averiguar si los consumidores del municipio encuestado se desplazan con alguna habitualidad a otros municipios para realizar las compras más importantes, la frecuencia de los desplazamientos y los tipos de comercios que visitan.

La encuesta se lleva a cabo en bastantes municipios, exceptuándose aquéllos cuya gravitación comercial resulta casi evidente, o coherente con otros municipios similares, después de la aplicación de modelos de gravitación comercial.

Actualización anual de las gravitaciones comerciales

La actualización anual de las gravitaciones comerciales de los consumidores no requiere la aplicación de la metodología expuesta anteriormente a un gran número de municipios. La investigación se centra en aquellos municipios cuya población ha podido modificar sus hábitos de desplazamiento para realizar sus compras más importantes (flujos comerciales) como consecuencia de variaciones en factores como: vías de comunicación, equipamientos comerciales, etc. Es decir, hay que identificar y analizar dichas variaciones y aplicar posteriormente modelos de gravitación y encuestas a los municipios afectados.

En los últimos años, se están originando importantes cambios en las gravitaciones comerciales de los consumidores que, si bien apenas afectan al número de áreas comerciales existentes (varían poco los municipios que son núcleos de atracción comercial cabeceras de área), sí que se crean nuevas subáreas comerciales, principalmente en las zonas costeras turísticas, y se producen sensibles modificaciones en las estructuras de las áreas y subáreas comerciales existentes: variación de los límites geográficos, aumento del número de gravitaciones comerciales múltiples (compartidas), etc. El desarrollo progresivo de las ofertas comerciales y de ocio, así como de las vías de comunicación, que se está extendiendo a municipios relativamente pequeños que

hasta hace pocos años no tenían interés económico, están revolucionando los hábitos de compra de los consumidores en lo que respecta a sus desplazamientos o gravitaciones comerciales a determinados municipios para realizar las compras de cierta importancia.

NOTAS

- Los datos para el modelo han sido sacados del Anuario Económico de la Caixa, para el período del 2009 al 2013; basado en las fuentes del Ministerio de Economía y Hacienda, Diputaciones Forales del País Vasco, Gobierno de Navarra y elaboración propia de la Caixa.

Al igual que en otras variables se ha asignado el valor de cada aeropuerto al de la provincia a la que pertenece. Puesto que los datos de los índices vienen dados por comunidades y provincias, de nuevo a los aeropuertos insulares se les ha aplicado un factor de corrección en función de su población total, como ya se explicó con anterioridad.

- Todos los antes expuestos son índices simples que se ha obtenido aplicando la siguiente fórmula:

$$\text{Índice} = \frac{\text{Cuota de la comunidad / provincia / municipio}}{\text{Total cuotas España}} \times 100.000$$

- Todos los índices están referidos a dos años antes del año descrito (2011 para 2013, 2010 para 2012, etc.).

- El IAE es un tributo que grava el mero ejercicio, en territorio nacional, de actividades empresariales, profesionales y artísticas. La cuota o impuesto se establece atendiendo a diferentes variables, con independencia de que la actividad genere beneficios con los que pagar los impuestos. La recaudación se lleva a cabo en los ayuntamientos (extensible a las diputaciones provinciales a través del recargo provincial). Debe advertirse que, a efectos de elaboración de los índices basados en la recaudación del IAE (actividad económica, industrial, comercial, comercial mayorista, comercial minorista, restauración y bares, y turístico), se ha partido de la "cuota de tarifa", no de la recaudación total (cuota total). Con la "cuota de tarifa" se trata de corregir el hecho de que exista (se produzca) distinta tributación ante idénticas actividades en distintos municipios, por mor del uso de índices correctores y de situación que tienen atribuidos

los ayuntamientos. De esta forma, elaborando dichos índices según la cuantía de la cuota de tarifa en vez de la cuota total, se pueden realizar análisis comparativos homogéneos entre municipios sobre la importancia de la industria, comercio, oferta turística, etc.

- Por otra parte, debe advertirse que, a partir del 1 de enero de 2003, están exentos del pago de la cuota del IAE todas las personas físicas y aquellas sociedades cuya cifra de negocios sea inferior a un millón de euros en el ejercicio anterior, teniendo en cuenta para calcular este importe el conjunto de actividades económicas que ejerza el contribuyente. En las provincias de Guipúzcoa y Vizcaya, la exención alcanza a los contribuyentes cuya cifra de negocios sea inferior a dos millones de euros. En Navarra, no se produjo exención en el referido año 2003 ni en los años siguientes (2004 a 2010) ("el impuesto satisfecho tendrá la consideración de pago a cuenta del IRPF o del Impuesto sobre Sociedades"). La información estadística sobre número de actividades y superficie no ha sido afectada por dichas exenciones del pago de cuota, pero sí los datos sobre cuota, que sólo se ha facilitado para el número de actividades no exentas del pago de cuota, por lo que han tenido que ser estimados (estimación que se ha visto reforzada con el cálculo realizado por las fuentes).

A continuación se representan en tablas, para cada año del estudio, todos los índices según el aeropuerto.

AÑO 2009

Aeropuerto	Índice comercial	Índice de restauración y bares	Índice turístico	Índice de actividad económica	Índice industrial
A Coruña	2.540	2.869	1.350	2.337	2.436
Alicante	4.206	4.526	4.995	3.409	3.453
Almería	1.198	1.254	1.988	1.186	1.036
Asturias	2.408	3.113	1.386	1.971	2.702
Badajoz	1.372	1.150	424	980	799
Barcelona	13.481	12.184	9.374	14.885	14.061
Bilbao	2.294	2.131	744	3.240	3.908
Burgos	774	841	476	998	1.650
El Hierro	23	21	82	16	11
Fuerteventura	267	198	893	179	112
Girona	1.387	1.289	3.015	1.653	1.640
Gran Canaria	2.113	1.566	7.054	1.415	885

Granada-Jaén	1.584	1.723	1.630	1.179	773
Ibiza	268	374	1.631	293	190
Jerez	2.340	2.375	2.951	2.184	2.086
La Gomera	46	42	162	33	22
La Palma	183	165	643	129	88
Lanzarote	348	258	1.162	233	146
León	915	1.235	570	1.008	1.251
Logroño	760	742	317	754	1.076
Madrid Barajas	14.462	15.248	10.378	17.044	10.787
Málaga	3.679	3.724	7.228	2.898	1.811
Melilla	217	108	49	89	63
Menorca	178	248	1.083	195	126
Murcia	3.032	3.038	1.542	2.506	2.827
Palma de Mallorca	1.650	2.300	10.041	1.806	1.168
Pamplona	1.628	1.310	824	2.328	2.770
Reus	1.478	1.620	2.759	1.983	3.020
Salamanca	689	960	600	797	1.006
San Sebastián	1.539	1.155	649	1.790	3.199
Santander	1.098	1.309	1.048	1.089	1.126
Santiago	2.540	2.869	1.350	2.337	2.436
Sevilla	4.005	3.510	2.833	3.100	2.237
Tenerife Norte	1.957	1.760	6.881	1.385	943
Tenerife Sur	1.957	1.760	6.881	1.385	943
Valencia	5.855	5.376	2.917	4.974	5.044
Valladolid	1.047	1.320	504	1.160	1.271
Vigo	2.065	2.024	1.074	1.672	1.704
Zaragoza	2.251	2.828	975	2.335	2.731

Tabla 10. Variables Índices 2009

AÑO 2010

Aeropuerto	Índice comercial	Índice de restauración y bares	Índice turístico	Índice de actividad económica	Índice industrial
A Coruña	2.499	2.849	1.299	2.258	2.443
Alicante	4.157	4.539	5.175	3.276	3.343
Almería	1.183	1.258	2.039	1.131	1.036
Asturias	2.471	3.102	1.352	2.145	2.691
Badajoz	1.346	1.147	420	974	846
Barcelona	13.556	12.233	10.204	14.907	13.755
Bilbao	2.386	2.140	699	3.276	4.025

Burgos	766	837	454	983	1.678
El Hierro	23	21	80	16	11
Fuerteventura	268	200	882	176	107
Girona	1.385	1.299	2.793	1.587	1.644
Gran Canaria	2.121	1.577	6.966	1.395	845
Granada-Jaén	1.567	1.673	1.790	1.137	767
Ibiza	273	376	1.528	277	175
Jerez	2.282	2.363	2.856	2.126	2.189
La Gomera	46	42	159	31	21
La Palma	183	165	631	125	83
Lanzarote	349	260	1.148	230	139
León	944	1.206	540	984	1.206
Logroño	752	745	312	802	1.157
Madrid Barajas	14.558	15.283	10.769	17.061	10.544
Málaga	3.708	3.736	7.088	3.100	1.940
Melilla	205	109	50	86	60
Menorca	181	250	1.014	184	116
Murcia	3.003	3.047	1.465	2.704	2.800
Palma de Mallorca	1.680	2.314	9.405	1.707	1.075
Pamplona	1.680	1.318	853	2.476	2.817
Reus	1.462	1.621	2.824	2.013	3.144
Salamanca	689	932	572	755	937
San Sebastián	1.568	1.167	638	1.919	3.381
Santander	1.121	1.308	1.029	1.074	1.110
Santiago	2.499	2.849	1.299	2.258	2.443
Sevilla	3.899	3.528	2.796	2.978	2.071
Tenerife Norte	1.957	1.769	6.749	1.333	890
Tenerife Sur	1.957	1.769	6.749	1.333	890
Valencia	5.812	5.392	2.984	4.813	4.954
Valladolid	1.073	1.275	537	1.109	1.252
Vigo	2.040	2.023	1.034	1.659	1.667
Zaragoza	2.282	2.839	1.066	2.470	2.805

Tabla 11. Variables Índices 2010

AÑO 2011

Aeropuerto	Índice comercial	Índice de restauración y bares	Índice turístico	Índice de actividad económica	Índice industrial
A Coruña	2.517	2.821	1.251	2.261	2.445
Alicante	4.127	4.500	4.969	3.377	3.306
Almería	1.170	1.241	2.078	1.094	1.026
Asturias	2.456	3.116	1.346	2.296	2.765

Badajoz	1.308	1.129	414	969	843
Barcelona	13.464	12.356	10.148	14.548	13.418
Bilbao	2.447	2.243	714	3.381	4.242
Burgos	767	828	394	1.025	1.725
El Hierro	23	20	83	16	10
Fuerteventura	270	192	902	181	107
Girona	1.386	1.241	2.953	1.507	1.588
Gran Canaria	2.130	1.521	7.128	1.429	842
Granada-Jaén	1.541	1.652	1.818	1.122	759
Ibiza	270	367	1.524	256	172
Jerez	2.282	2.357	2.753	2.179	2.241
La Gomera	46	41	166	32	21
La Palma	182	160	656	125	81
Lanzarote	351	251	1.175	235	139
León	954	1.215	532	994	1.217
Logroño	752	749	311	853	1.184
Madrid Barajas	14.797	15.598	10.833	16.730	10.470
Málaga	3.707	3.706	6.882	3.106	1.975
Melilla	193	109	49	84	58
Menorca	179	244	1.012	170	114
Murcia	3.009	3.025	1.553	2.869	2.777
Palma de Mallorca	1.661	2.262	9.380	1.575	1.057
Pamplona	1.617	1.349	889	2.623	2.695
Reus	1.439	1.571	2.861	2.065	3.201
Salamanca	769	934	510	758	908
San Sebastián	1.609	1.214	669	1.955	3.470
Santander	1.114	1.300	954	1.032	1.105
Santiago	2.517	2.821	1.251	2.261	2.445
Sevilla	3.871	3.519	2.920	2.863	2.020
Tenerife Norte	1.947	1.717	7.016	1.342	872
Tenerife Sur	1.947	1.717	7.016	1.342	872
Valencia	5.797	5.377	3.095	4.797	4.772
Valladolid	1.088	1.292	548	1.116	1.272
Vigo	2.032	1.988	945	1.625	1.699
Zaragoza	2.293	2.903	1.083	2.438	2.862

Tabla 12. Variables Índices 2011

AÑO 2012

Aeropuerto	Índice comercial	Índice de restauración y bares	Índice turístico	Índice de actividad económica	Índice industrial
A Coruña	2.534	2.755	1.247	2.270	2.482

Alicante	4.127	4.394	5.028	3.488	3.247
Almería	1.162	1.254	2.019	1.116	1.016
Asturias	2.437	3.042	1.281	2.289	2.758
Badajoz	1.309	1.103	358	975	895
Barcelona	13.290	12.591	10.260	14.584	13.018
Bilbao	2.453	2.357	712	3.428	4.339
Burgos	772	847	389	1.055	1.789
El Hierro	23	21	83	16	10
Fuerteventura	272	195	901	178	101
Girona	1.404	1.261	3.012	1.487	1.571
Gran Canaria	2.147	1.538	7.118	1.410	800
Granada-Jaén	1.526	1.672	1.825	1.096	761
Ibiza	269	372	1.541	257	165
Jerez	2.280	2.399	2.711	2.125	2.238
La Gomera	46	41	165	31	20
La Palma	181	163	653	123	78
Lanzarote	354	253	1.173	232	132
León	972	1.245	537	1.009	1.271
Logroño	752	733	306	847	1.189
Madrid Barajas	14.738	15.506	10.863	16.766	10.263
Málaga	3.639	3.766	6.881	3.019	2.034
Melilla	190	107	49	84	54
Menorca	178	247	1.023	170	109
Murcia	3.010	2.923	1.544	2.641	2.865
Palma de Mallorca	1.654	2.292	9.483	1.580	1.014
Pamplona	1.668	1.397	933	2.675	2.561
Reus	1.456	1.594	2.696	2.093	3.242
Salamanca	769	958	502	769	958
San Sebastián	1.675	1.260	675	2.034	3.521
Santander	1.120	1.282	931	1.022	1.082
Santiago	2.534	2.755	1.247	2.270	2.482
Sevilla	3.842	3.569	2.861	2.794	2.028
Tenerife Norte	1.938	1.749	6.990	1.318	831
Tenerife Sur	1.938	1.749	6.990	1.318	831
Valencia	5.780	5.203	3.009	4.727	4.791
Valladolid	1.098	1.325	555	1.096	1.287
Vigo	2.022	1.916	912	1.619	1.696
Zaragoza	2.340	2.843	1.314	2.431	2.881

Tabla 13. Variables Índices 2012

AÑO 2013

Aeropuerto	Índice comercial	Índice de restauración y bares	Índice turístico	Índice de actividad económica	Índice industrial
A Coruña	2.595	2.750	1.246	2.260	2.470
Alicante	4.165	4.385	5.075	3.426	3.207
Almería	1.209	1.242	1.931	1.107	1.027
Asturias	2.425	3.037	1.276	2.282	2.732
Badajoz	1.300	1.100	359	962	911
Barcelona	13.172	12.581	10.251	14.562	12.841
Bilbao	2.483	2.403	696	3.461	4.453
Burgos	772	850	371	1.073	1.822
El Hierro	23	21	84	16	10
Fuerteventura	273	193	882	178	98
Girona	1.394	1.254	3.107	1.467	1.530
Gran Canaria	2.157	1.526	6.970	1.406	777
Granada-Jaén	1.504	1.659	1.763	1.081	764
Ibiza	270	370	1.547	256	166
Jerez	2.225	2.393	2.775	2.110	2.251
La Gomera	45	41	167	31	19
La Palma	179	163	661	123	77
Lanzarote	355	251	1.149	232	128
León	986	1.250	509	1.012	1.284
Logroño	748	732	300	856	1.178
Madrid Barajas	14.739	15.607	10.836	16.717	10.199
Málaga	3.544	3.752	6.967	2.996	2.048
Melilla	189	107	48	80	54
Menorca	179	245	1.027	170	110
Murcia	3.022	2.905	1.628	2.662	2.844
Palma de Mallorca	1.663	2.275	9.524	1.574	1.023
Pamplona	1.725	1.444	968	2.739	2.603
Reus	1.507	1.584	2.732	2.110	3.270
Salamanca	761	963	471	774	972
San Sebastián	1.708	1.283	708	2.063	3.574
Santander	1.112	1.285	883	1.023	1.086
Santiago	2.595	2.750	1.246	2.260	2.470
Sevilla	3.855	3.549	2.777	2.828	2.037
Tenerife Norte	1.911	1.742	7.068	1.311	823
Tenerife Sur	1.911	1.742	7.068	1.311	823
Valencia	5.747	5.172	3.002	4.696	4.778
Valladolid	1.102	1.330	557	1.103	1.269
Vigo	2.004	1.901	894	1.618	1.689
Zaragoza	2.370	2.843	1.375	2.433	2.905

Tabla 14. Variables Índices 2013

3.3.12 PERNOCTACIONES

Los datos estadísticos sobre pernoctaciones turísticas por provincias y comunidades autónomas se presentan desglosados en: pernoctaciones turísticas regladas y pernoctaciones turísticas no regladas.

La importancia relativa del turismo entre comunidades autónomas, provincias y municipios se expresa en el Anuario Económico de la Caixa a través del índice turístico, que ya hemos explicado. Dicho índice se ha estimado a partir de la cuota turística del IAE, cuyos elementos tributarios son el número de habitaciones, ocupación (abierto o no todo el año) y categoría de los establecimientos turísticos controlados a efectos fiscales, constituyendo, por tanto, un buen indicador de la oferta turística reglada o legalizada.

El índice no tiene en cuenta la actividad turística que se canaliza a través de lugares no controlados o reglados, tales como apartamentos alquilados y campings no legalizados, apartamentos y viviendas particulares, etc., pese a su indudable importancia, puesta claramente de manifiesto en las estimaciones realizadas sobre el número de pernoctaciones turísticas (demanda turística) en este tipo de alojamientos. Se estima que, a nivel nacional, las pernoctaciones no regladas representan el 68% del total de pernoctaciones realmente existentes; o, lo que es lo mismo, las pernoctaciones no regladas son alrededor del doble que las regladas (2,1 veces superiores).

Por ese motivo, el objetivo de este apartado es divulgar las cifras obtenidas en la estimación para los diversos años de las pernoctaciones turísticas totales (regladas más no regladas) por provincias.

Cálculo de las pernoctaciones provinciales regladas

Son las pernoctaciones (de residentes de España y de residentes en el extranjero) que se producen en alojamientos turísticos reglados (desde el punto de vista fiscal) y controlados estadísticamente por el INE a través de sus cuatro encuestas: Encuesta de Ocupación Hotelera, Encuesta de Ocupación en Acampamentos Turísticos, Encuesta de Ocupación en Apartamentos Turísticos y Encuesta de Ocupación en Alojamientos de Turismo Rural.

Cálculo de las pernoctaciones provinciales no regladas

Son las pernoctaciones que se producen en lugares no controlados o reglados, tales como apartamentos no legalizados y viviendas particulares. Los datos de la serie se estiman mediante dos procedimientos distintos, según que los turistas sean extranjeros o españoles.

Las pernoctaciones no regladas de españoles se han estimado aplicando los coeficientes que relacionan las pernoctaciones no regladas con las regladas, obtenidos de la encuesta FAMILITUR del Instituto de Estudios Turísticos (IET), a las indicadas pernoctaciones regladas de españoles, del INE:

$$\text{Pernoctaciones no regladas de españoles} = \frac{\text{Pernoctaciones no regladas (FAMILITUR)}}{\text{Pernoctaciones regladas (FAMILITUR)}} \times \text{Pernoctaciones regladas de españoles (INE)}$$

Debe advertirse que la encuesta FAMILITUR (pernoctaciones de españoles regladas y no regladas), que realiza el Instituto de Estudios Turísticos (IET), experimentó un cambio de metodología muy importante en 2005, que ha dado lugar a que el cociente anterior (pernoctaciones no regladas/pernoctaciones regladas) sea sensiblemente inferior al de años anteriores, obteniéndose, por tanto, un número de pernoctaciones no regladas de españoles muy inferior.

Las pernoctaciones no regladas de extranjeros se han estimado con una metodología que comprende fundamentalmente tres fases. En la primera fase, se obtienen las viviendas de potencial uso turístico (viviendas secundarias y viviendas vacías) a partir del Censo de Población y Viviendas del INE de 2001. En una segunda fase, se estiman las viviendas de real uso turístico aplicando fundamentalmente a las anteriores (viviendas de potencial uso turístico) unos coeficientes de intervención de utilización de las mismas, con lo que se determinan las viviendas de real uso turístico y el número de plazas utilizadas. En la tercera fase se aplica al total de plazas de uso turístico el grado de ocupación de las mismas durante un año por parte de turistas extranjeros, obteniéndose en definitiva el número de pernoctaciones no regladas de turistas extranjeros. Los indicados coeficientes de utilización y ocupación, así como el número de plazas por vivienda, que se han aplicado a las viviendas de potencial uso turístico (viviendas secundarias y desocupadas), se obtuvieron en 1989, 1991 y 1992 mediante determinados estudios y encuestas realizadas por el IET a los poseedores de las referidas viviendas de potencial uso turístico. Se ha considerado que esta metodología

empleada para la determinación de las pernoctaciones no regladas de extranjeros, basada en el último Censo de Viviendas 2001 y en dichas encuestas del IET, conduce a unas estimaciones más precisas que las que se derivan directamente de la actual encuesta de EGATUR (IET), de cuyo informe ya se deducen las dificultades que todavía presenta esta encuesta debido a su gran complejidad, que afecta a las estimaciones de las pernoctaciones, independientemente de la gran utilidad de la encuesta EGATUR para el análisis en general de información muy diversa e importante sobre tipología, hábitos y comportamiento del turismo extranjero en España.

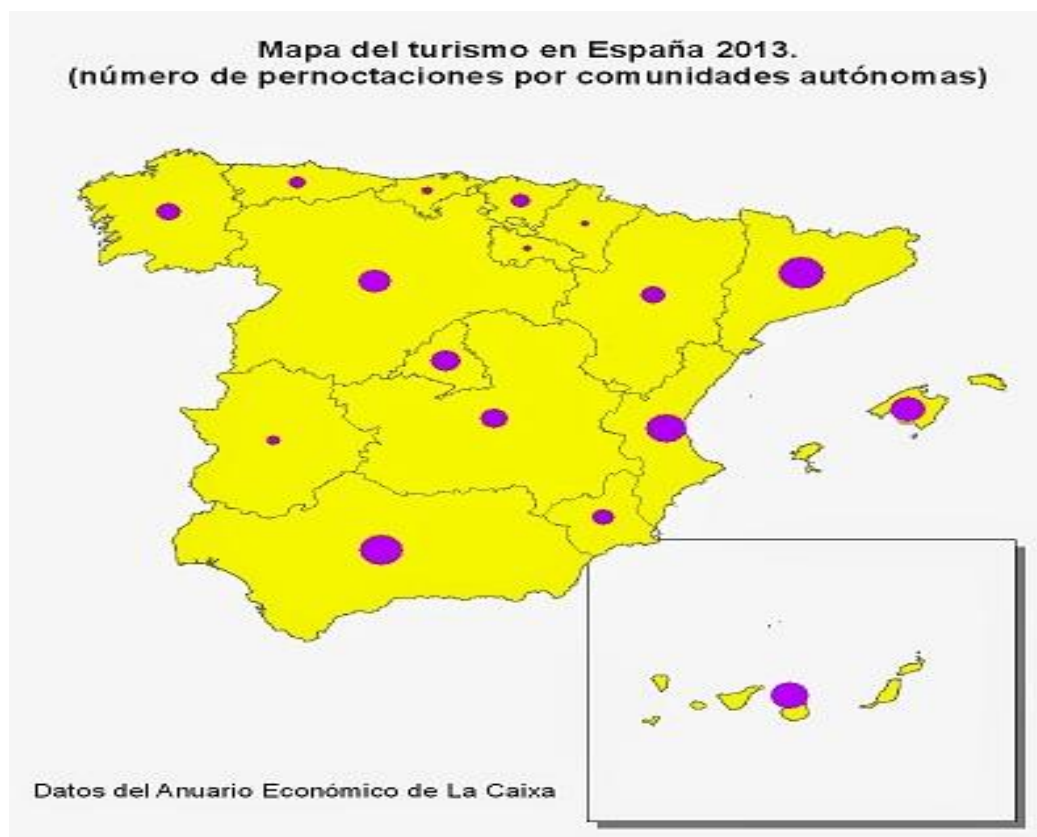


Figura 5. Pernoctaciones por comunidades autónomas

Como se ha explicado ya, los datos se han obtenido del Anuario Económico de la Caixa; el cual a su vez se ha basado en fuentes del INE y del Instituto de Estudios Turísticos.

Nuestra variable es “pernoctaciones totales”, suma de las regladas y no regladas.

Como podemos observar en la siguiente tabla por aeropuerto (provincia), los datos se repiten en los años 2009 y 2010; así como en los años 2011, 2012 y 2013.

Puesto que los datos de pernoctaciones vienen dados por comunidades y provincias, de nuevo a los aeropuertos insulares se les ha aplicado un factor de corrección en función de su población total, como se detalló en un principio.

Aeropuerto	2013	2012	2011	2010	2009
A Coruña	10.296.050	10.296.050	10.296.050	10.627.312	10.627.312
Alicante	82.042.464	82.042.464	82.042.464	85.533.480	85.533.480
Almería	16.427.020	16.427.020	16.427.020	18.399.164	18.399.164
Asturias	18.108.734	18.108.734	18.108.734	19.506.308	19.506.308
Badajoz	6.844.440	6.844.440	6.844.440	8.063.600	8.063.600
Barcelona	79.199.184	79.199.184	79.199.184	75.823.808	75.823.808
Bilbao	8.539.019	8.539.019	8.539.019	7.695.811	7.695.811
Burgos	8.173.930	8.173.930	8.173.930	8.072.092	8.072.092
El Hierro	641.934	641.934	641.934	717.300	717.300
Fuerteventura	7.661.338	7.661.338	7.661.338	8.665.064	8.665.064
Girona	50.860.128	50.860.128	50.860.128	52.384.132	52.384.132
Gran Canaria	60.540.207	60.540.207	60.540.207	68.471.692	68.471.692
Granada-Jaén	22.601.956	22.601.956	22.601.956	25.115.658	25.115.658
Ibiza	14.051.008	14.051.008	14.051.008	16.032.693	16.032.693
Jerez	34.572.240	34.572.240	34.572.240	36.265.976	36.265.976
La Gomera	1.277.755	1.277.755	1.277.755	1.427.769	1.427.769
La Palma	5.062.109	5.062.109	5.062.109	5.656.424	5.656.424
Lanzarote	9.975.375	9.975.375	9.975.375	11.282.267	11.282.267
León	7.804.567	7.804.567	7.804.567	7.835.081	7.835.081
Logroño	7.765.373	7.765.373	7.765.373	8.243.472	8.243.472
Madrid Barajas	45.976.204	45.976.204	45.976.204	47.538.978	47.538.978
Málaga	64.380.500	64.380.500	64.380.500	67.764.520	67.764.520
Melilla	826.438	826.438	826.438	925.656	925.656
Menorca	9.327.057	9.327.057	9.327.057	10.642.499	10.642.499
Murcia	18.993.522	18.993.522	18.993.522	20.973.374	20.973.374
Palma de Mallorca	86.481.263	86.481.263	86.481.263	98.678.153	98.678.153
Pamplona	8.954.211	8.954.211	8.954.211	8.700.936	8.700.936
Reus	46.252.580	46.252.580	46.252.580	47.948.432	47.948.432
Salamanca	7.437.945	7.437.945	7.437.945	7.356.705	7.356.705
San Sebastián	6.540.960	6.540.960	6.540.960	5.764.285	5.764.285
Santander	16.191.360	16.191.360	16.191.360	16.298.958	16.298.958
Santiago	10.296.050	10.296.050	10.296.050	10.627.312	10.627.312
Sevilla	15.109.798	15.109.798	15.109.798	15.632.435	15.632.435
Tenerife Norte	54.154.790	54.154.790	54.154.790	60.512.810	60.512.810
Tenerife Sur	54.154.790	54.154.790	54.154.790	60.512.810	60.512.810
Valencia	64.322.856	64.322.856	64.322.856	64.335.492	64.335.492
Valladolid	6.449.453	6.449.453	6.449.453	6.378.026	6.378.026

Vigo	10.481.157	10.481.157	10.481.157	10.877.518	10.877.518
Zaragoza	10.916.854	10.916.854	10.916.854	12.409.155	12.409.155

Tabla 15. Variable Pernoctaciones

3.3.13 EXPORTACIONES

Pese a ser posiblemente la variable menos relacionada con el tráfico aéreo de pasajeros o con el propio pasajero, se ha creído conveniente introducir esta variable, así como las importaciones, puesto que en España las exportaciones en los últimos años han experimentado un crecimiento del 47% (2009-2013); y puede ser interesante incluir ambas en el modelo estadístico.

Además España es el 20º exportador de mercancías del mundo y el 8º en comercio de servicios. Por otra parte, la balanza comercial estuvo cercana del equilibrio en 2013 y estará en equilibrio en 2014.

Según datos del Ministerio de Economía y Competitividad, las exportaciones experimentaron un crecimiento del 47% entre 2009 y 2013, con una aportación al crecimiento en 2012 del 2,5% del PIB. y del 1,5% en 2013; lo que hace pensar que el comportamiento de las exportaciones en los últimos años ha sido extraordinario.

Este buen comportamiento de las exportaciones en 2013 ha llevado a una reducción del déficit comercial del 55% en términos interanuales. En este periodo las exportaciones alcanzaron un nuevo record de 234.239 millones de euros, el nivel más alto desde que hay registros (1971), lo que supone un aumento del 5,2% respecto al año anterior. Al mismo tiempo, las importaciones se redujeron un 1,3% respecto al año anterior. La menor caída en 2013 respecto a la del año anterior (-2,8%) puede indicar una progresiva recuperación de la demanda interna. La balanza comercial previsiblemente estará en equilibrio en 2014, hecho histórico en la balanza comercial española.

Como se ha dicho anteriormente, esta evolución sitúa a España en el 20º país más exportador del mundo, según datos de la Organización Mundial de Comercio (OMC). Mientras que otros grandes exportadores como Alemania, Francia, Italia o Japón perdieron cuota de mercado, ésta se ha mantenido en España alrededor del 1,6% del total mundial.

Por primera vez en mucho tiempo, España tiene superávit comercial frente a la Unión Europea y a la Zona Euro, así como con Francia, Italia y Austria, entre otros, y el tradicional déficit comercial con Alemania se está reduciendo drásticamente. Esta situación ha ido mejorando a lo largo de 2013. Así en 2013, España registró un superávit comercial de 17.695 millones de euros frente a la Unión Europea y de 10.987 millones de euros frente a la Zona Euro, lo que supone un aumento del 40% y 42,3% respectivamente en comparación a los resultados de 2012.

La economía española cierra 2013 con una capacidad de financiación del 1,4% del PIB. Por primera vez desde 1997 España ha sido prestamista neta al resto del mundo en 2013.

Este es uno de los ajustes más rápidos realizados por un país, ya que después de 14 años de déficits por cuenta corriente, España tiene superávit y este ajuste se ha hecho vía un incremento de las exportaciones debido a la mejora en la calidad y la competitividad de las empresas españolas, no recurriendo a una devaluación, dado que al ser miembros de la zona euro, eso no es posible.

Los datos se han obtenido del fichero de Comercio Exterior de Mercancías del Ministerio de Economía y Competitividad (Secretaría de Estado de Comercio). Todos los datos son definitivos, excepto los de 2013 que son provisionales.

A continuación tenemos la tabla que relaciona cada aeropuerto con sus exportaciones (en millones de euros) para los años de estudio. Al igual que en las demás variables, a cada aeropuerto se le asigna los datos de la provincia a la que pertenece; y puesto que los datos de las exportaciones vienen dados por comunidades y provincias, de nuevo a los aeropuertos insulares se les ha aplicado un factor de corrección en función de su población total, como se detalló en un principio.

Aeropuerto	2013	2012	2011	2010	2009
A Coruña	8.403,0	8.018,3	6.885,2	6.021,9	5.519,2
Alicante	4.504,9	4.111,4	3.739,7	3.715,1	3.739,2
Almería	2.656,0	2.439,3	2.176,7	2.091,9	1.868,3
Asturias	3.711,6	3.837,6	3.767,4	3.368,8	2.573,4
Badajoz	1.206,8	1.290,7	1.146,5	955,9	865,9
Barcelona	45.280,2	45.444,3	42.321,8	38.073,9	32.480,8
Bilbao	6.489,5	8.491,1	8.477,3	7.829,2	6.144,0

Burgos	3.131,0	3.146,5	2.921,4	2.513,1	2.314,4
El Hierro	8,3	8,6	6,8	7,0	5,9
Fuerteventura	178,2	170,9	177,5	127,7	95,7
Girona	4.272,4	4.244,7	4.046,6	3.562,3	3.139,6
Gran Canaria	1.408,5	1.350,3	1.402,8	1.009,5	756,3
Granada-Jaén	911,6	877,8	832,1	690,4	719,5
Ibiza	107,0	130,0	110,5	111,8	152,6
Jerez	8.316,5	8.178,4	7.573,0	5.528,7	4.160,1
La Gomera	16,5	17,2	13,6	13,9	11,8
La Palma	65,3	68,0	53,8	55,2	46,9
Lanzarote	232,1	222,5	231,1	166,3	124,6
León	1.183,7	1.531,2	1.277,6	878,6	732,8
Logroño	1.498,6	1.474,8	1.489,9	1.299,9	1.126,1
Madrid Barajas	30.446,6	27.941,7	26.722,3	21.331,9	18.997,4
Málaga	1.420,0	1.392,0	1.493,3	1.348,8	1.035,9
Melilla	34,2	77,8	4,0	3,6	6,8
Menorca	71,0	86,3	73,4	74,2	101,3
Murcia	9.299,9	8.956,2	5.470,1	4.994,9	4.367,8
Palma de Mallorca	658,5	800,0	680,1	688,2	939,5
Pamplona	7.382,4	7.235,8	8.302,3	7.402,2	5.477,9
Reus	7.076,2	7.545,5	7.142,5	5.900,5	4.704,4
Salamanca	870,9	733,9	904,6	564,8	616,7
San Sebastián	8.464,7	7.069,1	6.664,5	5.618,1	5.207,5
Santander	2.413,8	2.713,4	2.678,6	2.283,6	1.832,4
Santiago	8.403,0	8.018,3	6.885,2	6.021,9	5.519,2
Sevilla	4.574,0	3.995,6	3.425,7	3.361,0	2.897,0
Tenerife Norte	698,2	727,4	575,3	590,6	501,5
Tenerife Sur	698,2	727,4	575,3	590,6	501,5
Valencia	13.594,5	11.473,6	11.408,1	10.407,6	9.149,7
Valladolid	4.618,8	3.592,1	3.846,1	3.219,4	2.580,4
Vigo	8.437,2	7.010,1	8.692,2	7.847,7	7.417,4
Zaragoza	7.558,7	7.705,0	8.093,8	7.437,1	6.312,0

Tabla 16. Variable Exportaciones

3.3.14 IMPORTACIONES

Como no podía ser de otra manera, si tenemos en cuenta las exportaciones de España en el modelo, debemos a su vez introducir otra variable con las importaciones, para poder completar la balanza comercial exterior de mercancías.

A diferencia de las exportaciones, en 2013 las importaciones en España cayeron un 2,82% respecto al año anterior. Las compras al exterior representan el 24,94% de

nuestro PIB, por lo que nos encontramos en el puesto 41, de 181 países, del ranking de importaciones respecto al PIB, ordenado de menor a mayor porcentaje.

Las importaciones en 2013 supusieron 255.163,0 millones de euros; y en este caso España ocupa el puesto número 166 de la lista de importaciones mundiales, ordenadas de menor a mayor valor.

En 2013 hubo déficit en la Balanza comercial ya que, a pesar de producirse un incremento de las exportaciones, fueron menores que las importaciones.

Si miramos la evolución de las importaciones en España en los últimos años se observa que ha caído respecto a 2012, como ya hemos visto, aunque han subido respecto a 2003 cuando fueron de 184.408,0 millones de euros, que suponía un 23,55% de su PIB.

Como el presente proyecto es un estudio acerca del tráfico de pasajeros, cabe señalar, que no es objeto de este apartado analizar a que países exportamos o importamos, ni qué tipo de mercancías son las más o menos comunes.

Al igual que en la variable anterior, los datos se han obtenido del fichero de Comercio Exterior de Mercancías del Ministerio de Economía y Competitividad (Secretaría de Estado de Comercio). Del mismo modo que antes, todos los datos son definitivos, excepto los de 2013 que son provisionales.

Debajo se detalla la tabla que relaciona cada aeropuerto con sus exportaciones (en millones de euros) para los años de estudio. El procedimiento para cada aeropuerto por provincia y para los que pertenecen a las distintas islas es similar al del apartado anterior.

Aeropuerto	2013	2012	2011	2010	2009
A Coruña	7.638,4	8.381,7	7.753,1	6.979,3	5.708,1
Alicante	2.996,5	2.967,9	3.165,6	3.108,8	2.743,3
Almería	2.137,8	1.283,0	1.007,1	446,2	601,4
Asturias	3.352,3	3.868,1	4.193,2	3.713,8	2.864,7
Badajoz	777,7	631,3	720,5	672,1	585,0
Barcelona	52.408,6	54.941,1	57.190,8	54.111,9	47.270,7
Bilbao	3.025,0	10.074,1	11.234,6	9.855,6	7.832,9
Burgos	2.347,4	2.887,2	2.485,3	1.933,2	1.965,4
El Hierro	27,1	31,1	29,3	30,1	23,0
Fuerteventura	174,1	178,6	190,4	180,9	145,2
Girona	2.020,3	2.017,9	2.228,3	2.046,8	1.799,4

Gran Canaria	1.375,9	1.411,0	1.504,2	1.429,2	1.147,6
Granada-Jaén	526,1	550,6	631,1	854,8	612,3
Ibiza	173,9	181,0	192,4	200,2	147,8
Jerez	13.029,9	13.845,8	12.876,0	10.634,4	7.438,4
La Gomera	54,0	62,0	58,3	59,9	45,8
La Palma	213,8	245,6	230,9	237,4	181,5
Lanzarote	226,7	232,5	247,9	235,5	189,1
León	369,8	395,3	420,8	445,0	354,1
Logroño	1.056,9	995,3	1.121,5	981,4	791,1
Madrid Barajas	47.287,1	50.338,6	53.390,3	51.413,4	47.370,1
Málaga	1.204,9	1.274,3	1.440,1	1.492,2	1.266,8
Melilla	248,4	225,7	160,1	124,7	140,3
Menorca	115,4	120,2	127,7	132,9	98,1
Murcia	11.913,1	12.700,0	10.301,2	7.542,3	6.433,3
Palma de Mallorca	1.070,4	1.114,1	1.184,4	1.232,0	910,0
Pamplona	3.836,8	4.323,9	5.360,5	4.493,3	3.569,9
Reus	11.375,5	12.036,8	11.354,9	10.441,5	7.778,2
Salamanca	525,1	643,9	633,3	727,6	700,6
San Sebastián	10.328,1	3.223,7	3.412,9	3.234,1	2.699,5
Santander	1.724,3	1.754,5	1.990,0	1.957,8	1.511,0
Santiago	7.638,4	8.381,7	7.753,1	6.979,3	5.708,1
Sevilla	3.336,2	2.767,6	2.549,6	2.655,0	2.551,4
Tenerife Norte	2.287,2	2.627,2	2.470,4	2.539,8	1.942,1
Tenerife Sur	2.287,2	2.627,2	2.470,4	2.539,8	1.942,1
Valencia	12.949,8	12.144,4	13.182,4	12.502,0	10.673,9
Valladolid	6.606,9	6.064,0	5.865,4	4.953,1	4.050,5
Vigo	5.421,2	5.134,4	5.150,1	5.769,9	6.034,8
Zaragoza	6.218,6	6.087,6	7.250,8	6.489,5	5.599,0

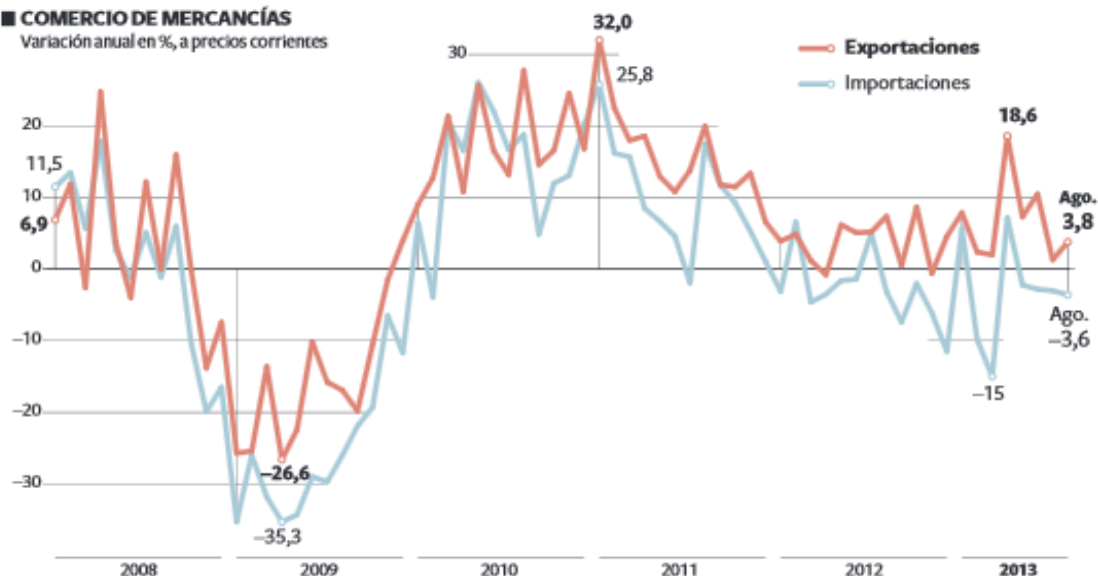
Tabla 17. Variable Importaciones

Por último, en los dos siguientes gráficos tenemos la evolución de las exportaciones e importaciones así como el saldo comercial para los años de estudio.

El sector exterior en España

■ COMERCIO DE MERCANCÍAS

Variación anual en %, a precios corrientes



■ SALDO COMERCIAL (exportaciones – importaciones)

En millones de euros. Entre enero y agosto de cada año



Fuente: Ministerio de Economía.

Gráfico 9. Evolución de las exportaciones e importaciones

3.3.15 MERCADO POTENCIAL TOTAL

Como se explica en el Anuario Económico de la Caixa, para estudiar la distribución geográfica de los flujos comerciales, debe tenerse presente que ni el municipio ni la provincia ni la comunidad autónoma son divisiones territoriales apropiadas para un análisis de esta naturaleza. Por ello, para la configuración del Anuario se ha determinado qué municipios (cabeceras de áreas y de subárea) ejercen atracción comercial sobre otros y cuáles se sienten atraídos o gravitan comercialmente sobre aquéllos.

El Anuario incluye los flujos comerciales o desplazamientos de los consumidores en España, de unos municipios ("municipios de origen") a otros ("municipios destino"), que son centros de atracción comercial denominados cabecera de área comercial, para realizar sus compras más importantes. La información se presenta organizada en función del concepto de "Área comercial", entendiendo como tal la zona geográfica común que queda delimitada por los habitantes de un conjunto de municipios en su proceso de compra de bienes de consumo en un municipio cabecera de área. Las áreas comerciales definidas se refieren a 1 de enero de 2012, fecha en la que se han identificado en el territorio español 77 áreas comerciales, que a su vez comprenden 268 subáreas comerciales.

La mayor parte de los consumidores españoles acuden a otros municipios, en mayor o menor medida, para realizar sus compras. En ocasiones, el motivo principal de los desplazamientos no es la compra, sino otras causas (capitalidad, servicios administrativos, servicios financieros, servicios sanitarios, ocio, etc.), y de paso, realizan compras.

Para evaluar la capacidad comercial de las áreas comerciales se estima el mercado potencial de los municipios que constituyen el núcleo central de las 77 áreas, además de las 268 subáreas comerciales. También se estima el mercado potencial de los municipios de más de 50.000 habitantes que no son cabecera de área ni de subárea (28 municipios).

El mercado potencial se subdivide en tres partes fundamentales en función de la procedencia del comprador:

- ☐ Mercado potencial local
- ☐ Mercado potencial del resto del área
- ☐ Mercado potencial turístico

Se entiende por mercado potencial el volumen de ventas anual del comercio minorista de un municipio durante un año. Estas ventas se han estimado indirectamente a través del gasto realizado por los consumidores en el comercio minorista situado en el municipio. Para los municipios cabecera de área comercial se ha hecho una estimación del Mercado Potencial Total: ventas del comercio detallista a los consumidores domiciliados en el municipio (Mercado Potencial Local) y a los residentes en el resto del área comercial y turistas (Mercado Potencial del Resto del Área y Turismo). En el caso de los municipios que son cabecera de subárea y de los de más de 50.000

habitantes que no son cabecera de área ni de subárea, la estimación del mercado potencial sólo abarca el mercado local (Mercado Potencial Local), por la dificultad de estimar el mercado potencial del resto del área y del turismo.

El mercado potencial de cada cabecera de área se presenta desglosado, además, atendiendo a la tipología del bien adquirido. En conjunto, las variables que se determinan se recogen en el cuadro adjunto.

Mercado potencial Productos de alimentación	Mercado potencial Productos no alimentarios		Mercado potencial total (alimentación y no alimentación)
Local	Local	Vestido y calzado	Local
		Hogar	
		Resto	
Resto área y turismo	Resto área y turismo		Resto área y turismo
Total alimentación	Total no alimentación		Total

Tabla 18. Mercado potencial según el bien adquirido

Como es habitual en los estudios e informes sobre distribución comercial, la actividad comercial minorista excluye las ventas de estancos (excepto tabaco, que está incluido en el grupo de Alimentación), farmacias, automóviles y carburantes, pese a que la Clasificación Nacional de Actividades Económicas (CNAE) las considera comercio al por menor.

El grupo de “Alimentación” comprende todos los productos alimentarios adquiridos en el comercio minorista y las bebidas (alcohólicas y no alcohólicas) consumidas en el hogar, así como el tabaco. “Vestido y calzado” comprende: prendas y complementos de vestir de hombre, mujer y niño; calzado; ropa y calzado deportivo. “Hogar” comprende: muebles; textil hogar; electrodomésticos; material eléctrico del hogar; artículos de cocina, menaje, vajillas y cristalerías; droguería y limpieza; utensilios de jardín. “Resto” (de no alimentación) comprende: artículos recreativos (radio, tv., vídeos, discos, artículos de deporte, juguetes, instrumentos de música, fotografía, etc.); papelería, libros, periódicos y revistas; perfumería y cosmética; relojería, joyería y bisutería; artículos de fumador; artículos de viaje; plantas y flores y animales de compañía; etc.

Además en el Anuario se proporciona un ranking del mercado potencial total (alimentación y no alimentación) de los 77 municipios cabecera de área comercial.

El mercado potencial (ventas del comercio minorista) permite el conocimiento de la participación de una gran empresa o centro comercial en dicho mercado potencial, y también puede ser un indicador aproximado de la potencialidad de actividades no comerciales, tales como bancarias, de seguros, sanitarias, etc., debido a que dichas actividades económicas suelen estar muy correlacionadas con el volumen de ventas del comercio minorista.

Debe tenerse en cuenta que el mercado potencial que se indica en el Anuario es un indicador orientativo de las ventas totales del comercio minorista, ya que no existe un censo del comercio minorista y sus ventas se han estimado indirectamente a partir del gasto que realizan los consumidores en el comercio minorista, y este gasto se ha obtenido a partir de la Encuesta de Presupuestos Familiares del INE, que ha sido la base para estimar el gasto por habitante que se realiza en el comercio minorista con mayor desagregación territorial: estratos de población provinciales (conjunto de municipios de menos de 10.000 habitantes, de 10.000 a 20.000, 20.000 a 50.000, 50.000 a 100.000, y más de 100.000).

Asimismo, en el Anuario se estiman las ventas del comercio minorista ubicado en un municipio cabecera de área a las personas procedentes del resto del área comercial, así como las ventas que se realizan a los turistas.

El Anuario Económico de La Caixa ha sido elaborado con datos propios y del INE (Instituto Nacional de Estadística). La delimitación de las áreas y subáreas comerciales, así como la gravitación entre ellas está explicado anteriormente de manera detallada en el apartado correspondiente a los índices (después del índice industrial).

Los datos introducidos en el modelo estadístico son los correspondientes al mercado potencial total (en millones de euros), el cual es suma del mercado potencial local y del mercado potencial resto de área y turismo. En caso de desearlo, en el propio Anuario se pueden encontrar todos los datos desglosados; así como las divisiones territoriales realizadas y el porqué de estas, rankings y porcentajes.

A cada aeropuerto se le ha asignado su municipio cabecera de área comercial, que en la mayoría de ocasiones coincide con su capital de provincia (los datos engloban toda la provincia); sin embargo, en algunos casos no se da así, sino que en una misma provincia hay varias cabeceras de área comercial. Por tanto, para cada aeropuerto se suman todas las cabeceras de área comercial existentes en esa provincia. Los aeropuertos (provincias) afectados son los siguientes:

- A Coruña: Ferrol y A Coruña.
- Asturias: Avilés, Gijón y Oviedo.
- Badajoz: Don Benito-Villanueva de la S. y Badajoz.
- Barcelona: Vic, Manresa y Barcelona.
- Girona: Olot, Figueres y Girona.
- Jerez: incluye Algeciras-La Línea de la Concepción, Cádiz-San Fernando y Jerez de la Frontera.
- León: Ponferrada y León.
- Málaga: Ronda y Málaga.
- Murcia: Lorca, Cartagena y Murcia.
- Salamanca: Ciudad Rodrigo y Salamanca.
- Vigo: Pontevedra y Vigo.
- Zaragoza: Calatayud y Zaragoza.

Además, solo a los aeropuertos insulares Canarios se les ha aplicado el factor de corrección explicado antes; ya que en este caso si hay datos para las Islas Baleares.

A continuación se expone la tabla resumen del mercado potencial total (en millones de euros) para los años de estudio según aeropuerto.

Aeropuerto	2013	2012	2011	2010	2009
A Coruña	1.697	1.673	1.636	1.635	1.693
Alicante	1.901	1.818	1.984	2.016	2.184
Almería	914	862	901	924	896
Asturias	2.763	3.062	2.884	2.915	2.988
Badajoz	1.097	1.293	1.248	1.165	1.119
Barcelona	10.307	10.528	11.142	11.325	10.955
Bilbao	2.183	2.282	2.191	2.360	2.358
Burgos	877	1.425	1.326	1.512	1.445
El Hierro	16	16	16	19	21
Fuerteventura	163	170	183	192	210
Girona	969	1.030	968	1.027	1.082
Gran Canaria	1.285	1.344	1.447	1.517	1.661
Granada-Jaén	1.217	1.224	1.453	1.484	1.535
Ibiza	387	361	349	387	428
Jerez	2.305	2.487	2.519	2.624	2.718
La Gomera	33	32	33	38	42
La Palma	129	125	129	149	166
Lanzarote	212	222	238	250	274

León	1.158	1.551	1.572	1.764	1.697
Logroño	733	806	773	762	791
Madrid Barajas	15.232	16.011	16.264	16.788	16.336
Málaga	2.962	3.041	3.183	3.424	3.068
Melilla	328	321	340	366	304
Menorca	186	177	174	191	214
Murcia	3.216	3.478	3.463	3.957	3.940
Palma de Mallorca	1.685	2.087	2.094	2.381	2.470
Pamplona	1.153	1.224	1.246	1.314	1.248
Reus	837	848	930	965	997
Salamanca	891	827	823	969	869
San Sebastián	1.125	1.149	1.231	1.183	1.167
Santander	889	915	993	973	931
Santiago	675	662	607	660	631
Sevilla	3.556	3.707	3.593	4.034	4.190
Tenerife Norte	1.382	1.338	1.384	1.592	1.772
Tenerife Sur	1.382	1.338	1.384	1.592	1.772
Valencia	3.922	4.070	4.542	4.824	4.569
Valladolid	1.722	2.197	2.173	2.453	2.499
Vigo	1.851	1.916	1.885	1.832	2.002
Zaragoza	3.205	3.547	3.519	3.771	4.072

Tabla 19. Variable Mercado potencial total

3.3.16 CUOTA DE MERCADO

La cuota de mercado de los municipios se elabora mediante números índices que expresan comparativamente la capacidad de compra o de consumo de los municipios e indican la participación que corresponde a cada municipio sobre una base nacional de 100.000 unidades.

El valor del indicador se obtiene en función de 6 variables: población, teléfonos fijos (de uso doméstico), automóviles, camiones (camiones y furgonetas), oficinas bancarias y actividades comerciales minoristas, referidas al 1 de enero de 2012 (para el año 2013). Es decir, la capacidad de compra o consumo de un municipio se mide, no sólo en función de la importancia de la población, sino también en función del poder adquisitivo de la misma, que viene representado por las cinco variables restantes indicadas.

El modelo que se ha aplicado es el siguiente:

$$I_c = \frac{I_p + \frac{I_t + I_a + I_{cam} + I_b + I_{com}}{5}}{2}$$

donde:

I_c = Índice compuesto llamado "cuota de mercado"

I_p = Índice simple de la población

I_t = Índice simple de los teléfonos fijos

I_a = Índice simple de los automóviles

I_{cam} = Índice simple de los camiones (camiones y furgonetas)

I_b = Índice simple de las oficinas bancarias

I_{com} = Índice simple de las actividades comerciales minoristas

Debe advertirse que, en algunos municipios, la ratio de número de vehículos de motor (automóviles y camiones, fundamentalmente) por 1.000 habitantes resulta muy elevada, lo cual puede deberse, no a la existencia de una gran capacidad adquisitiva de sus habitantes, sino al hecho de existir, en el municipio, agencias de vehículos de alquiler o que el coste de matriculación de los automóviles sea más bajo que en su entorno. Esta circunstancia podría originar una distorsión metodológica en el cálculo de la cuota de mercado de algunos municipios (a título de ejemplo, cabe citar el municipio de Robledo de Chavela, de la provincia de Madrid, donde el parque de automóviles matriculados asciende a 68.276, con una población de sólo 4.058 habitantes). En estos casos, el Anuario ha llevado a cabo una depuración a efectos del cálculo de la cuota de mercado.

La cuota de mercado constituye una orientación adecuada para valorar y ponderar la cantidad de productos y servicios que, teóricamente y en igualdad de condiciones, pueden absorber los municipios, especialmente cuando se trata de productos o servicios que son objeto de una distribución homogénea. Por lo tanto, puede resultar de utilidad para la gestión y planificación comercial de las empresas; además, al tratarse de un solo indicador resulta de fácil interpretación y aplicación para el modelo estadístico que se va a implementar.

Los datos, que tenemos tabulados justo debajo, de nuevo han sido obtenidos del Anuario Económico de la Caixa; asignando a cada aeropuerto su valor provincial, excepto en el caso de los aeropuertos insulares, donde se aplican los factores de corrección ya explicados.

Aeropuerto	2013	2012	2011	2010	2009
A Coruña	2.415	2.415	2.412	2.413	2.413
Alicante	3.990	3.990	3.987	4.001	4.020
Almería	1.480	1.480	1.481	1.481	1.473
Asturias	2.316	2.316	2.316	2.312	2.302
Badajoz	1.499	1.499	1.501	1.495	1.484
Barcelona	11.483	11.483	11.502	11.563	11.650
Bilbao	2.402	2.402	2.393	2.375	2.357
Burgos	844	844	844	848	848
El Hierro	24	24	24	24	24
Fuerteventura	236	236	235	234	233
Girona	1.721	1.721	1.721	1.741	1.745
Gran Canaria	1.867	1.867	1.855	1.850	1.842
Granada-Jaén	1.922	1.922	1.928	1.933	1.929
Ibiza	324	324	323	322	320
Jerez	2.441	2.441	2.439	2.426	2.421
La Gomera	47	47	47	47	47
La Palma	187	187	188	187	186
Lanzarote	308	308	306	305	303
León	1.107	1.107	1.111	1.105	1.104
Logroño	746	746	746	743	742
Madrid Barajas	13.540	13.540	13.519	13.519	13.538
Málaga	3.333	3.333	3.321	3.312	3.312
Melilla	162	162	158	153	149
Menorca	215	215	214	214	213
Murcia	3.020	3.020	3.020	3.014	3.010
Palma de Mallorca	1.992	1.992	1.987	1.980	1.971
Pamplona	1.465	1.465	1.463	1.446	1.434
Reus	1.766	1.766	1.770	1.784	1.788
Salamanca	782	782	784	781	778
San Sebastián	1.525	1.525	1.518	1.506	1.493
Santander	1.279	1.279	1.277	1.270	1.260
Santiago	2.415	2.415	2.412	2.413	2.413
Sevilla	3.809	3.809	3.806	3.796	3.805
Tenerife Norte	1.999	1.999	2.010	2.005	1.992
Tenerife Sur	1.999	1.999	2.010	2.005	1.992
Valencia	5.411	5.411	5.428	5.457	5.477
Valladolid	1.128	1.128	1.129	1.126	1.127
Vigo	1.984	1.984	1.989	1.989	1.999
Zaragoza	2.082	2.082	2.077	2.078	2.080

Tabla 20. Variable Cuota de mercado

3.4 VARIABLES AEROPORTUARIAS

Para este tipo de variables es muy importante destacar que no deben ser tomadas con la misma confianza que las variables socioeconómicas, pues las propias variables aeroportuarias están muy interrelacionadas con la variable dependiente “pasajeros” desde el momento inicial de la obtención de datos para cada aeropuerto. Además algunas de ellas son resultado de la síntesis de otras dos; y aunque se haya creído oportuno introducirlas en el modelo y analizarlas posteriormente, deben ser tratadas de forma diferente, siempre interpretándolas con cautela como se indicará en los resultados finales.

3.4.1 COMPAÑÍAS AÉREAS

Los datos han sido obtenidos de la página web de AENA (Sección Aeropuertos-Listado de compañías aéreas). Es importante destacar que los datos aportados en esta variable no tienen una fiabilidad completa, pues dependiendo de la diversa bibliografía consultada se atribuyen más o menos compañías aéreas a cada aeropuerto; hasta tal punto, que la propia página web de AENA ha variado dichos números en un período de tres meses. Esto puede deberse a que algunas compañías dejan de operar en ciertos aeropuertos, mientras que en otros aumenta el número de estas, dependiendo de la época buscada. En la siguiente tabla tenemos el número de aerolíneas que operan en cada aeropuerto español seleccionado a fecha de Octubre de 2014 (realización del presente proyecto), aunque se incluirá en el modelo para el año 2013.

Aeropuerto	Nº aerolíneas	Aeropuerto	Nº aerolíneas
A Coruña	6	Madrid Barajas	88
Alicante	38	Málaga	44
Almería	25	Melilla	3
Asturias	8	Menorca	42
Badajoz	2	Murcia	8
Barcelona	107	Palma de Mallorca	83
Bilbao	22	Pamplona	3
Burgos	3	Reus	9
El Hierro	1	Salamanca	2
Fuerteventura	64	San Sebastián	3
Girona	13	Santander	4

Gran Canaria	60	Santiago	12
Granada-Jaén	11	Sevilla	23
Ibiza	59	Tenerife Norte	14
Jerez	22	Tenerife Sur	75
La Gomera	2	Valencia	39
La Palma	14	Valladolid	7
Lanzarote	70	Vigo	7
León	1	Zaragoza	6
Logroño	2		

Tabla 21. Variable Compañías

En la tabla adjunta podemos observar un ranking de las aerolíneas que operan en España según pasajeros, asientos y factor de ocupación por mercados entre Enero y Diciembre de 2013.

RANKING DE LAS PRINCIPALES COMPAÑÍAS EN EL TRANSPORTE AÉREO EN ESPAÑA POR MERCADOS

Rk.	Compañía	Total			Doméstico		Unión Europea		Extra-UE	
		Cuota 2013	Pasajeros 2013	FO 2013	Pasajeros 2013	FO 2013	Pasajeros 2013	FO 2013	Pasajeros 2013	FO 2013
1	RYANAIR	17,4%	27.457.284	80%	4.962.643	76%	21.588.899	82%	905.742	74%
2	VUELING	10,2%	16.095.436	76%	7.598.556	78%	7.764.000	75%	732.880	68%
3	EASYJET	7,6%	11.977.131	86%	145	46%	10.446.314	87%	1.530.672	83%
4	IBERIA	6,9%	10.895.913	68%	2.328.820	63%	3.685.825	68%	4.881.268	71%
5	GRUPO AIR BERLIN	6,3%	9.871.040	80%	967.048	75%	8.373.108	81%	530.884	80%
6	AIR EUROPA	5,7%	8.955.981	79%	4.785.046	79%	2.431.977	78%	1.738.958	81%
7	THOMSON	2,7%	4.241.776	93%	212	90%	4.240.283	93%	1.281	78%
8	THOMAS COOK	2,5%	3.893.037	87%	1.202	44%	3.505.795	87%	386.040	86%
9	GRUPO TUIFLY	2,4%	3.791.770	88%	126	26%	3.537.503	88%	254.141	87%
10	AIR NOSTRUM	2,4%	3.774.577	61%	2.627.609	61%	1.009.212	61%	137.756	63%
11	LUFTHANSA	2,3%	3.627.560	80%	115	83%	3.627.445	80%	-	-
12	MONARCH	2,2%	3.531.595	87%	213	33%	3.530.177	87%	1.205	75%
13	NORWEGIAN	2,1%	3.242.647	82%	327	85%	1.986.729	82%	1.255.591	81%
14	IBERIA EXPRESS	2,1%	3.240.332	71%	2.581.535	69%	654.064	76%	4.733	62%
15	JET2.COM	2,0%	3.124.139	87%	520	79%	3.123.619	87%	-	-
16	TRANSAVIA	1,9%	2.971.550	85%	569	37%	2.969.439	85%	1.542	74%
17	GRUPO AF/KLM	1,8%	2.797.291	80%	-	-	2.797.291	80%	-	-
18	CONDOR	1,8%	2.796.126	90%	1.090	80%	2.795.036	90%	-	-
19	GRUPO BINTER	1,6%	2.525.750	69%	2.458.190	69%	14.656	65%	52.904	53%
20	BRITISH AIRWAYS	1,2%	1.879.296	76%	-	-	1.879.296	77%	-	-
21	SWISS	1,1%	1.767.267	76%	-	-	-	-	1.767.267	76%
22	AER LINGUS	0,8%	1.203.576	80%	-	-	1.203.576	80%	-	-
23	GERMANWINGS	0,7%	1.138.558	80%	196	27%	1.138.362	80%	-	-
24	WIZZ AIR	0,7%	1.113.901	85%	-	-	1.046.406	85%	67.495	84%
25	SAS	0,7%	1.102.491	81%	-	-	530.894	84%	571.597	78%
26	TAP	0,6%	955.474	72%	-	-	955.404	72%	70	36%
27	ALITALIA	0,5%	843.789	73%	78	41%	843.479	73%	232	20%
28	TURKISH AIRLINES	0,5%	731.483	63%	-	-	328	56%	731.155	63%
29	BRUSSELS AIRLINES	0,5%	726.582	72%	-	-	726.582	72%	-	-
30	TRAVEL SERVICE	0,5%	716.605	82%	5.177	61%	688.836	83%	22.592	67%
31	AMERICAN AIRLINES	0,4%	650.025	82%	-	-	-	-	650.025	82%
32	EMIRATES	0,4%	596.131	54%	-	-	-	-	596.131	54%
33	FINNAIR	0,4%	593.851	87%	-	-	593.851	87%	-	-
34	AEROFLOT	0,3%	534.164	70%	-	-	170	92%	533.994	70%
35	AVIANCA	0,3%	528.269	66%	-	-	-	-	528.269	66%
36	LATAM AIRLINES	0,3%	522.231	57%	-	-	92.300	56%	429.931	57%
37	DELTA AIRLINES	0,3%	502.295	80%	-	-	-	-	502.295	80%
38	TRANSAERO AIRLINES	0,3%	432.694	63%	-	-	-	-	432.694	63%
39	QATAR AIRWAYS	0,3%	414.002	66%	-	-	-	-	414.002	66%
40	GERMANIA	0,2%	375.592	71%	-	-	372.546	72%	3.046	61%
	RESTO	7,3%	11.508.937	-	293.181	-	5.584.311	-	5.631.445	-
	Total	100,0%	157.648.148	78%	28.612.598	73%	103.737.713	81%	25.297.837	72%

Tabla 22. Ranking de compañías por mercados

3.4.2 DESTINOS

Nuevamente, el número de destinos se ha sacado de la página web de AENA (Sección Aeropuertos-Destinos y Nuevas Rutas). Al igual que ocurría con el número de compañías aéreas, los datos sobre destinos prácticamente varían mes a mes, ya que se incluyen rutas nuevas y otras dejan de estar operativas según la época del año y de la localización. A continuación se muestran los destinos por aeropuerto a fecha de Octubre de 2014 (realización del presente proyecto), aunque se incluirá en el modelo para el año 2013.

Aeropuerto	Destinos	Aeropuerto	Destinos
A Coruña	6	Madrid Barajas	162
Alicante	73	Málaga	87
Almería	7	Melilla	4
Asturias	10	Menorca	6
Badajoz	2	Murcia	5
Barcelona	138	Palma de Mallorca	53
Bilbao	30	Pamplona	1
Burgos	1	Reus	3
El Hierro	2	Salamanca	1
Fuerteventura	75	San Sebastián	2
Girona	16	Santander	7
Gran Canaria	125	Santiago	16
Granada-Jaén	5	Sevilla	29
Ibiza	9	Tenerife Norte	16
Jerez	7	Tenerife Sur	120
La Gomera	1	Valencia	41
La Palma	15	Valladolid	1
Lanzarote	82	Vigo	4
León	1	Zaragoza	9
Logroño	1		

Tabla 23. Variable Destinos

En la tabla adjunta podemos observar un ranking de las principales rutas en España por mercados entre Enero y Diciembre de 2013.

RANKING DE LAS PRINCIPALES RUTAS EN EL TRANSPORTE AÉREO EN ESPAÑA POR MERCADOS

Rk.	Ruta	Cuota merc 2013	Pasajeros 2013	F.O. 2013	Δ Pax 13/12	Nº Cías*	Cía Dominante	Cuota Cía Dom	FO Cía Dom.
DOMÉSTICO	1 MADRID-BARCELONA	7,7%	2.211.413	67%	-13,3%	4	IBERIA	49,1%	61%
	2 PALMA DE MALLORCA-BARCELONA	4,9%	1.388.409	73%	-4,9%	4	VUELING	45,2%	71%
	3 PALMA DE MALLORCA-MADRID	4,3%	1.224.144	79%	-15,1%	4	AIR EUROPA	50,4%	82%
	4 GRAN CANARIA-MADRID	4,0%	1.152.763	76%	-11,4%	4	AIR EUROPA	42,1%	84%
	5 TENERIFE NORTE-MADRID	3,9%	1.102.686	74%	-5,6%	4	AIR EUROPA	57,8%	81%
	6 IBIZA-BARCELONA	2,8%	812.384	78%	-4,6%	4	VUELING	72,9%	81%
	7 BARCELONA-SEVILLA	2,8%	802.507	80%	-18,0%	4	VUELING	70,2%	79%
	8 MENORCA-BARCELONA	2,2%	621.031	73%	-5,4%	3	VUELING	76,9%	75%
	9 BARCELONA-MÁLAGA	2,1%	613.434	81%	-22,9%	4	VUELING	69,2%	83%
	10 MADRID-BILBAO	2,1%	606.310	60%	-27,6%	3	IBERIA	67,1%	62%
UNIÓN EUROPEA	1 MADRID-LONDRES	2,3%	2.383.457	69%	-0,6%	6	EASYJET	34,3%	88%
	2 BARCELONA-LONDRES	2,1%	2.174.069	80%	7,3%	7	EASYJET	39,9%	88%
	3 MADRID-PARÍS	2,0%	2.060.940	79%	-11,8%	8	EASYJET	27,2%	85%
	4 BARCELONA-PARIS	2,0%	2.050.512	83%	-0,8%	10	EASYJET	31,8%	85%
	5 MÁLAGA-LONDRES	1,5%	1.605.729	89%	0,8%	9	MONARCH	27,0%	90%
	6 PALMA DE MALLORCA-LONDRES	1,2%	1.287.981	85%	9,1%	7	EASYJET	38,4%	86%
	7 MADRID-ROMA	1,2%	1.223.876	81%	-9,7%	8	IBERIA	37,7%	70%
	8 BARCELONA-AMSTERDAM	1,1%	1.187.200	84%	-5,0%	7	EASYJET	25,5%	89%
	9 ALICANTE-LONDRES	1,1%	1.185.750	86%	8,0%	6	EASYJET	33,7%	84%
	10 PALMA DE MALLORCA-DUSSELDORF	1,1%	1.102.577	84%	10,6%	6	RYANAIR	27,9%	91%
UNIÓN EUROPEA	1 BARCELONA-MOSCÚ	4,1%	1.048.873	75%	20,1%	12	AEROFLOT	21,1%	70%
	2 MADRID-BUENOS AIRES	3,1%	772.651	84%	-1,2%	5	IBERIA	26,3%	65%
	3 MADRID-NUEVA YORK	3,0%	752.131	74%	-4,2%	3	IBERIA	36,9%	59%
	4 MADRID-GINEBRA	2,3%	576.320	71%	-10,7%	5	IBERIA	32,7%	75%
	5 BARCELONA-ZURICH	2,2%	563.261	74%	0,5%	3	AIR EUROPA	37,9%	79%
	6 MADRID-LIMA	2,2%	556.017	67%	-4,1%	3	IBERIA	50,7%	87%
	7 BARCELONA-GINEBRA	2,0%	508.719	76%	-3,7%	2	AVIANCA	51,3%	66%
	8 MADRID-BOGOTÁ	2,0%	500.517	73%	-13,3%	3	IBERIA	43,0%	66%
	9 MADRID-ZURICH	1,9%	486.805	71%	-1,4%	9	NORWEGIAN	36,2%	85%
	10 MADRID-CIUDAD DE MÉXICO	1,9%	483.587	83%	-1,7%	4	EL AL- ISRAEL	44,0%	76%

* Operadores con más de 1.000 pasajeros en la ruta.

Tabla 24. Ranking de rutas por mercados

3.4.3 DESTINOS AL AEROPUERTO MÁS CERCANO

La siguiente variable trata de analizar la posible relación entre los destinos que tiene un aeropuerto determinado y los destinos del aeropuerto más cercano a este. Hay que señalar que España, con algo más de 46 millones de habitantes y 17 autonomías divididas en 50 provincias, más las ciudades autónomas de Ceuta y Melilla, tiene a su

disposición 52 aeropuertos, el 90% de ellos gestionado por AENA. Apenas unos pocos, como el de Ciudad Real, Lérida o Castellón, han sido impulsados por las comunidades autónomas, entidades locales y el sector privado. De todos los que gestiona la entidad pública, solo ocho aeródromos son oficialmente rentables y más de una veintena dispone de otro aeropuerto a menos de una hora por carretera. Alemania, con casi el doble de población que la española (81 millones de habitantes), no necesita más de 39 aeropuertos.

A pesar de que Madrid y Barcelona son los aeropuertos que más tráfico de personas y mercancías alcanzan, un informe de la Fundación de Estudios de Economía Aplicada (Fedea) basado en datos públicos, asegura que únicamente ocho de los 47 aeropuertos gestionados por AENA presentan resultados operativos positivos: Palma de Mallorca, Málaga, Gran Canaria, Alicante, Tenerife Sur, Girona, Bilbao y Murcia.

De todos los 39 deficitarios, Fedea salva a otros trece considerados “necesarios” por absorber gran parte del tráfico de pasajeros y mercancías, aunque no sean eficientes. El debate que se abre es si existe o no la necesidad de renovar su modelo de gestión para conseguir resultados operativos positivos.

Con estos datos, la balanza del mapa aeroportuario español no se sostiene con el precio tan alto que a España le supone mantener 26 aeródromos totalmente ruinosos y que, casi todos, podrían disponer de otro vuelo a menos de una hora en coche. El de Castellón, por ejemplo, con un coste inicial de 151 millones de euros y sin recibir ni un solo vuelo muchos meses después de su inauguración, está a menos de cien kilómetros del de Valencia.

Otro ejemplo es el aeropuerto de Huesca, en el que se invirtieron 60 millones de euros para acoger 160.000 pasajeros anuales, al final pasaron 2.781 pasajeros en todo 2011; y en el año 2013 tan sólo 273 pasajeros. Un número ridículo de visitantes que disponen del aeródromo más cercano, el de Zaragoza, a 97 kilómetros de distancia.

El caso del Pirineo aragonés no es ninguna excepción. Más de la mitad tienen al menos otro aeropuerto cercano a menos de 130 kilómetros de distancia, como Girona, Barcelona, Reus, Valladolid, Salamanca, Alicante, Murcia, Granada y Málaga. En otros casos, el pasajero puede elegir entre dos (Santiago de Compostela, Vigo, Coruña) a

menos de hora y media por asfalto. El caso más llamativo es el de Vitoria, rodeada de cinco aeródromos, sin contar el suyo: Bilbao (a 73 kilómetros), San Sebastián (117), Pamplona (119), Burgos (116) y Logroño (116).

Con los datos del tráfico aéreo de los últimos años y teniendo en cuenta que muchos de ellos se mantienen gracias a las subvenciones de las comunidades autónomas, toca reflexionar sobre si la escasez de pasajeros justifica mantener abiertos algunos de los aeródromos más deficitarios. Mantener el aeropuerto “fantasma” de Castellón, por citar un caso, cuesta 3.600 euros diarios. Hay que replantear si su funcionamiento influye o no en el turismo de la zona y si existen otras alternativas más viables para acceder a esta.

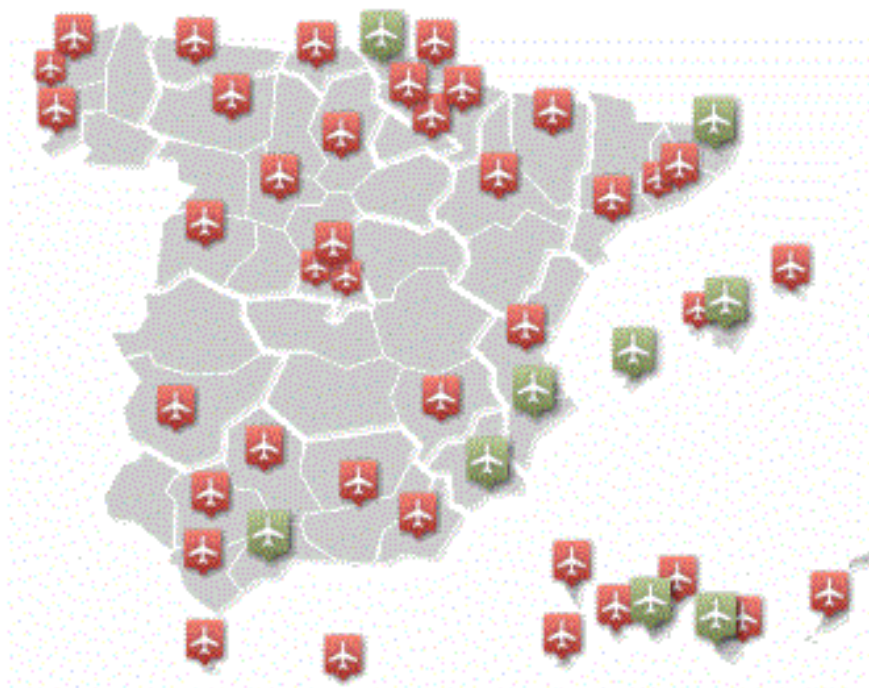


Figura 6. Aeropuertos con resultados operativos positivos y negativos

En la imagen de arriba tenemos en verde los aeropuertos con resultados operativos positivos y en rojo los aeropuertos con resultados operativos negativos. Por lo que se puede decir que en España tan solo hay 9 aeropuertos rentables, aunque alguno de los no rentables tiene un gran número de pasajeros y rutas importantes.

Como ya se ha comentado en la introducción, algunos de estos aeropuertos no están incluidos en el modelo estadístico, pues tienen menos de 10.000 pasajeros al año, sin ningún destino regular o no están gestionados por AENA.

El coeficiente aplicado a cada aeropuerto viene de la fórmula siguiente, y serán dichos coeficientes los que se introduzcan en el modelo estadístico:

$$\text{Coeficiente} = \frac{\text{número de destinos}}{\text{número de destinos al aeropuerto más cercano}}$$

Los datos son los obtenidos en el apartado anterior (Destinos), sacados de la página web de AENA (Sección Aeropuertos-Destinos y Nuevas Rutas), y los kilómetros se han consultado en una calculadora de distancias de internet.

En la tabla posterior tenemos el cálculo del coeficiente por aeropuerto para 2013, así como el nombre del aeropuerto más cercano.

Aeropuerto	Coeficiente	Aeropuerto más cercano
A Coruña	0,375	Santiago
Alicante	14,6	Murcia
Almería	1,4	Granada
Asturias	10	León
Badajoz	0,069	Sevilla
Barcelona	8,625	Girona
Bilbao	4,286	Santander
Burgos	1	Valladolid
El Hierro	0,017	Tenerife Sur
Fuerteventura	0,915	Lanzarote
Girona	0,116	Barcelona
Gran Canaria	1,042	Tenerife Sur
Granada-Jaén	0,057	Málaga
Ibiza	0,17	Palma de Mallorca
Jerez	0,241	Sevilla
La Gomera	0,008	Tenerife Sur
La Palma	0,125	Tenerife Sur
Lanzarote	1,093	Fuerteventura
León	1	Valladolid
Logroño	1	Pamplona
Madrid Barajas	162	Valladolid
Málaga	17,4	Granada
Melilla	0,046	Málaga
Menorca	0,113	Palma de Mallorca
Murcia	0,068	Alicante
Palma de Mallorca	5,889	Ibiza
Pamplona	0,5	San Sebastián
Reus	0,022	Barcelona
Salamanca	1	Valladolid

San Sebastián	0,067	Bilbao
Santander	0,233	Bilbao
Santiago	2,667	A Coruña
Sevilla	4,143	Jerez
Tenerife Norte	0,128	Gran Canaria
Tenerife Sur	0,96	Gran Canaria
Valencia	0,562	Alicante
Valladolid	1	Salamanca, Burgos y León
Vigo	0,25	Santiago
Zaragoza	9	Logroño y Pamplona

Tabla 25. Variable Destinos al aeropuerto más cercano

Para el caso de los aeropuertos peninsulares se ha tomado la distancia más corta por carretera. Como se puede observar, los aeropuertos de Valladolid y Zaragoza tienen varios aeropuertos cerca con el mismo coeficiente.

En los siguientes casos la distancia es prácticamente la misma al aeropuerto elegido que a otro cercano, pero varía el coeficiente. Ante estos casos siempre se ha elegido el aeropuerto con más destinos (coeficiente menor). Estos son los aeropuertos y coeficientes descartados:

- Barcelona: Reus (Coef.:46)
- Bilbao: San Sebastián (Coef.:15)
- Burgos: Logroño (Coef.:1)
- Pamplona: Logroño (Coef.:1)
- San Sebastián: Pamplona (Coef.:2)

Para los aeropuertos insulares de Gran Canaria, La Palma, La Gomera y El Hierro se ha cogido como aeropuerto más cercano el de Tenerife Sur, por tener más pasajeros y destinos que Tenerife Norte (aunque la distancia es similar), por lo que el coeficiente elegido, de nuevo, es más pequeño.

Para el aeropuerto de Palma de Mallorca se ha elegido como más cercano el de Ibiza por la misma razón, en detrimento del de Menorca (Coef.:8,83), que está a la misma distancia.

Del mismo modo para Melilla se ha optado por el aeropuerto de Málaga en vez del de Almería (Coef.:0,57).

Como se puede observar, un mayor coeficiente significará que el aeropuerto tiene más destinos que el más cercano, y viceversa. Si dicho coeficiente es 1, hay los mismos destinos en el aeropuerto elegido que en su más cercano.

3.4.4 DISTANCIA Y MAYORES DESTINOS

Hasta ahora se ha analizado el número de destinos de un aeropuerto y el de su más cercano, pero en ocasiones al pasajero le interesa desplazarse algo más lejos para poder disfrutar de un mayor número de destinos (sobre todo para viajes internacionales e intercontinentales). Esta variable trata de cuantificar una relación entre la distancia a un gran aeropuerto y los destinos que tiene. Para ello se han elegido los 6 aeropuertos con más tráfico de pasajeros de España, que son: Madrid, Barcelona, Palma de Mallorca, Málaga, Gran Canaria y Alicante. Además estos 6 aeropuertos, como es lógico, también están entre los primeros en número de destinos y tienen un amplio reparto por la geografía del país.

El proceso seguido es, para cada aeropuerto del modelo, elegir el más cercano de entre esos 6 mayores aeropuertos y multiplicar la distancia entre ambos (en kilómetros) por el número de destinos que tenga el aeropuerto escogido; y de ahí obtener un coeficiente que poder introducir en el modelo estadístico.

El coeficiente para cada uno de los 6 aeropuertos con más tráfico es simplemente su número de destinos (como si su distancia fuese de 1 km).

De nuevo los datos son los obtenidos en el apartado “Destinos”, sacados de la página web de AENA (Sección Aeropuertos-Destinos y Nuevas Rutas), y los kilómetros se han consultado en una calculadora de distancias de internet.

Es importante destacar que en un primer momento la elección de esta variable fue “Kilómetros/Destinos”, pues los coeficientes parecían guardar una mayor lógica en el ratio Kms-Destinos; pero al introducirlos en el modelo estadístico los resultados eran peores que los actuales que se tienen en los resultados finales.

A continuación tenemos la tabla con el coeficiente por aeropuerto en el año 2013, siendo la fórmula la siguiente:

Coeficiente = Kilómetros a uno de los 6 aeropuertos con más pasajeros × Número de destinos en ese aeropuerto

Aeropuerto	Coeficiente	Aeropuerto	Coeficiente
A Coruña	95742	Madrid Barajas	162
Alicante	73	Málaga	87
Almería	17661	Melilla	18270
Asturias	77760	Menorca	11660
Badajoz	65448	Murcia	5329
Barcelona	138	Palma de Mallorca	53
Bilbao	64314	Pamplona	63504
Burgos	39528	Reus	12972
El Hierro	46000	Salamanca	36450
Fuerteventura	25375	San Sebastián	71280
Girona	14490	Santander	66420
Gran Canaria	125	Santiago	99144
Granada-Jaén	10875	Sevilla	17835
Ibiza	7950	Tenerife Norte	18375
Jerez	19749	Tenerife Sur	23750
La Gomera	34875	Valencia	12483
La Palma	47625	Valladolid	32562
Lanzarote	38000	Vigo	97524
León	54756	Zaragoza	41262
Logroño	53298		

Tabla 26. Variable Distancia y mayores destinos

La siguiente tabla nos indica que aeropuertos engloba cada uno de los 6 con más tráfico de pasajeros, y la distancia entre ellos.

Aeropuerto 1	Aeropuerto 2	Distancia (Kms)
Madrid	A Coruña	591
	Asturias	480
	Badajoz	404
	Bilbao	397
	Burgos	244
	León	338
	Logroño	329
	Pamplona	392
	Salamanca	225
	San Sebastián	440
	Santander	410
	Santiago	612
	Valladolid	201

	Vigo	602
Barcelona	Girona	105
	Reus	94
	Zaragoza	299
Palma de Mallorca	Ibiza	150
	Menorca	220
Málaga	Almería	203
	Granada-Jaén	125
	Jerez	227
	Melilla	210
	Sevilla	205
Gran Canaria	El Hierro	368
	Fuerteventura	203
	La Gomera	279
	La Palma	381
	Lanzarote	304
	Tenerife Norte	147
	Tenerife Sur	190
Alicante	Murcia	73
	Valencia	171

Tabla 27. Aeropuertos y distancias

3.4.5 P.I.B. PER CÁPITA POR PAÍSES

En una de las primeras variables explicadas se analizaba la renta per cápita de España, según la provincia en la que se ubicara el aeropuerto; en esta nueva variable se ha creído oportuno analizar de algún modo la renta per cápita, esta vez, de los países que nos rodean (y que más se viaja desde o hacia España); pues es lógico pensar que a mayor o menor poder adquisitivo de los pasajeros de dichos países, mayor o menor puede ser el tráfico aéreo.

En un principio se pensó en hacer una media de la renta entre los primeros países a los que se vuela desde España o con destino España, pero se descartó la idea por no estar en absoluto fundamentada (cada país aporta un número muy diferente de pasajeros).

Después la idea fue sacar un coeficiente relacionado con los pasajeros a/desde cada país (Pasajeros/Renta), pero esto implicaba tener en cuenta la variable dependiente dentro de esta, algo que evidentemente nos llevó a error en el modelo estadístico.

Tras analizar aeropuerto por aeropuerto, quedó de manifiesto que generalmente en cada uno de ellos hay un país que aporta el mayor número de pasajeros (junto con la propia España); así que la decisión final fue introducir en el modelo la renta per cápita del país

que más pasajeros tuviera a/desde cada aeropuerto. Esto también implica fijarse previamente en la variable dependiente (pasajeros), pero no necesariamente en su número concreto, sino solamente ver qué país es el que más aporta y cómo evoluciona a lo largo de los años.

Para algunos aeropuertos el primer país en número de pasajeros transportados es extranjero (no España), y ese es el elegido para el modelo; pero como para la mayoría de aeropuertos el mayor número de pasajeros transportados es España, se ha cogido la renta del primer país extranjero (segundo país después de España).

Para el resto de aeropuertos, o no hay destinos fuera de España, o el número de pasajeros transportados fuera de nuestras fronteras es ínfimo (menor del 10% del total de pasajeros por aeropuerto); por lo que se toma la renta per cápita media de nuestro país.

En la siguiente tabla podemos ver los países con más de un millón de pasajeros tanto en salidas como llegadas desde/a España para el año 2013. Reino Unido y Alemania suman juntas casi los mismos pasajeros que España, siendo los dos países que más aportan con diferencia; así mismo, la mayoría de países de la lista son de la Unión Europea. En otra columna se adjunta el P.I.B. per cápita de cada país (en U.S. \$ a precios actuales).

En el gráfico que sigue a la tabla se pueden observar claras disparidades entre la renta de los distintos países; la bibliografía actual sobre estas disparidades Norte-Sur (que se dan de manera generalizada) es abundante. En Europa existe una no convergencia clara entre Sur y Norte Comunitario, al igual que en España.

PAÍS ESCALA	Pasajeros Total	P.I.B. PER CÁPITA
Total	186.754.749	
ESPAÑA	57.618.090	29.118
REINO UNIDO	32.984.248	39.337
ALEMANIA	23.129.830	45.085
ITALIA	9.657.204	34.619
FRANCIA	9.402.103	41.421
HOLANDA	5.533.002	47.617
SUIZA	4.935.269	80.477
BELGICA	4.452.621	45.387
IRLANDA	3.049.654	47.400
NORUEGA	2.982.836	100.819
SUECIA	2.909.713	58.269
PORTUGAL	2.733.593	21.035
ESTADOS UNIDOS	2.693.382	53.143
FEDERACION RUSA	2.644.166	14.612
DINAMARCA	2.352.360	58.894
FINLANDIA	1.413.787	47.219
AUSTRIA	1.342.257	49.054
MARRUECOS	1.310.439	3.109
POLONIA	1.174.604	13.432
RUMANIA	1.035.521	9.499

Tabla 28. Relación de pasajeros y renta per cápita por países

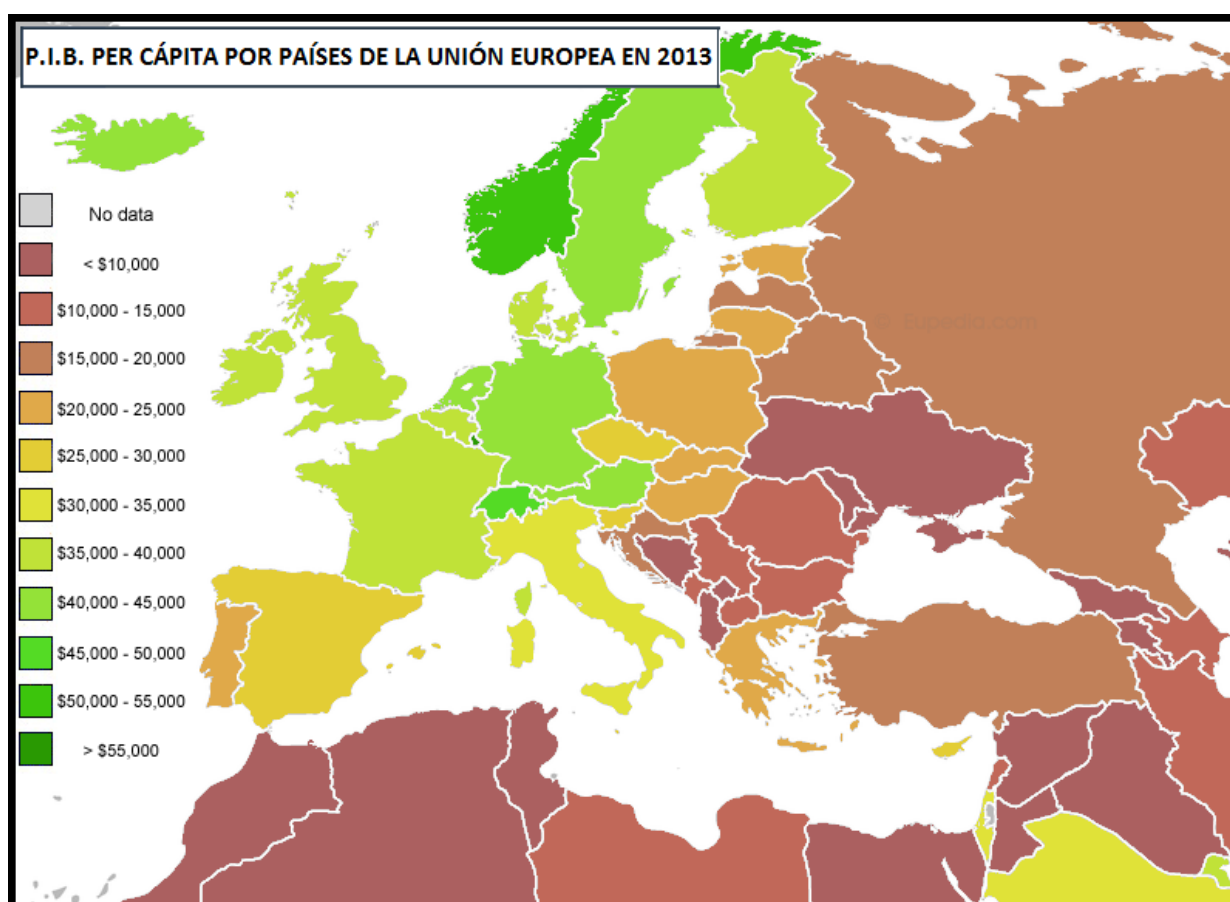


Figura 7. P.I.B. per cápita por países de la UE para 2013

La fuente de los datos es la página web www.bancomundial.org, extraídos en un archivo descargable de su base de datos. En dicha web se pueden consultar todos los países y multitud de años; y en la tabla sucesiva tenemos el país con más pasajeros según el aeropuerto (consultados en la página web de AENA) y la renta per cápita de dicho país en 2013, en base a los criterios anteriormente explicados.

Aeropuerto	País con más pasajeros	P.I.B. per cápita (\$)
A Coruña	Reino Unido	39.337
Alicante	Reino Unido	39.337
Almería	Reino Unido	39.337
Asturias	Reino Unido	39.337
Badajoz	España	29.118
Barcelona	Reino Unido	39.337
Bilbao	Alemania	45.085
Burgos	España	29.118
El Hierro	España	29.118
Fuerteventura	Alemania	45.085
Girona	Alemania	45.085
Gran Canaria	Alemania	45.085
Granada-Jaén	España	29.118
Ibiza	Reino Unido	39.337
Jerez	Alemania	45.085
La Gomera	España	29.118
La Palma	Alemania	45.085
Lanzarote	Reino Unido	39.337
León	España	29.118
Logroño	España	29.118
Madrid Barajas	Italia	34.619
Málaga	Reino Unido	39.337
Melilla	España	29.118
Menorca	Reino Unido	39.337
Murcia	Reino Unido	39.337
Palma de Mallorca	Alemania	45.085
Pamplona	España	29.118
Reus	Reino Unido	39.337
Salamanca	España	29.118
San Sebastián	España	29.118
Santander	Reino Unido	39.337
Santiago	Suiza	80.477
Sevilla	Francia	41.421
Tenerife Norte	España	29.118
Tenerife Sur	Reino Unido	39.337
Valencia	Italia	34.619

Valladolid	España	29.118
Vigo	Francia	41.421
Zaragoza	Italia	34.619

Tabla 29. Variable P.I.B. per cápita por países

3.4.6 COMPAÑÍAS LOW COST CON MÁS PASAJEROS

Las Compañías de Bajo Coste, en su término inglés “Low Cost Carriers”, también llamadas “no frills” o “discount carriers” (sin adornos o de descuento), están consideradas como aerolíneas que ofrecen un servicio a bajo coste que no solo trasladan consigo una oportunidad para el pasajero de transportarse con un menor gasto, sino que traen un nuevo modelo de gestión y de desarrollo comercial, esto conlleva, a un nuevo perfil sociodemográfico del turista, delimitando un nuevo perfil de pasajeros que visita nuestro país. A partir de aquí, la diferenciación de un perfil que aporte a los investigadores y al mundo empresarial más información para el turismo procedente del extranjero, beneficiará las estrategias que se dirijan al futuro turístico de España. Es de notar, que estas compañías han evolucionado en España y en el resto del mundo, consiguiendo muy buenos resultados tanto en el mundo aeronáutico como en el mundo aeroportuario, y claro está, en el propio pasajero. Es por ello interesante analizar el low cost como una variable más del modelo estadístico, aunque como se verá a continuación cuantificar dicha variable es una tarea complicada.

El aumento del protagonismo de las compañías de bajo coste está teniendo consecuencias en los hábitos de la demanda internacional. Este aumento se suma a la mejora de nuevas tecnologías, lo que hace que el panorama turístico internacional este evolucionando a un ritmo muy alto.

El mercado, en este sentido, está en una situación de cambio, pues el número de aerolíneas que operan bajo la etiqueta de “low cost” varía constantemente. Cada año aparecen nuevas compañías, algunas dejan de operar e incluso líneas aéreas clasificadas hasta el momento como tradicionales transforman su estrategia e implantan modelos de gestión y comercialización más próximos a las de bajo coste.

En Septiembre de 2014 los aeropuertos españoles recibieron 7,1 millones de pasajeros internacionales, cifra un 4,8% superior a la registrada hace un año. Las compañías

aéreas de bajo coste (CBC) transportaron el 49,1% de los pasajeros, registrando un avance del 10,5%. Por su parte, las compañías tradicionales fueron responsables del 50,9% del tráfico restante, experimentando una mínima reducción del -0,1% en relación al mismo mes de 2013.

Las compañías aéreas Ryanair, Easyjet y Vueling siguen liderando el ranking de llegadas en CBC, aglutinando entre las tres el 65,6% de los movimientos de pasajeros generados por este tipo de compañías, y el 32,2% del total del flujo aéreo internacional recibido en España. La primera compañía presentó estabilidad respecto al año anterior, y las dos últimas registraron aumentos interanuales de pasajeros.

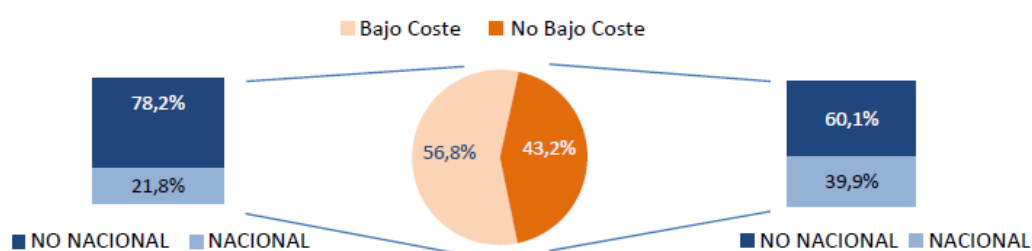
En el periodo enero-septiembre se recibieron cerca de 54 millones de pasajeros internacionales, un 5,9% más que hace un año. Reino Unido, Alemania e Italia originaron algo más de la mitad de este flujo (51,5%). El tráfico de pasajeros de CBC experimentó una expansión (10,3%) notablemente superior al de las compañías tradicionales (2,2%).

En cuanto a las Comunidades Autónomas de destino; Baleares, Cataluña y la C. de Madrid fueron los destinos que más pasajeros por vía aérea concentraron, el 65,4% del total conjuntamente, mostrando esta última, el mayor aumento interanual del 9,7% este mes de 2014.

Canarias y la Comunidad de Madrid recibieron el 19,4% del flujo de pasajeros que viajaron en alguna CBC, y destacaron por ser los destinos que presentaron los mayores crecimientos interanuales este mes.

En el gráfico siguiente podemos observar la evolución del tráfico aéreo por tipo de compañía, entre Enero y Diciembre del año 2013.

DISTRIBUCIÓN DEL TRÁFICO AÉREO EN ESPAÑA POR TIPO DE COMPAÑÍA



EVOLUCIÓN DEL BAJO COSTE EN EL TRÁFICO AÉREO EN ESPAÑA

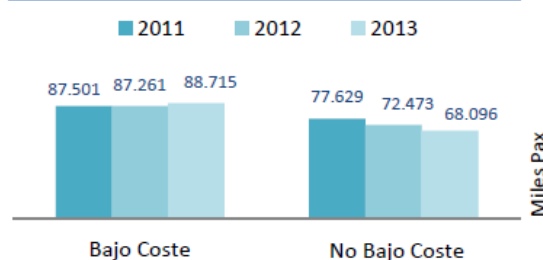


Gráfico 10. Tráfico aéreo en España por tipo de compañía

ESTRUCTURA POR TIPO Y ORIGEN DE LAS COMPAÑÍAS DEL MERCADO AÉREO EN ESPAÑA

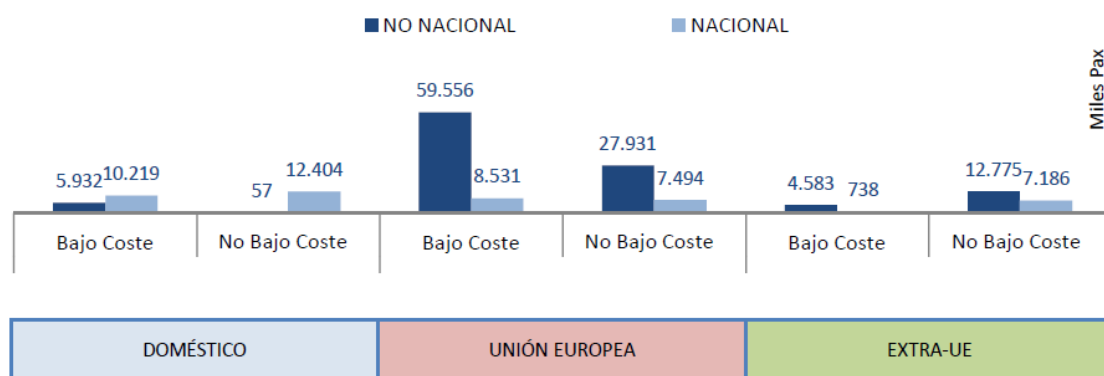


Gráfico 11. Estructura de las compañías del mercado aéreo en España

Como ya se ha podido leer, cuantificar esta variable ha supuesto diversos retos y dificultades, pues no se puede tener en cuenta el número de pasajeros que viajan en low cost, ya que esto supondría introducir la variable dependiente dentro de una independiente; lo cual falsearía los resultados y no daría opciones a una predicción futura.

Además, al igual que ocurre con otras variables de esta parte del proyecto, encontrar datos fiables es muy complicado, pues en cada aeropuerto continuamente aparecen y

desaparecen aerolíneas de bajo coste, incluso algunas de ellas cambian su estrategia de mercado a lo largo del tiempo.

Al tener que relacionar la variable low cost con cada aeropuerto, otra de las opciones barajadas fue analizar cuantos aviones por aerolínea “duermen” en los diversos aeropuertos; pero nuevamente son datos casi imposibles de conseguir de manera pública.

Así que finalmente, investigando en la página web de AENA (fundamental en el presente proyecto) y en diversa bibliografía, se observó que hay 5 compañías de bajo coste que suman aproximadamente el 85% del tráfico aéreo español; estas son (año 2013):

- Ryanair (32,4 millones de pasajeros).
- Vueling (23,5 millones de pasajeros).
- Easyjet (10,5 millones de pasajeros).
- Air Berlin (10 millones de pasajeros).
- Iberia Express (5,8 millones de pasajeros).

Esto se puede ver de manera clara en la distribución de pasajeros por tipo de compañía aérea de bajo coste en Septiembre de 2014, con en el gráfico adjunto a continuación.

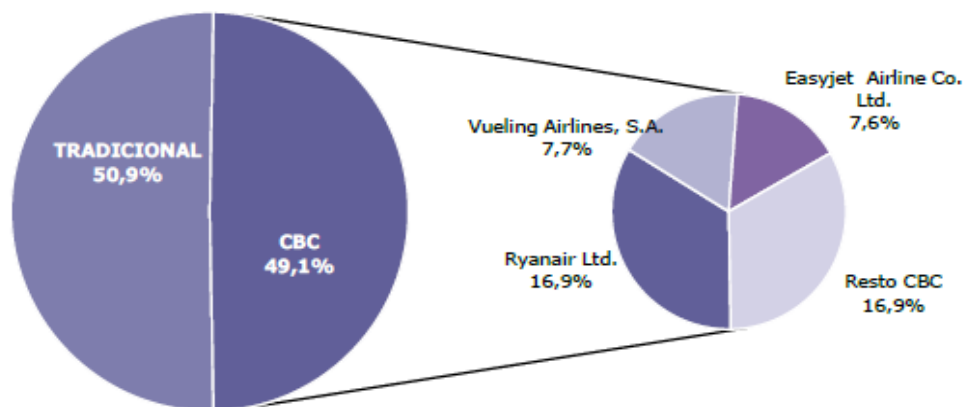


Gráfico 12. Porcentaje de pasajeros por compañía low cost en Septiembre de 2014

Además alguna o varias de estas 5 compañías low cost están presentes en la mayoría de los aeropuertos españoles; por lo que la elección de esta variable fue puntuar a cada aeropuerto de 0 a 5 en función del número de estas que estuviera presente en cada uno de ellos.

Para esto, se ha ido analizando cada aeropuerto en la página web de AENA, sacando en cuales están operativas dichas low cost, y teniendo en cuenta que tengan al menos algún destino fijo y un número representativo de pasajeros por aeropuerto. Un claro ejemplo de lo descrito es el aeropuerto de Zaragoza (entre otros), donde en el año 2013 Ryanair tuvo 272.000 pasajeros y Vueling tan solo 900. De este modo en esta variable se le ha asignado un 1 a este aeropuerto (Ryanair), pues los pasajeros que aportó Vueling no son significativos. Como norma general, para tener o no en cuenta en esta variable a una compañía de bajo coste, se pide que ésta tenga como mínimo entorno al 4-5% del tráfico de pasajeros total del aeropuerto.

Así, en la siguiente tabla se observa cómo queda dicha variable según el aeropuerto para el 2013.

Aeropuerto	Low Cost (5)	Aeropuerto	Low Cost (5)
A Coruña	1	Madrid Barajas	5
Alicante	5	Málaga	5
Almería	4	Melilla	0
Asturias	4	Menorca	4
Badajoz	0	Murcia	2
Barcelona	5	Palma de Mallorca	5
Bilbao	4	Pamplona	1
Burgos	0	Reus	1
El Hierro	0	Salamanca	0
Fuerteventura	5	San Sebastián	1
Girona	1	Santander	2
Gran Canaria	5	Santiago	5
Granada-Jaén	1	Sevilla	5
Ibiza	5	Tenerife Norte	3
Jerez	3	Tenerife Sur	4
La Gomera	0	Valencia	4
La Palma	2	Valladolid	2
Lanzarote	5	Vigo	2
León	0	Zaragoza	1
Logroño	0		

Tabla 30. Variable Compañías low cost con más pasajeros

De manera ilustrativa se muestra la siguiente gráfica, donde se puede ver la principal compañía low cost por aeropuerto, así como el número de pasajeros que tuvo en 2013.

Aeropuerto	1ª Low Cost	Nº de Pasajeros de la 1ª Low Cost
A Coruña	Vueling	368.798
Alicante	Ryanair	2.673.107
Almería	Ryanair	103.580
Asturias	Vueling	268.327
Badajoz	X	0
Barcelona	Vueling	12.172.815
Bilbao	Vueling	1.327.059
Burgos	X	0
El Hierro	X	0
Fuerteventura	Ryanair	592.763
Girona	Ryanair	2.474.201
Gran Canaria	Ryanair	971.990
Granada-Jaén	Vueling	356.144
Ibiza	Ryanair	1.322.843
Jerez	Ryanair	144.969
La Gomera	X	0
La Palma	Air Berlin	37.395
Lanzarote	Ryanair	1.083.335
León	X	0
Logroño	X	0
Madrid Barajas	Ryanair	3.848.096
Málaga	Ryanair	2.980.058
Melilla	X	0
Menorca	Vueling	603.034
Murcia	Ryanair	597.292
Palma de Mallorca	Air Berlin	6.157.373
Pamplona	Vueling	17.509
Reus	Ryanair	622.814
Salamanca	X	0
San Sebastián	Vueling	86.758
Santander	Ryanair	782.588
Santiago	Ryanair	1.029.970
Sevilla	Ryanair	1.464.874
Tenerife Norte	Vueling	394.865
Tenerife Sur	Ryanair	1.484.885
Valencia	Ryanair	1.803.020
Valladolid	Ryanair	144.783
Vigo	Vueling	134.211
Zaragoza	Ryanair	271.723

Tabla 31. Principal compañía low cost por aeropuerto

3.4.7 COMPAÑÍAS LOW COST TOTALES

Siguiendo con la dinámica explicada en la variable anterior, se ha decidido introducir otra nueva relacionada también con las compañías de bajo coste. La anterior pretende ser una buena predicción de que implantando o no una de estas compañías más grandes, el aeropuerto puede tener más o menos pasajeros en un futuro próximo.

Pero, ¿y si en un aeropuerto se implanta una compañía de bajo coste que no pertenece a estas 5 grandes?; para dar respuesta a esta pregunta lo mejor es saber cuántas low cost operan en cada aeropuerto, para poder estudiar el impacto que tienen en el tráfico total de pasajeros; aunque detrás de esto hay otros muchos factores relacionados y para poder analizarlos con total confianza habría que hacer un exhaustivo análisis de mercado.

De nuevo en la múltiple bibliografía analizada, los datos no pueden ser tomados como fiables en su totalidad, pues al igual que antes, varía mucho el número de low cost por aeropuerto según la época consultada; así como la propia definición que tienen las diversas compañías sobre el bajo coste (esto se puede ver tanto en el gráfico como en la tabla del final).

Lo que es abrumante, es la completa evolución del tráfico aéreo según el tipo de compañía (low cost o tradicional) en nuestro país en los últimos 8 años, como se puede ver a continuación (más abajo en número de vuelos).

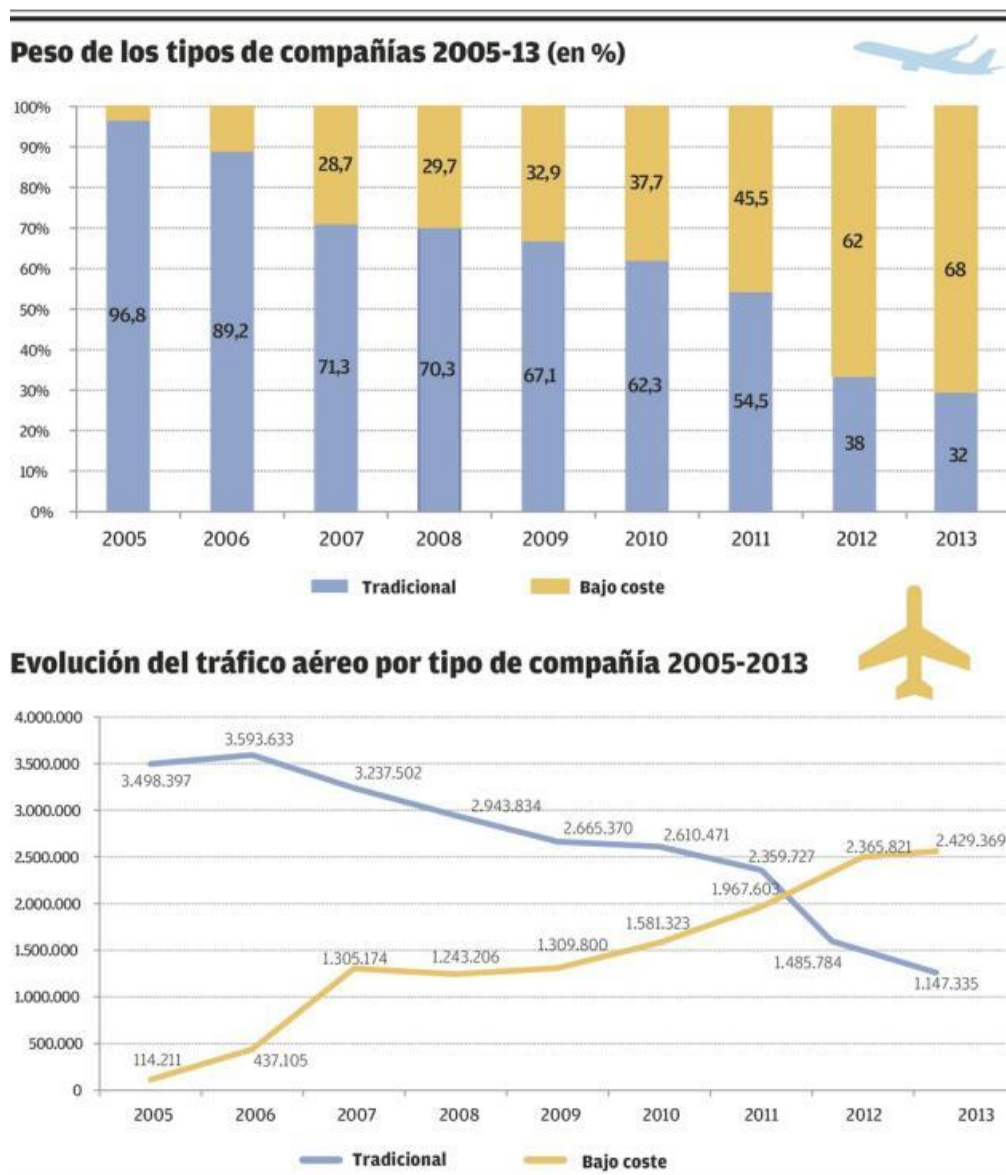


Gráfico 13. Porcentaje por tipos de compañías y evolución según el número de vuelos de 2005 a 2013

Los datos para esta variable se han sacado de la página web de AENA y de www.blavofly.es (sección Aeropuertos-Compañías aéreas de bajo coste que operan en cada uno de ellos), además del Instituto de Estudios Turísticos. A continuación se puede ver un gráfico aclaratorio realizado por el propio IET acerca del número de compañías low cost por aeropuerto español; y debajo la tabla con los datos de la variable que será introducida en el modelo estadístico. Para algún aeropuerto se puede apreciar la no concordancia en número, debido a lo explicado anteriormente.

Además en la tabla “Ranking de compañías que operan en España” dentro de la variable Compañías, se puede ver en qué posición esta cada una de las low cost, así como su número de pasajeros y factor de ocupación por mercados para el año 2013.

Aerolíneas Low Cost en aeropuertos españoles

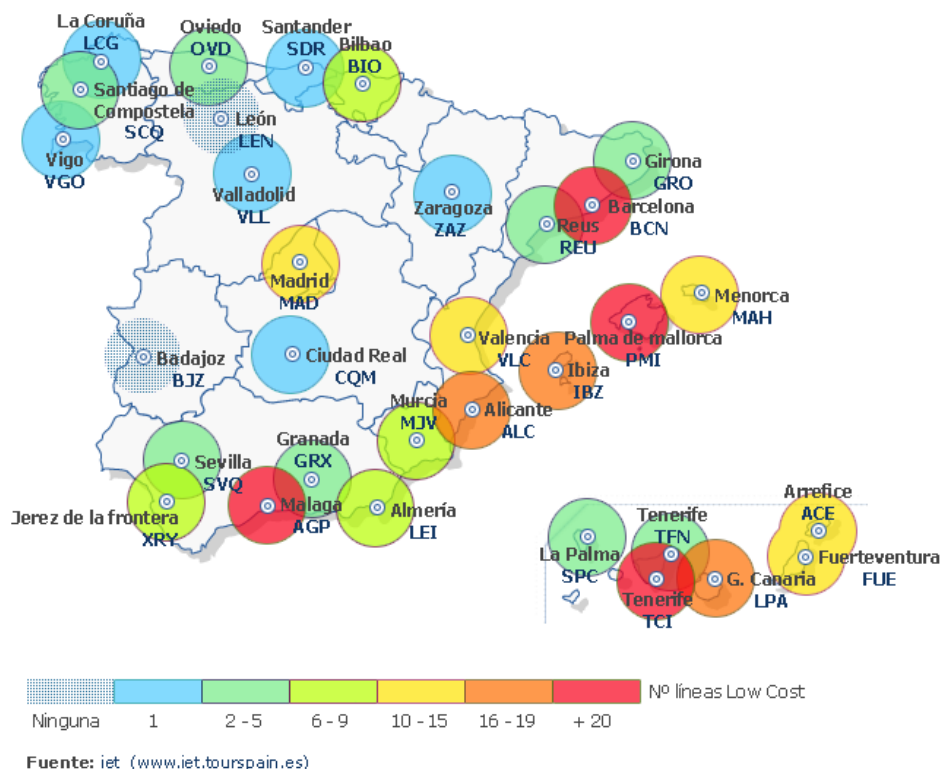


Figura 8. Compañías low cost por aeropuerto

Aeropuerto	Low Cost Totales	Aeropuerto	Low Cost Totales
A Coruña	1	Madrid Barajas	10
Alicante	13	Málaga	18
Almería	8	Melilla	0
Asturias	4	Menorca	11
Badajoz	0	Murcia	7
Barcelona	16	Palma de Mallorca	19
Bilbao	4	Pamplona	1
Burgos	0	Reus	3
El Hierro	0	Salamanca	0
Fuerteventura	11	San Sebastián	1
Girona	4	Santander	2
Gran Canaria	14	Santiago	5
Granada-Jaén	2	Sevilla	5
Ibiza	15	Tenerife Norte	3
Jerez	7	Tenerife Sur	15
La Gomera	0	Valencia	8
La Palma	4	Valladolid	2
Lanzarote	13	Vigo	2
León	0	Zaragoza	1

Logroño	0		
---------	---	--	--

Tabla 32. Variable Compañías low cost totales

3.4.8 HUBS

El modelo de distribución Hub-and-Spoke es un sistema de conexiones que permite reducir el número de rutas para comunicar los aeropuertos entre sí. Consiste en concentrar el tráfico en determinados aeropuertos de gran capacidad o “Hubs” que se encargan de enlazar los de menor capacidad o importancia “Spokes”.

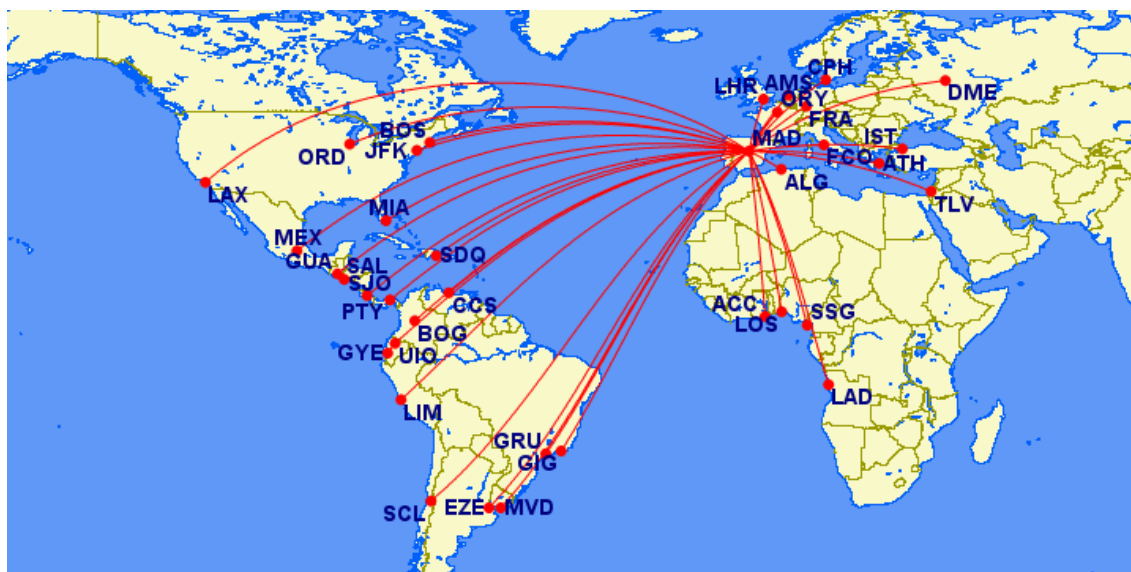


Figura 9. Aeropuerto de Madrid Barajas como hub

Para entender mejor este sistema, se puede considerar como una rueda de bicicleta en la cual todo el tráfico se mueve a lo largo de los radios conectados con el centro.

Los sistemas Hub-and-Spoke no son solo aplicados en la industria de la aviación comercial, sino que desde décadas se lleva usando en la industria del transporte en general, ya sea por mar, tierra o aire.

Los beneficios del uso de este tipo de distribución del tráfico son múltiples, además de mejorar la eficiencia de las comunicaciones, permite concentrar determinadas operaciones que requieren numerosos recursos como el control del flujo de pasajeros y equipajes, en aeropuertos de gran capacidad. Este hecho libera a los aeropuertos más simples de llevar a cabo operaciones puntuales que su sistema aeroportuario no es capaz de llevar a cabo. Además facilita la creación de aeropuertos nuevos.

Por otro lado, este sistema de comunicaciones entre aeropuertos presenta diversos problemas relativos a la flexibilidad de las operaciones. Retrasos tanto en el aeropuerto central como en los exteriores, pueden dar lugar a retrasos en toda la red. Además, la programación del tráfico puede ser muy exigente para los controladores del hub. Se requiere una cuidadosa sincronización para mantener la red funcionando de manera eficiente.

Los sistemas Hub-and-Spoke surgieron como resultado de liberación de las compañías aéreas de las restricciones impuestas por el gobierno llevado a cabo en Estados Unidos (Airline Deregulation Act of 1978).

Esta ley motivó la creación de un nuevo sistema de regulación del tráfico aéreo basado en el sistema Hub-and-Spoke, siguiendo el camino comenzado por algunas compañías como Delta AirLines en la década de los 50.

La situación previa a la implantación de este sistema, se basaba en general en la operación punto a punto, lo cual no era eficiente económicamente puesto que muchos vuelos no llegaban a llenarse.

En la imagen que vemos a continuación se observa una comparación entre la situación de las comunicaciones antes de la ley de 1978 y después, una vez implantado el sistema Hub-and-Spoke.

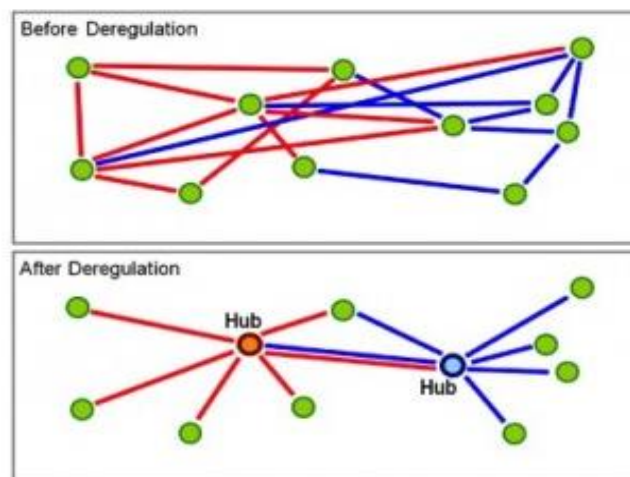


Figura 10. Antes y después del modelo hub and spoke

Hay que destacar que la implantación de este sistema de comunicaciones fue llevado a cabo por las propias aerolíneas, que escogían según diversos criterios los aeropuertos

que consideraban más preparados y mejor situados para convertirlos en aeropuertos hub (no necesariamente los aeropuertos con más tráfico de pasajeros, aunque terminan siéndolo en la mayoría de los casos).

La Terminal 4 de Madrid Barajas sirve como hub para la compañía Iberia, Heathrow en Londres cumple la misma función para British Airways, Amsterdam para KLM, el parisino Charles de Gaulle para Air France o Fráncfort para Lufthansa. El punto en común de estas aerolíneas en la explotación de estos espacios es que, junto a sus socios, operan en exclusividad en su hub.

La posesión de un hub es muy importante para las compañías, porque al centralizar su tráfico en un punto, se ahorran el gasto en “rodadura”, esto es lo que se produce por el traslado desde el punto de aterrizaje hasta la pista (las compañías que no disponen de hub en el aeropuerto dependen de que les adjudique una puerta de desembarco, lo que no siempre ocurre justo al aterrizar, por lo que tienen que estar “paseándose” por la pista, con el gasto en combustible y desgaste de las ruedas que supone).

Además, estos centros de distribución son claves para los aeropuertos porque se convierte en un nodo de conexiones, genera más tráfico de pasajeros y se ve beneficiado porque esa compañía focaliza en ese aeropuerto todas sus inversiones.

Esta variable, al igual que las enclavadas en esta sección, es de difícil análisis en un modelo como el que se pretende hacer; pues su cuantificación es difícil. La forma elegida para hacerlo es asignando un 1 a aquellos aeropuertos considerados como “hub”, y con un 0 al resto.

A su vez, dependiendo de la bibliografía consultada, se considera o no como hub a un mismo aeropuerto. Como apoyo se ha consultado en la página web de AENA para cada aeropuerto, el número de pasajeros en tránsito (aquellos que llegan a un aeropuerto en un vuelo internacional, para continuar su viaje en otro vuelo de conexión internacional sin necesidad de salir del área aeroportuaria).

Aunque AENA solo catalogue a los aeropuertos de Madrid y Barcelona como hubs, en el presente proyecto consideraremos un hub a aquel aeropuerto español donde una compañía aérea haya establecido éste como su base de operaciones (punto de

interconexión de pasajeros y mercancías). Tan solo 4 aeropuertos de los analizados cumplen con este requisito; estos son:

- Madrid Barajas (con Iberia).
- Barcelona (con Vueling).
- Palma de Mallorca (con Air Berlin).
- Gran Canaria (con Tui Fly).

En el siguiente mapa de España tenemos los aeropuertos hubs en rojo, en naranja los no peninsulares, en verde los turísticos peninsulares, en azul los principales peninsulares (no turísticos) y en amarillo los regionales. Además el tamaño del icono cuadrado nos da una idea del número de pasajeros de cada uno de estos. Para los hubs, en 2013, Madrid Barajas tuvo casi 40 millones de pasajeros, seguido de Barcelona (35,2), Palma de Mallorca (22,7) y Gran Canaria (9,7).

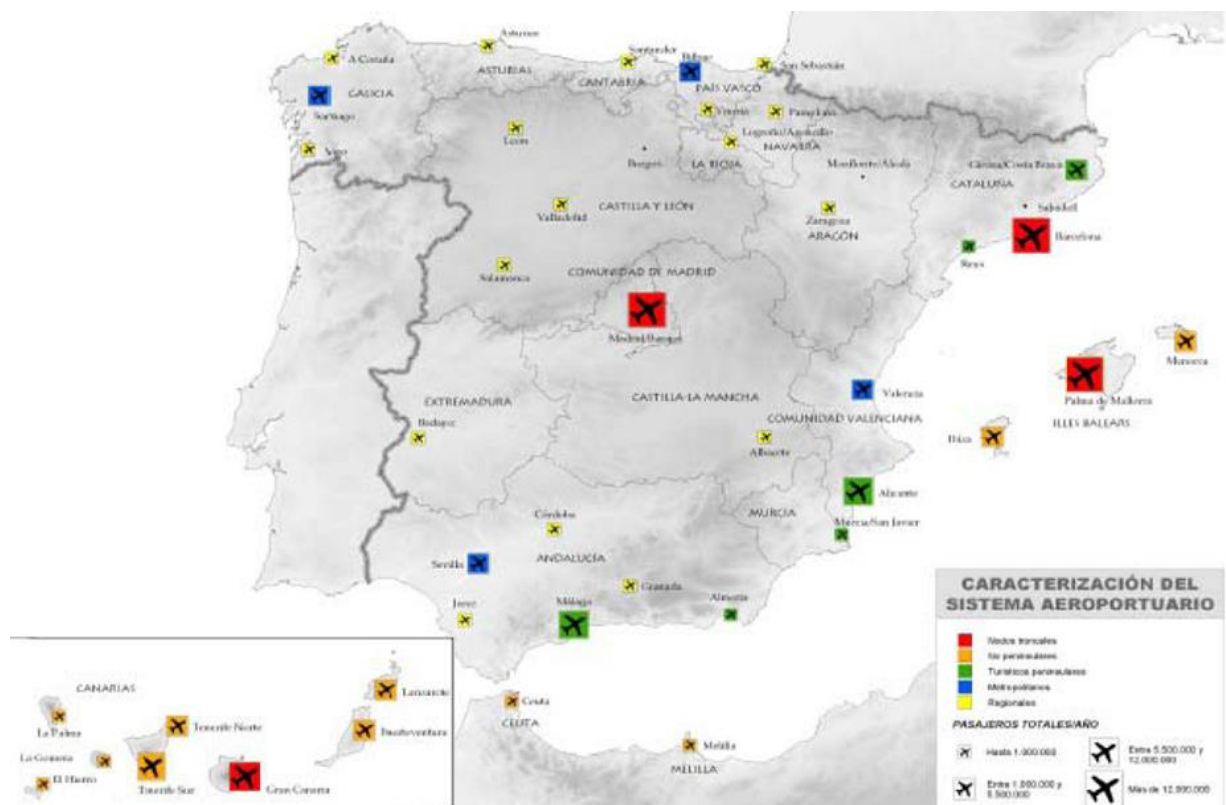


Figura 11. Caracterización de cada aeropuerto español

Como ya se ha explicado, los datos para esta variable introducidos en el modelo del año 2013 son:

Aeropuerto	HUBS	Aeropuerto	HUBS
A Coruña	0	Madrid Barajas	1
Alicante	0	Málaga	0
Almería	0	Melilla	0
Asturias	0	Menorca	0
Badajoz	0	Murcia	0
Barcelona	1	Palma de Mallorca	1
Bilbao	0	Pamplona	0
Burgos	0	Reus	0
El Hierro	0	Salamanca	0
Fuerteventura	0	San Sebastián	0
Girona	0	Santander	0
Gran Canaria	1	Santiago	0
Granada-Jaén	0	Sevilla	0
Ibiza	0	Tenerife Norte	0
Jerez	0	Tenerife Sur	0
La Gomera	0	Valencia	0
La Palma	0	Valladolid	0
Lanzarote	0	Vigo	0
León	0	Zaragoza	0
Logroño	0		

Tabla 33. Variable Hubs

3.4.9 POBLACIÓN EN LOS ENTORNOS AEROPORTUARIOS POR DISTANCIAS DISUASORIAS

Tras haber recopilado los datos de la población de cada aeropuerto por provincias o islas en la variable “Población”, en esta nueva variable se tratará de analizar si puede ser más conveniente acotar la población entorno a las zonas aeroportuarias en función de su capacidad de tráfico aéreo de pasajeros.

Esta variable se ha incluido en esta sección debido a que su estudio será solo para el año 2009, pues los datos han sido obtenidos de un proyecto encontrado en internet, en el cual se detallan las poblaciones para cada aeropuerto según la distancia disuasoria para dicho año. Como se explica a continuación es un proceso tedioso, por lo que se considera suficiente el análisis para el año 2009.

Además, esta variable también es difícil de tratar, pues elegir una franja geográfica disuasoria por cada aeropuerto en función de su volumen de pasajeros, supone saber o presuponer de antemano cuántos viajarán desde un aeropuerto. Aunque puede ser válido si nos fijamos en los pasajeros del año anterior o de toda una época pasada, para fijar dicha zona disuasoria aeroportuaria (ya que realmente no estamos incluyendo en los datos de entrada la variable dependiente).

El procedimiento empleado es el siguiente, se describe un radio de acción alrededor de los aeropuertos que será de 50, 100, 150 o 200Km en función de su importancia (ver número de pasajeros en el año 2009 y anteriores). Este radio debe ser entendido como distancia disuasoria, es decir, una distancia a partir de la cual se presupone que los clientes prefieren otro aeropuerto o simplemente otro medio de transporte.

Dentro de cada área obtenida se hace un barrido, de forma que se registran todos los municipios importantes. Una vez obtenido este listado se efectúa una estimación de la cifra de población que reside en esa área utilizando los datos ofrecidos por el INE (Instituto Nacional de Estadística) para el año 2009 de cada uno de los municipios de España. Para la estimación de población en las áreas de interés se procede sumando o restando poblaciones, es decir, tomando como base el dato de la provincia o provincias que están dentro de dicha área y agregando las cifras de población de los municipios que abarca de provincias colindantes. En cada caso se procede a las combinaciones de sumas y restas de población buscando siempre cometer el mínimo error.

A modo de ejemplo, se expone el caso del aeropuerto de Almería, donde se ha definido una distancia disuasoria de 50Km por tratarse de un aeropuerto de relativa poca importancia dentro del panorama nacional.

En la imagen siguiente se aprecia hasta dónde se extiende esta área. Para la estimación de la población asociada a este aeropuerto se ha procedido de la siguiente manera; en primer lugar se toma como base la población de la provincia de Almería (695.560 habitantes) y en segundo lugar se resta a esa cantidad la suma de las poblaciones de los municipios que están fuera del área azul (Adra, Berja, Láujar de Andrax, Fiñara, Serón, Oria, Macael, Albox, Taberno, Zurgena, Huercal-Ourea, Vera, Los Gallardos y Pulpí) que en este caso resulta 115.618 habitantes. En definitiva, se obtiene una población asociada al aeropuerto de Almería de 579.942 habitantes.



Figura 12. Área disuasoria del aeropuerto de Almería

La imagen anterior ha sido obtenida del software Google Earth, así como las que se muestran a continuación como ejemplos de diferentes radios de áreas disuasorias.



Figura 13. Área disuasoria del aeropuerto de Madrid Barajas



Figura 14. Área disuasoria del aeropuerto de Barcelona

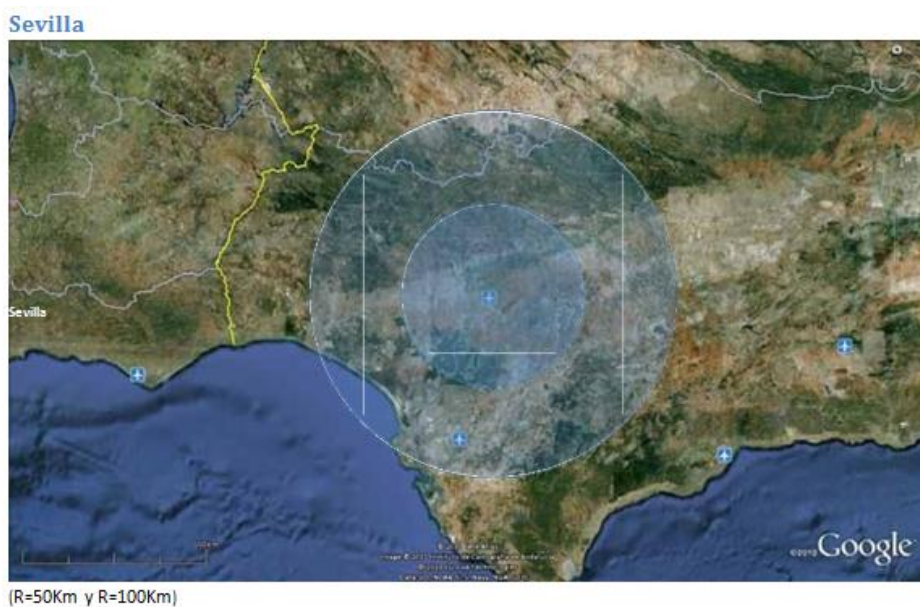


Figura 15. Área disuasoria del aeropuerto de Sevilla



Figura 16. Área disuasoria del aeropuerto de León

En la siguiente tabla se detallan, para el año 2009, las poblaciones obtenidas por entorno aeroportuario y para diferentes distancias con el cálculo explicado.

Aeropuerto	Población	Área Disuasoria (Kms)
A Coruña	1.054.479	50
Alicante	3.449.053	100
Almería	579.942	50
Asturias	1.110.484	100
Badajoz	313.336	50
Barcelona	7.512.381	150
Bilbao	1.494.665	50
Burgos	341.672	50
El Hierro	9.159	50
Fuerteventura	245.105	100
Girona	6.117.931	100
Gran Canaria	838.397	50
Granada-Jaén	3.376.008	100
Ibiza	151.440	50
Jerez	3.382.503	100
La Gomera	19.413	50
La Palma	86.996	50
Lanzarote	245.105	50
León	399.870	50
Logroño	345.700	50
Madrid Barajas	8.604.924	200
Málaga	6.309.484	150

Melilla	76.034	50
Menorca	91.855	50
Murcia	3.208.495	100
Palma de Mallorca	868.152	50
Pamplona	578.580	50
Reus	6.522.657	100
Salamanca	327.282	50
San Sebastián	618.138	50
Santander	577.819	50
Santiago	1.524.254	100
Sevilla	2.704.075	100
Tenerife Norte	1.027.914	100
Tenerife Sur	1.027.914	100
Valencia	3.296.192	100
Valladolid	977.327	100
Vigo	917.494	50
Zaragoza	917.180	50

Tabla 34. Variable Población en los entornos aeroportuarios por distancias disuasorias

El caso de las islas debe tratarse de manera diferente debido a las particularidades geográficas que presentan. En general y para islas de tamaño mediano y pequeño se ha considerado como potenciales clientes sólo los residentes en la propia isla; es el caso de Lanzarote, Hierro, la Palma, la Gomera y Gran Canaria en las Islas Canarias y de Palma de Mallorca y Menorca en Baleares. En el caso de Tenerife, para los dos aeropuertos (Tenerife Norte y Tenerife Sur) se tiene en cuenta la misma población, y se añade además la población de las tres islas más cercanas (Hierro, la Gomera y la Palma). Del mismo modo, para Fuerteventura se ha tenido en cuenta la población de Lanzarote. En el caso del aeropuerto de Ibiza se tiene en cuenta la población de la propia isla así como la de Formentera.

En el caso de Melilla debido a su particularidad fronteriza se utiliza sólo la población de la propia ciudad autónoma.

4. ANÁLISIS DE REGRESIÓN LINEAL

4.1 INTRODUCCIÓN

El análisis de regresión lineal es una técnica estadística utilizada para estudiar la relación entre variables. Se adapta a una amplia variedad de situaciones. En la investigación social, el análisis de regresión se utiliza para predecir un amplio rango de fenómenos, desde medidas económicas hasta diferentes aspectos del comportamiento humano. En el contexto de la investigación de mercados puede utilizarse para determinar en cuál de diferentes medios de comunicación puede resultar más eficaz invertir; o para predecir el número de ventas de un determinado producto. En física se utiliza para caracterizar la relación entre variables o para calibrar medidas. Etc.

Tanto en el caso de dos variables (regresión simple) como en el de más de dos variables (regresión múltiple), el análisis de regresión lineal puede utilizarse para explorar y cuantificar la relación entre una variable llamada dependiente o criterio (Y) y una o más variables llamadas independientes o predictoras ($X_1, X_2, \dots, X_k$), así como para desarrollar una ecuación lineal con fines predictivos. Además, el análisis de regresión lleva asociados una serie de procedimientos de diagnóstico (análisis de los residuos, puntos de influencia) que informan sobre la estabilidad e idoneidad del análisis y que proporcionan pistas sobre cómo perfeccionarlo.

Nuestro objetivo es el de proporcionar los fundamentos del análisis de regresión. Al igual que en los capítulos precedentes, no haremos hincapié en los aspectos más técnicos del análisis, sino que intentaremos fomentar la comprensión de cuándo y cómo utilizar el análisis de regresión lineal, y cómo interpretar los resultados. También prestaremos atención a otras cuestiones como el chequeo de los supuestos del análisis de regresión y la forma de proceder cuando se incumplen.

4.1.1 LA RECTA DE REGRESIÓN

Un diagrama de dispersión ofrece una idea bastante aproximada sobre el tipo de relación existente entre dos variables. Pero, además, un diagrama de dispersión también

puede utilizarse como una forma de cuantificar el grado de relación lineal existente entre dos variables: basta con observar el grado en el que la nube de puntos se ajusta a una línea recta.

Ahora bien, aunque un diagrama de dispersión permite formarse una primera impresión muy rápida sobre el tipo de relación existente entre dos variables, utilizarlo como una forma de cuantificar esa relación tiene un serio inconveniente: la relación entre dos variables no siempre es perfecta o nula; de hecho, habitualmente no es ni lo uno ni lo otro.

Supongamos que disponemos de un pequeño conjunto de datos con información sobre 35 casos y que estamos interesados en estudiar la relación entre la variable Y, y la variable X. Un buen punto de partida para formarnos una primera impresión de esa relación podría ser la representación de la nube de puntos, tal como muestra el diagrama de dispersión de la figura siguiente.

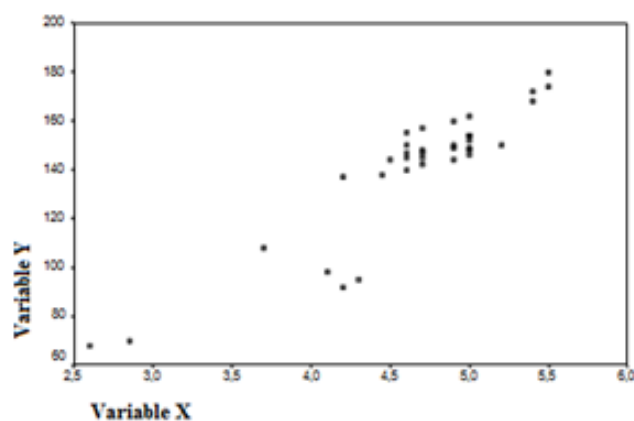


Gráfico 14. Nube de puntos

El eje vertical muestra la variable Y, y el horizontal la variable X (podría ser al revés). A simple vista, parece existir una relación positiva entre ambas variables: conforme aumenta X, aumenta Y.

La mayor parte de los casos de la muestra se agrupan entre el 4,5 y el 5 de la escala de la variable X. Se podría haber extendido el rango de la muestra incluyendo más casos de mayor rango de escala.

¿Cómo podríamos describir los datos que acabamos de proponer? Podríamos decir simplemente que el aumento de una variable va acompañado del aumento de la otra;

pero esto, aunque correcto, es poco específico. ¿Cómo podríamos obtener una descripción más concreta de los resultados? Podríamos, por ejemplo, listar los datos concretos de que disponemos; pero esto, aunque preciso, no resulta demasiado informativo.

Podríamos hacer algo más interesante. Por ejemplo, describir la pauta observada en la nube de puntos mediante una función matemática simple, tal como una línea recta. A primera vista, una línea recta podría ser un buen punto de partida para describir resumidamente la nube de puntos de la figura anterior.

Puesto que una línea recta posee una fórmula muy simple, $Y_i = B_0 + B_1 X_i$, podemos comenzar obteniendo los coeficientes B_0 y B_1 que definen la recta. El coeficiente B_1 es la pendiente de la recta: el cambio medio que se produce en la variable Y (Y_i) por cada unidad de cambio que se produce en la variable X (X_i). El coeficiente B_0 es el punto en el que la recta corta el eje vertical. Conociendo los valores de estos dos coeficientes, nuestro interlocutor podría reproducir la recta y describir con ella la relación existente entre las variables. Aunque no entremos todavía en detalles de cómo obtener los valores de B_0 y B_1 , sí podemos ver cómo es esa recta.

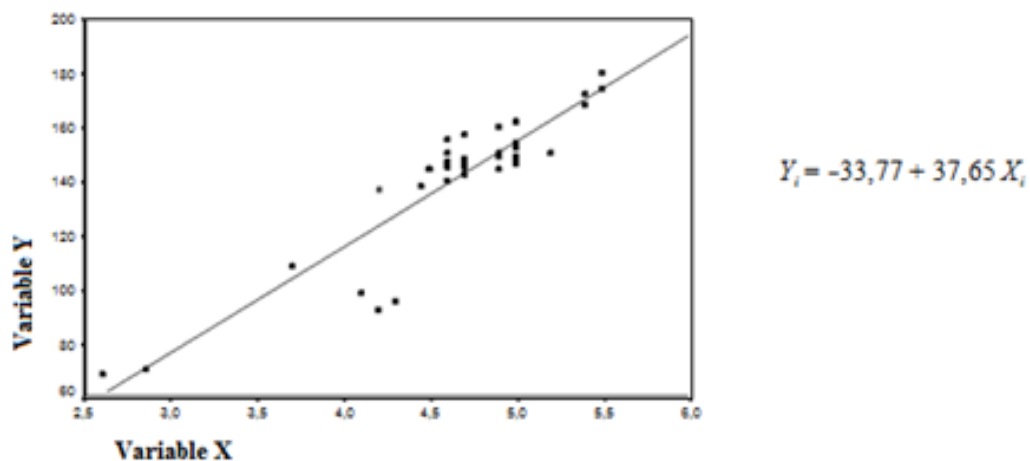


Gráfico 15. Recta de regresión

Vemos que, en general, la recta hace un seguimiento bastante bueno de los datos. La fórmula de la recta aparece a la derecha del diagrama. La pendiente de la recta (B_1) indica que, en promedio, a cada incremento de una unidad de X_i le corresponde un incremento de 37,65 en Y_i . Además, el origen de la recta aporta información sobre lo que podría ocurrir si extrapolamos hacia abajo la pauta observada en los datos hasta

llegar a una variable X cero, al hacer esto estaríamos efectuando pronósticos en un rango de valores que va más allá de lo que abarcan los datos disponibles, y eso es algo extremadamente arriesgado en el contexto del análisis de regresión, sin ampliar el rango de casos de análisis.

4.1.2 LA MEJOR RECTA DE REGRESIÓN

En una situación ideal (e irreal) en la que todos los puntos de un diagrama de dispersión se encontraran en una línea recta, no tendríamos que preocuparnos de encontrar la recta que mejor resume los puntos del diagrama. Simplemente uniendo los puntos entre sí obtendríamos la recta con mejor ajuste a la nube de puntos. Pero en una nube de puntos más realista (como la de la figura) es posible trazar muchas rectas diferentes. Obviamente, no todas ellas se ajustarán igualmente bien a la nube de puntos. Se trata de encontrar la recta capaz de convertirse en el mejor representante del conjunto total de puntos.

Existen diferentes procedimientos para ajustar una función simple, cada uno de los cuales intenta minimizar una medida diferente del grado de ajuste. La elección preferida ha sido, tradicionalmente, la recta que hace mínima la suma de los cuadrados de las distancias verticales entre cada punto y la recta. Esto significa que, de todas las rectas posibles, existe una y sólo una que consigue que las distancias verticales entre cada punto y la recta sean mínimas (las distancias se elevan al cuadrado porque, de lo contrario, al ser unas positivas y otras negativas, se anularían unas con otras al sumarlas).

4.1.3 BONDAD DE AJUSTE

Además de acompañar la recta con su fórmula, podría resultar útil disponer de alguna indicación precisa del grado en el que la recta se ajusta a la nube de puntos. De hecho, la mejor recta posible no tiene por qué ser buena.

Imaginemos una situación como la presentada en el diagrama de abajo, en el que la recta consigue un ajuste bastante más pobre que en el caso de antes, ya que no parece existir la misma pauta de asociación detectada en las variables anteriores. Así pues, aunque siempre resulta posible, cualquiera que sea la nube de puntos, obtener la recta

mínimo-cuadrática, necesitamos información adicional para determinar el grado de fidelidad con que esa recta describe la pauta de relación existente en los datos.

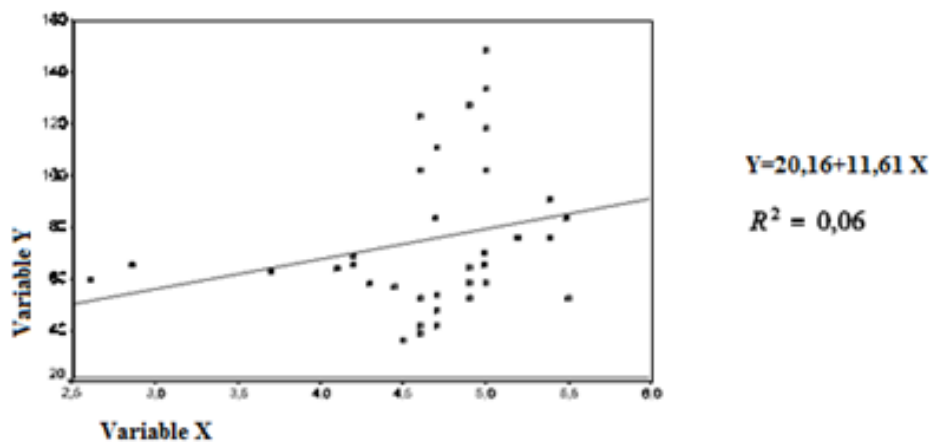


Gráfico 16. Bondad de ajuste

¿Cómo podemos cuantificar ese mejor o peor ajuste de la recta? Hay muchas formas de resumir el grado en el que una recta se ajusta a una nube de puntos. Podríamos utilizar la media de los residuos, o la media de los residuos en valor absoluto, o las medianas de alguna de esas medidas, o alguna función ponderada de esas medidas, etc.

Una medida de ajuste que ha recibido gran aceptación en el contexto del análisis de regresión es el coeficiente de determinación R^2 : el cuadrado del coeficiente de correlación múltiple. Se trata de una medida estandarizada que toma valores entre 0 y 1 (0 cuando las variables son independientes y 1 cuando entre ellas existe relación perfecta).

Este coeficiente posee una interpretación muy intuitiva: representa el grado de ganancia que podemos obtener al predecir una variable basándonos en el conocimiento que tenemos de otra u otras variables.

4.2 ANÁLISIS DE REGRESIÓN LINEAL SIMPLE CON EL PROGRAMA SPSS

4.2.1 BONDAD DE AJUSTE

Puesto que sólo tenemos dos variables, el coeficiente de correlación múltiple no es otra cosa que el valor absoluto del coeficiente de correlación de Pearson entre esas dos variables. Su cuadrado (R cuadrado) es el coeficiente de determinación:

$$R^2 = 1 - \frac{\text{Suma de cuadrados de los residuos}}{\text{Suma de cuadrados total}}$$

(los residuos son las diferencias existentes entre las puntuaciones observadas y los pronósticos obtenidos con la recta). Tal como hemos señalado ya, R^2 expresa la proporción de varianza de la variable dependiente que está explicada por la variable independiente. Tomemos como ejemplo el caso del salario de una serie de empresas (como variable dependiente), y el salario inicial (como variable independiente). Como se puede observar, R toma un valor muy alto (su máximo es 1); y R^2 nos indica que el 77,5 % de la variación de salario está explicada por salario inicial. Es importante resaltar en este momento que el análisis de regresión no permite afirmar que las relaciones detectadas sean de tipo causal: sólo es posible hablar de grado de relación.

Resumen del modelo.

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,880	,775	,774	8.115,36

Tabla 35. Resumen del modelo

R cuadrado corregida es una corrección a la baja de R^2 que se basa en el número de casos y de variables independientes:

$$R_{\text{corregida}}^2 = R^2 - [p(1 - R^2)/(n - p - 1)]$$

(p se refiere al número de variables independientes). En una situación con pocos casos y muchas variables independientes, R^2 puede ser artificialmente alta. En tal caso, el valor de R^2 corregida será sustancialmente más bajo que el de R^2 . En nuestro ejemplo, como hemos introducido 473 casos y una sola variable independiente, los dos valores de R^2 (el corregido y el no corregido) son prácticamente iguales.

El error típico de la estimación (al que llamaremos Se) es la desviación típica de los residuos, es decir, la desviación típica de las distancias existentes entre las puntuaciones

en la variable dependiente (Y_i) y los pronósticos efectuados con la recta de regresión ($\hat{Y}_i$) aunque no exactamente, pues la suma de las distancias al cuadrado están divididas por $n-2$:

$$\text{Error típico de estimación} = S_e = \sqrt{\sum (Y_i - \hat{Y}_i)^2 / (n - 2)}$$

En realidad, este error típico es la raíz cuadrada de la media cuadrática residual de antes. Representa una medida de la parte de variabilidad de la variable dependiente que no es explicada por la recta de regresión. En general, cuanto mejor es el ajuste, más pequeño es este error típico.

Resumen del ANOVA.

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	106831048750,124	1	106831048750,124	1622,118	,000
	Residual	31085446686,216	472	65858997,217		
	Total	137916495436,340	473			

Tabla 36. Resumen del ANOVA

La tabla resumen del ANOVA nos informa sobre si existe o no relación significativa entre las variables. El estadístico F permite contrastar la hipótesis nula de que el valor poblacional de R es cero, lo cual, en el modelo de regresión simple, equivale a contrastar la hipótesis de que la pendiente de la recta de regresión vale cero. El nivel crítico (Sig.) indica que, si suponemos que el valor poblacional de R es cero, es improbable (probabilidad = 0,000) que R, en esta muestra, tome el valor 0,88. Lo cual implica que R es mayor que cero y que, en consecuencia, ambas variables están linealmente relacionadas.

4.2.2 ECUACIÓN DE REGRESIÓN

La tabla siguiente muestra los coeficientes de la recta de regresión. La columna etiquetada Coeficientes no estandarizados contiene los coeficientes de regresión parcial que definen la ecuación de regresión en puntuaciones directas.

Coefficientes de regresión parcial.

	Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.
	B	Error típ.	Beta		
(Constante)	1928,206	888,680		2,170	,031
Salario inicial	1,909	,047	,880	40,276	,000

Tabla 37. Coeficientes parciales

El coeficiente correspondiente a la Constante es el origen de la recta de regresión (lo que hemos llamado B_0):

$$B_0 = \bar{Y} - B_1 \bar{X}$$

Y el coeficiente correspondiente a Salario inicial es la pendiente de la recta de regresión (lo que hemos llamado B_1):

$$B_1 = \frac{\sum X_i Y_i - \sum X_i \sum Y_i}{n \sum X_i^2 - (\sum X_i)^2}$$

B_1 indica el cambio medio que corresponde a la variable dependiente (salario) por cada unidad de cambio de la variable independiente (salario inicial). Según esto, la ecuación de regresión queda de la siguiente manera:

Pronóstico en salario = 1928,206 + 1,909 salario inicial

A cada valor de salario inicial le corresponde un pronóstico en salario basado en un incremento constante (1928,206) más 1,909 veces el valor de salario inicial.

4.2.3 COEFICIENTES DE REGRESIÓN ESTANDARIZADOS

Los coeficientes Beta (coeficientes de regresión parcial estandarizados) son los coeficientes que definen la ecuación de regresión cuando ésta se obtiene tras estandarizar las variables originales, es decir, tras convertir las puntuaciones directas en típicas. Se obtiene de la siguiente manera:

$$\beta_1 = B_1 (S_x / S_y)$$

En el análisis de regresión simple, el coeficiente de regresión estandarizado correspondiente a la única variable independiente presente en la ecuación coincide exactamente con el coeficiente de correlación de Pearson. En regresión múltiple, según veremos, los coeficientes de regresión estandarizados permiten valorar la importancia relativa de cada variable independiente dentro de la ecuación.

4.2.4 PRUEBAS DE SIGNIFICACIÓN

Finalmente, los estadísticos t y sus niveles críticos (Sig.) nos permiten contrastar las hipótesis nulas de que los coeficientes de regresión valen cero en la población. Estos estadísticos t se obtienen dividiendo los coeficientes de regresión B_0 y B_1 entre sus correspondientes errores típicos:

$$t_{B_0} = \frac{B_0}{S_{B_0}} \quad \text{y} \quad t_{B_1} = \frac{B_1}{S_{B_1}}$$

siendo:

$$S_{B_0} = S_e \sqrt{\frac{1}{n} + \frac{\bar{X}^2}{\sum (X_i - \bar{X})^2}} \quad \text{y} \quad S_{B_1} = \frac{S_e}{\sqrt{\sum (X_i - \bar{X})^2}}$$

Estos estadísticos t se distribuyen según el modelo de probabilidad t de Student con n-2 grados de libertad. Por tanto, pueden ser utilizados para decidir si un determinado coeficiente de regresión es significativamente distinto de cero y, en consecuencia, si la variable independiente está significativamente relacionada con la dependiente.

Puesto que en regresión simple sólo trabajamos con una variable independiente, el resultado del estadístico t es equivalente al del estadístico F de la tabla del ANOVA (de hecho, $t^2 = F$).

A partir de los resultados de la tabla anterior, podemos llegar a las siguientes conclusiones:

1. El origen poblacional de la recta de regresión (B_0) es significativamente distinto de cero (generalmente, contrastar la hipótesis " $B_0 = 0$ " carece de utilidad, pues no contiene información sobre la relación entre X_i e Y_i).
2. La pendiente poblacional de la recta de regresión (el coeficiente de regresión B_1 correspondiente a salario inicial) es significativamente distinta de cero, lo cual nos permite concluir que entre salario y salario inicial existe relación lineal significativa.

4.3 ANÁLISIS DE REGRESIÓN LINEAL MÚLTIPLE CON EL PROGRAMA SPSS

El procedimiento Regresión Lineal en el programa estadístico SPSS permite utilizar más de una variable independiente y, por tanto, permite llevar a cabo análisis de regresión múltiple. Pero en el análisis de regresión múltiple, la ecuación de regresión ya no define una recta en el plano, sino un hiperplano en un espacio multidimensional.

Imaginemos un análisis de regresión como antes, con salario como variable dependiente y salario inicial y experiencia previa como variables independientes. La figura adjunta muestra el diagrama de dispersión de salario sobre las variables independientes, y el plano de regresión en un espacio tridimensional.

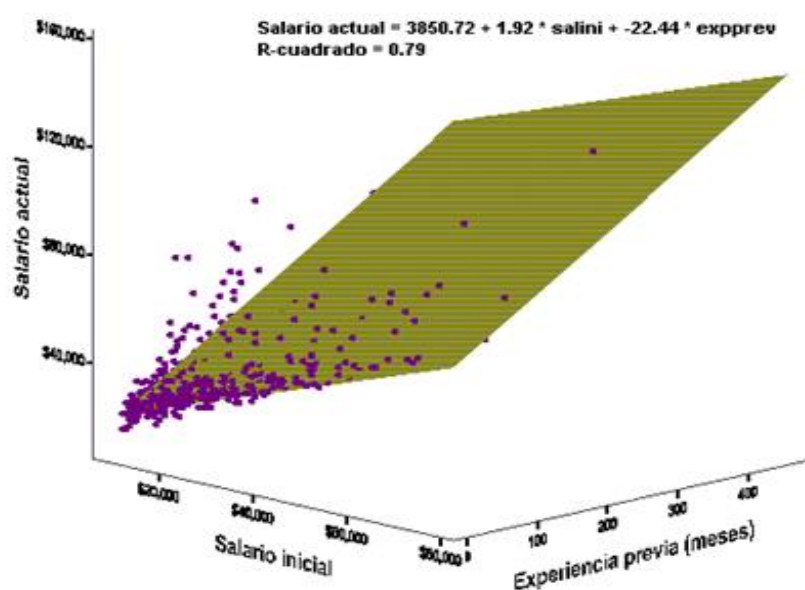


Gráfico 17. Hiperplano

Con una variable dependiente y dos independientes, necesitamos tres ejes para poder representar el correspondiente diagrama de dispersión. Y si en lugar de dos variables independientes utilizáramos tres, sería necesario un espacio de cuatro dimensiones para poder construir el diagrama de dispersión. Y un espacio de cinco dimensiones para poder construir el diagrama correspondiente a cuatro variables independientes. Etc.

Por tanto, con más de una variable independiente, la representación gráfica de las relaciones presentes en un modelo de regresión resulta poco intuitiva, muy complicada y nada útil. Es más fácil y práctico partir de la ecuación del modelo de regresión lineal:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \epsilon$$

De acuerdo con este modelo o ecuación, la variable dependiente (Y) se interpreta como una combinación lineal de un conjunto de K variables independientes (X_k), cada una de las cuales va acompañada de un coeficiente (β_k) que indica el peso relativo de esa variable en la ecuación. La ecuación incluye además una constante (β_0) y un componente aleatorio (los residuos: ϵ) que recoge todo lo que las variables independientes no son capaces de explicar.

En el caso que nos ocupa en este proyecto, el modelo teórico original se expresa en forma no lineal (exponencial) de modo que para abordar su estimación mediante métodos lineales, se “linealiza”, generándose una expresión en logaritmos.

El modelo de demanda no lineal de elasticidades constantes es el siguiente:

$$D = \beta_0 X_1^{\beta_1} X_2^{\beta_2} \dots X_n^{\beta_n}$$

Donde las β_i son las elasticidades de la demanda D. Tomando logaritmos, el modelo resultante es lineal en las elasticidades y su ajuste es trivial (pasa por resolver un sistema lineal de ecuaciones):

$$\log D = \log \beta_0 + \beta_1 \log X_1 + \beta_2 \log X_2 + \dots + \beta_n \log X_n$$

Por tanto los datos que introducimos en el modelo deben tener aplicados logaritmos, y si una vez halladas las elasticidades, quisiéramos predecir futuros modelos de demanda deshacer dichos logaritmos.

Al aplicar logaritmos en un modelo de demanda, se observa como las correlaciones entre las nuevas variables siguen una distribución normal, ya que sus nubes de puntos en un gráfico tienen una forma aproximadamente “elíptica”. Por tanto $\log(X_i)$ sigue una distribución normal.

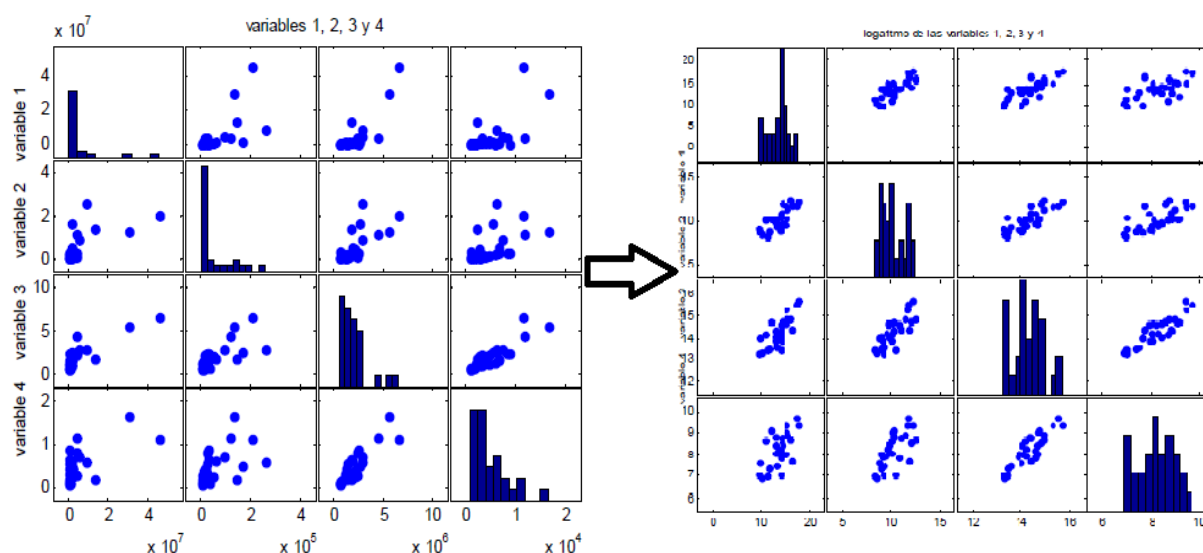


Gráfico 18. Nubes de puntos por variables tras aplicar logaritmos a estas

Este modelo, al igual que cualquier otro modelo estadístico, se basa en una serie de supuestos (linealidad, independencia, normalidad, homocedasticidad y no-colinealidad) que estudiaremos en detalle más adelante.

La ecuación de regresión mínimo-cuadrática se construye estimando los valores de los coeficientes beta del modelo de regresión. Estas estimaciones se obtienen intentando hacer que las diferencias al cuadrado entre los valores observados (Y) y los pronosticados ($\hat{Y}$) sean mínimas:

$$\hat{Y} = B_0 + B_1 X_1 + B_2 X_2 + \dots + B_k X_k$$

Esta es la notación que seguiremos usando para la explicación estadística, pero como se ha explicado antes con la transformación en logaritmos, los términos de esta ecuación para nuestro modelo de demanda son: $Y = \log Y$, y $X_i = \log X_i$.

Al igual que en el análisis de regresión simple del apartado anterior, vamos a seguir utilizando salario (salario actual) como variable dependiente. Pero ahora vamos a incluir 3 variables independientes en el modelo: salini (salario inicial), expprev (experiencia previa) y educ (nivel educativo).

4.3.1 BONDAD DE AJUSTE

Tomadas juntas las tres variables independientes incluidas en el análisis explican un 80 % de la varianza de la variable dependiente, pues R^2 corregida = 0,80. Además, el error típico de los residuos (8.115,36 en el análisis de regresión simple) ha disminuido algo (7.631,92 en el análisis de regresión múltiple), lo que indica una pequeña mejora en el ajuste. De nuevo, el valor corregido de R^2 es casi idéntico al valor no corregido.

Resumen del modelo.

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,895	,802	,800	7,631.92

Tabla 38. Resumen del modelo

El estadístico F de la tabla siguiente contrasta la hipótesis nula de que el valor poblacional de R es cero y, por tanto, nos permite decidir si existe relación lineal significativa entre la variable dependiente y el conjunto de variables independientes tomadas juntas. El valor del nivel crítico Sig. = 0,000 indica que sí existe relación lineal significativa. Podemos afirmar, por tanto, que el hiperplano definido por la ecuación de regresión ofrece un buen ajuste a la nube de puntos.

Resumen del ANOVA.

	Suma de cuadrados	gl	Media cuadrática	F	Sig.
Regresión	110540801465,350	3	36846933821,783	632,607	,000
Residual	27375693970,990	470	58246157,385		
Total	137916495436,340	473			

Tabla 39. Resumen del ANOVA

4.3.2 ECUACIÓN DE REGRESIÓN

La tabla de coeficientes de regresión parcial de la próxima tabla contiene toda la información necesaria para construir la ecuación de regresión mínimo-cuadrática.

Coeficientes de regresión parcial.

	Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.
	B	Error típ.	Beta		
(Constante)	-3661,517	1935,490		-1,892	,059
Salario inicial	1,749	,060	,806	29,198	,000
Experiencia previa (meses)	-16,730	3,605	-,102	-4,641	,000
Nivel educativo (años)	735,956	168,689	,124	4,363	,000

Tabla 40. Coeficientes parciales

En la columna encabezada Coeficientes no estandarizados se encuentran los coeficientes (B_k) que forman parte de la ecuación en puntuaciones directas:

$$\text{Pronóstico en salario} = -3.661,517 + 1,749 \text{ salini} - 16,730 \text{ expprev} + 735,956 \text{ educ}$$

Estos coeficientes no estandarizados se interpretan en los términos ya conocidos. Por ejemplo, el coeficiente correspondiente a la variable salini, que vale 1,749, indica que, si el resto de variables se mantienen constantes, a un aumento de una unidad (un euro) en salini le corresponde, en promedio, un aumento de 1,749 euros en salario.

Es necesario señalar que estos coeficientes no son independientes entre sí. De hecho, reciben el nombre de coeficientes de regresión parcial porque el valor concreto estimado para cada coeficiente se ajusta teniendo en cuenta la presencia del resto de variables independientes. Conviene, por tanto, interpretarlos con cautela.

El signo del coeficiente de regresión parcial de una variable puede no ser el mismo que el del coeficiente de correlación simple entre esa variable y la dependiente. Esto es debido a los ajustes que se llevan a cabo para poder obtener la mejor ecuación posible. Aunque existen diferentes explicaciones para justificar el cambio de signo de un coeficiente de regresión, una de las que deben ser más seriamente consideradas es la que se refiere a la presencia de un alto grado de asociación entre algunas de las variables independientes (colinealidad). Trataremos esta cuestión más adelante.

4.3.3 COEFICIENTES DE REGRESIÓN ESTANDARIZADOS

Los coeficientes Beta están basados en las puntuaciones típicas y, por tanto, son directamente comparables entre sí. Indican la cantidad de cambio, en puntuaciones típicas, que se producirá en la variable dependiente por cada cambio de una unidad en la correspondiente variable independiente (manteniendo constantes el resto de variables independientes).

Estos coeficientes proporcionan una pista muy útil sobre la importancia relativa de cada variable independiente en la ecuación de regresión. En general, una variable tiene tanto más peso (importancia) en la ecuación de regresión cuanto mayor (en valor absoluto) es su coeficiente de regresión estandarizado. Observando los coeficientes Beta de la tabla anterior vemos que la variable salini es la más importante; después, educ; por último, expprev. Lo ya dicho sobre la no independencia de los coeficientes de regresión parcial no estandarizados también vale aquí.

4.3.4 PRUEBAS DE SIGNIFICACIÓN

Las pruebas t y sus niveles críticos (últimas dos columnas de la tabla anterior: t y Sig.) sirven para contrastar la hipótesis nula de que un coeficiente de regresión vale cero en la población. Niveles críticos (Sig.) muy pequeños (generalmente menores que 0,05) indican que debemos rechazar esa hipótesis nula.

Un coeficiente de cero indica ausencia de relación lineal, de modo que los coeficientes significativamente distintos de cero nos informan sobre qué variables son relevantes en la ecuación de regresión. Observando el nivel crítico asociado a cada prueba t, vemos que las tres variables utilizadas poseen coeficientes significativamente distintos de cero (en todas, Sig. = 0,000). Todas ellas, por tanto, contribuyen de forma significativa a explicar lo que ocurre con la variable dependiente.

4.3.5 INFORMACIÓN COMPLEMENTARIA

Además de la ecuación de regresión y de la calidad de su ajuste, un análisis de regresión no debe renunciar a la obtención de algunos estadísticos descriptivos elementales como la matriz de correlaciones, la media y la desviación típica de cada variable y el número de casos con el que se está trabajando, etc.

Estimaciones. Ofrece las estimaciones de los coeficientes de regresión parcial no estandarizados (B) y estandarizados (Beta), junto con las pruebas de significación t individuales para contrastar las hipótesis de que el valor poblacional de esos coeficientes es cero.

Ajuste del modelo. Muestra el coeficiente de correlación múltiple, su cuadrado corregido y no corregido, y el error típico de los residuos. Esta opción también incluye la tabla resumen del ANOVA, la cual contiene el estadístico F para contrastar la hipótesis $R = 0$.

Intervalos de confianza. Estos intervalos nos informan sobre los límites entre los que podemos esperar que se encuentre el valor poblacional de cada coeficiente de regresión. Los límites se obtienen sumando y restando 1,96 errores típicos al valor del correspondiente coeficiente de regresión (decimos 1,96 porque el SPSS trabaja, por defecto, con un nivel de confianza de 0,95).

Intervalos de confianza muy amplios indican que las estimaciones obtenidas son poco precisas y, probablemente, inestables (cosa que suele ocurrir, por ejemplo, cuando existen problemas de colinealidad; estudiaremos esta cuestión más adelante, en el apartado dedicado a los supuestos del modelo de regresión).

Coeficientes de regresión parcial, incluyendo los Intervalos de confianza.

	Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.	Intervalo de confianza para B al 95%	
	B	Error típ.	Beta			Límite inferior	Límite superior
(Constante)	-3661,5	1935,490		-1,892	,059	-7464,803	141,768
Salario inicial	1,749	,060	,806	29,198	,000	1,631	1,866
Experiencia previa	-16,730	3,605	-,102	-4,641	,000	-23,814	-9,646
Nivel educativo	735,956	168,689	,124	4,363	,000	404,477	1067,434

Tabla 41. Coeficientes parciales e intervalos de confianza

Matriz de covarianza. Muestra una matriz con las covarianzas y correlaciones existentes entre los coeficientes de regresión parcial. Vemos que, efectivamente, los coeficientes de regresión parcial no son independientes entre sí.

Correlaciones entre los coeficientes de regresión.

		Nivel educativo	Experiencia previa	Salario inicial
Correlaciones	Nivel educativo	1,000	,363	-,667
	Experiencia previa	,363	1,000	-,274
	Salario inicial	-,667	-,274	1,000
Covarianzas	Nivel educativo	28456,057	220,958	-6,737
	Experiencia previa	220,958	12,997	-5,908E-02
	Salario inicial	-6,737	-5,908E-02	3,587E-03

Tabla 42. Correlaciones

Descriptivos. Ofrece la media y la desviación típica de cada variable y el número de casos utilizados en el análisis.

Además, ofrece la matriz de correlaciones entre el conjunto de variables utilizadas en el análisis. En la matriz de correlaciones, cada coeficiente de correlación aparece acompañado de su correspondiente nivel crítico (el cual permite decidir sobre la hipótesis de que el coeficiente de correlación vale cero en la población) y del número de casos sobre el que se ha calculado cada coeficiente.

Lógicamente, en la diagonal de la matriz de correlaciones aparecen unos, pues la relación entre una variable y ella misma es perfecta.

Estadísticos descriptivos.

		Media	Desviación típ.	N
Salario	actual	34,419.57	17,075.66	474
Salario	inicial	17,016.09	7,870.64	474
Experiencia	previa	95,86	104,59	474
Nivel educativo		13,49	2,88	474

Tabla 43. Estadísticos descriptivos

Correlaciones entre variables.

		Salario actual	Salario inicial	Experiencia previa	Nivel educativo
Salario actual	Correlación de Pearson	1,000	,880	-,097	,661
	Sig. (unilateral)		,000	,017	,000
	N	474	474	474	474
Salario inicial	Correlación de Pearson	,880	1,000	,045	,633
	Sig. (unilateral)	,000		,163	,000
	N	474	474	474	474
Experiencia previa	Correlación de Pearson	-,097	,045	1,000	-,252
	Sig. (unilateral)	,017	,163		,000
	N	474	474	474	474
Nivel educativo	Correlación de Pearson	,661	,633	-,252	1,000
	Sig. (unilateral)	,000	,000	,000	
	N	474	474	474	474

Tabla 44. Correlaciones

Correlaciones parcial y semiparcial. Esta opción permite obtener los coeficientes de correlación parcial y semiparcial entre la variable dependiente y cada variable independiente. Un coeficiente de correlación parcial expresa el grado de relación existente entre dos variables tras eliminar de ambas el efecto debido a terceras variables. En el contexto del análisis de regresión, los coeficientes de correlación parcial expresan el grado de relación existente entre cada variable independiente y la variable dependiente tras eliminar de ambas el efecto debido al resto de variables independientes incluidas en la ecuación.

Un coeficiente de correlación semiparcial expresa el grado de relación existente entre dos variables tras eliminar de una de ellas el efecto debido a terceras variables. En el contexto del análisis de regresión, estos coeficientes expresan el grado de relación existente entre la variable dependiente y la parte de cada variable independiente que no está explicada por el resto de variables independientes.

Coeficientes de regresión parcial y coeficientes de correlación parcial y semiparcial.

	Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.	Correlaciones		
	B	Error típ.	Beta			Orden cero	Parcial	Semi-parcial
(Constante)	3661,517	1935,490		-1,892	,059			
Salario inicial	1,749	,060	,806	29,198	,000	,880	,803	,600
Experiencia previa	-16,730	3,605	-,102	-4,641	,000	-,097	-,209	-,095
Nivel educativo	735,956	168,689	,124	4,363	,000	,661	,197	,090

Tabla 45. Coeficientes parciales y semiparciales

Junto con los coeficientes de correlación parcial y semiparcial, aparecen las correlaciones de orden cero, es decir, los coeficientes de correlación calculados sin tener en cuenta la presencia de terceras variables. Comparando entre sí estos coeficientes (de orden cero, parcial y semiparcial) pueden encontrarse pautas de relación interesantes. En los datos de la tabla anterior ocurre, por ejemplo, que la relación entre la variable dependiente salario actual y la variable independiente nivel educativo vale 0,661. Sin embargo, al eliminar de salario actual y de nivel educativo el efecto atribuible al resto de variables independientes (salario inicial y experiencia previa), la relación baja hasta 0,197 (parcial); y cuando el efecto atribuible a salario inicial y experiencia previa se elimina sólo de salario actual, la relación baja hasta 0,090 (semiparcial). Lo cual está indicando que la relación entre estas dos últimas variables podría ser espúrea, pues puede explicarse casi por completo recurriendo a las otras dos variables independientes.

4.3.6 SUPUESTOS DEL MODELO DE REGRESIÓN LINEAL

Los supuestos de un modelo estadístico se refieren a una serie de condiciones que deben darse para garantizar la validez del modelo. Al efectuar aplicaciones prácticas del modelo de regresión, nos veremos en la necesidad de examinar muchos de estos supuestos.

1. Linealidad. La ecuación de regresión adopta una forma particular. En concreto, la variable dependiente es la suma de un conjunto de elementos: el origen de la recta, una combinación lineal de variables independientes o predictoras y los residuos. El incumplimiento del supuesto de linealidad suele denominarse error de especificación. Algunos ejemplos son: omisión de variables independientes importantes, inclusión de variables independientes irrelevantes, no linealidad (la relación entre las variables independientes y la dependiente no es lineal), parámetros cambiantes (los parámetros no permanecen constantes durante el tiempo que dura la recogida de datos), no aditividad (el efecto de alguna variable independiente es sensible a los niveles de alguna otra variable independiente), etc.

2. Independencia. Los residuos son independientes entre sí, es decir, los residuos constituyen una variable aleatoria (recordemos que los residuos son las diferencias entre

los valores observados y los pronosticados). Es frecuente encontrarse con residuos autocorrelacionados cuando se trabaja con series temporales.

3. Homocedasticidad. Para cada valor de la variable independiente (o combinación de valores de las variables independientes), la varianza de los residuos es constante.

4. Normalidad. Para cada valor de la variable independiente (o combinación de valores de las variables independientes), los residuos se distribuyen normalmente con media cero.

5. No-colinealidad. No existe relación lineal exacta entre ninguna de las variables independientes. El incumplimiento de este supuesto da origen a colinealidad o multicolinealidad.

Sobre el cumplimiento del primer supuesto puede obtenerse información a partir de una inspección del diagrama de dispersión: si tenemos intención de utilizar el modelo de regresión lineal, lo razonable es que la relación entre la variable dependiente y las independientes sea de tipo lineal (veremos que existen gráficos parciales que permiten obtener una representación de la relación neta existente entre dos variables). El quinto supuesto, no-colinealidad, no tiene sentido en regresión simple, pues es imprescindible la presencia de más de una variable independiente. Veremos que existen diferentes formas de diagnosticar la presencia de colinealidad. El resto de los supuestos, independencia, homocedasticidad y normalidad, están estrechamente asociados al comportamiento de los residuos. Por tanto, un análisis cuidadoso de los residuos puede informarnos sobre el cumplimiento de los mismos.

4.3.7 ANÁLISIS DE LOS RESIDUOS

Llamamos residuos a las diferencias entre los valores observados y los pronosticados: $(Y_i - \hat{Y}_i)$.

Los residuos son muy importantes en el análisis de regresión. En primer lugar, nos informan sobre el grado de exactitud de los pronósticos: cuanto más pequeño es el error típico de los residuos, mejores son los pronósticos, o lo que es lo mismo, mejor se ajusta la recta de regresión a la nube de puntos. En segundo lugar, el análisis de las

características de los casos con residuos grandes (sean positivos o negativos; es decir, grandes en valor absoluto) puede ayudarnos a detectar casos atípicos y, consecuentemente, a perfeccionar la ecuación de regresión a través de un estudio detallado de los mismos.

Independencia

El verdadero interés de los residuos hay que buscarlo en el hecho de que el análisis de los mismos nos proporciona información crucial sobre el cumplimiento de varios supuestos del modelo de regresión lineal: independencia, homocedasticidad, normalidad y linealidad.

Uno de los supuestos básicos del modelo de regresión lineal es el de independencia entre los residuos (supuesto éste particularmente relevante cuando los datos se han recogido siguiendo una secuencia temporal). El estadístico de Durbin-Watson (1951) proporciona información sobre el grado de independencia existente entre ellos:

$$DW = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2}$$

(e_i se refiere a los residuos: $e_i = Y_i - \hat{Y}_i$). El estadístico DW oscila entre 0 y 4, y toma el valor 2 cuando los residuos son independientes. Los valores menores que 2 indican autocorrelación positiva y los mayores que 2 autocorrelación negativa. Podemos asumir independencia entre los residuos cuando DW toma valores entre 1,5 y 2,5.

Resumen del modelo.

R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
.895	.802	.800	7.631.92	1.579

Tabla 46. Resumen del modelo

Puesto que el valor $DW = 1,579$ se encuentra entre 1,5 y 2,5, podemos asumir que los residuos son independientes.

Homocedasticidad

El procedimiento Regresión lineal dispone de una serie de gráficos que permiten, entre otras cosas, obtener información sobre el grado de cumplimiento de los supuestos de homocedasticidad y normalidad de los residuos.

Las variables elegidas permiten obtener diferentes gráficos de dispersión en SPSS, dependiendo de una de las siguientes formas:

DEPENDENT: variable dependiente de la ecuación de regresión.

ZPRED (pronósticos tipificados): pronósticos divididos por su desviación típica. Son pro-nósticos transformados en puntuaciones z (con media 0 y desviación típica 1).

ZRESID (residuos tipificados): residuos divididos por su desviación típica. El tamaño de cada residuo tipificado indica el número de desviaciones típicas que se aleja de su media, de modo que, si están normalmente distribuidos (cosa que asumimos en el análisis de regresión), el 95 % de estos residuos se encontrará en el rango $(-1,96, +1,96)$, lo cual permite identificar fácilmente casos con residuos grandes.

DRESID (residuos eliminados o corregidos): residuos obtenidos al efectuar los pronósticos eliminando de la ecuación de regresión el caso sobre el que se efectúa el pronóstico. El residuo correspondiente a cada caso se obtiene a partir del pronóstico efectuado con una ecuación de regresión en la que no se ha incluido ese caso. Son muy útiles para detectar puntos de influencia (casos con gran peso en la ecuación de regresión).

ADJPRED (pronósticos corregidos): pronósticos efectuados con una ecuación de regresión en la que no se incluye el caso pronosticado (ver residuos eliminados o corregidos). Diferencias importantes entre PRED y ADJPRED delatan la presencia de puntos de influencia (casos con gran peso en la ecuación de regresión).

SRESID (residuos estudentizados): residuos divididos por su desviación típica, basada ésta en cómo de próximo se encuentra un caso a su(s) media(s) en la(s) variable(s) independiente(s). Al igual que ocurre con los residuos estandarizados (a los que se parecen mucho), los estudentizados están escalados en unidades de desviación típica. Se distribuyen según el modelo de probabilidad t de Student con $n-p-1$ grados de libertad

(p se refiere al número de variables independientes). Con muestras grandes, aproximadamente el 95 % de estos residuos debería encontrarse en el rango $(-2, +2)$.

SDRESID (residuos corregidos estudentizados): residuos corregidos divididos por su desviación típica. Útiles también para detectar puntos de influencia.

Algunas de estas variables permiten identificar puntos de influencia (los estudiaremos más adelante), pero hay, entre otras, dos variables cuyo diagrama de dispersión informa sobre el supuesto de homocedasticidad o igualdad de varianzas: ZPRED y ZRESID. El supuesto de igualdad de varianzas implica que la variación de los residuos debe ser uniforme en todo el rango de valores pronosticados. O, lo que es lo mismo, que el tamaño de los residuos es independiente del tamaño de los pronósticos, de donde se desprende que el diagrama de dispersión no debe mostrar ninguna pauta de asociación entre los pronósticos y los residuos.

Observando el diagrama de dispersión de abajo podemos ver que, aunque los residuos y los pronósticos parecen ser independientes (pues la nube de puntos no sigue ninguna pauta de asociación clara, ni lineal ni de otro tipo), no está claro que las varianzas sean homogéneas. Más bien parece que conforme va aumentando el valor de los pronósticos también lo va haciendo la dispersión de los residuos: los pronósticos menores que la media (con puntuación típica por debajo de cero) están más concentrados que los pronósticos mayores que la media (con puntuación típica mayor que cero).

Diagrama de dispersión de *pronósticos tipificados* por *residuos tipificados*.

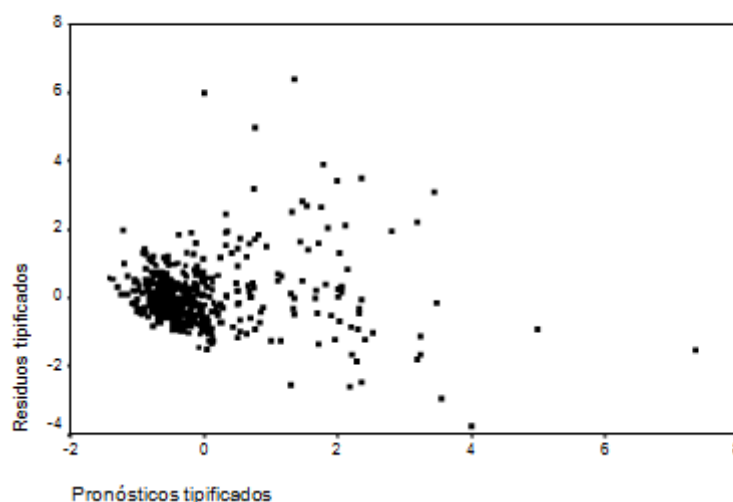


Gráfico 19. Pronósticos tipificados VS residuos tipificados

Cuando un diagrama de dispersión delata la presencia de varianzas heterogéneas, puede utilizarse una transformación de la variable dependiente para resolver el problema (tal como una transformación logarítmica o una transformación raíz cuadrada). No obstante, al utilizar una transformación de la variable dependiente, no deben descuidarse los problemas de interpretación que añade el cambio de escala.

El diagrama de dispersión de las variables ZPRED y ZRESID posee la utilidad adicional de permitir detectar relaciones de tipo no lineal entre las variables. Si la relación es, de hecho, no lineal, el diagrama puede contener indicios sobre otro tipo de función de ajuste: por ejemplo, los residuos estandarizados podrían, en lugar de estar homogéneamente dispersos, seguir un trazado curvilíneo.

Normalidad

Histograma. Ofrece un histograma de los residuos tipificados con una curva normal superpuesta. La curva se construye tomando una media de 0 y una desviación típica de 1, es decir, la misma media y la misma desviación típica que los residuos tipificados.

En el histograma de la figura adjunta podemos observar, en primer lugar, que la parte central de la distribución acumula muchos más casos de los que existen en una curva normal. En segundo lugar, la distribución es algo asimétrica: en la cola positiva de la distribución existen valores más extremos que en la negativa (esto ocurre cuando uno o varios errores muy grandes, correspondientes por lo general a valores atípicos, son contrarrestados con muchos residuos pequeños de signo opuesto). La distribución de los residuos, por tanto, no parece seguir el modelo de probabilidad normal, de modo que los resultados del análisis deben ser interpretados con cautela.

Figura 18.9. *Histograma de los residuos tipificados.*

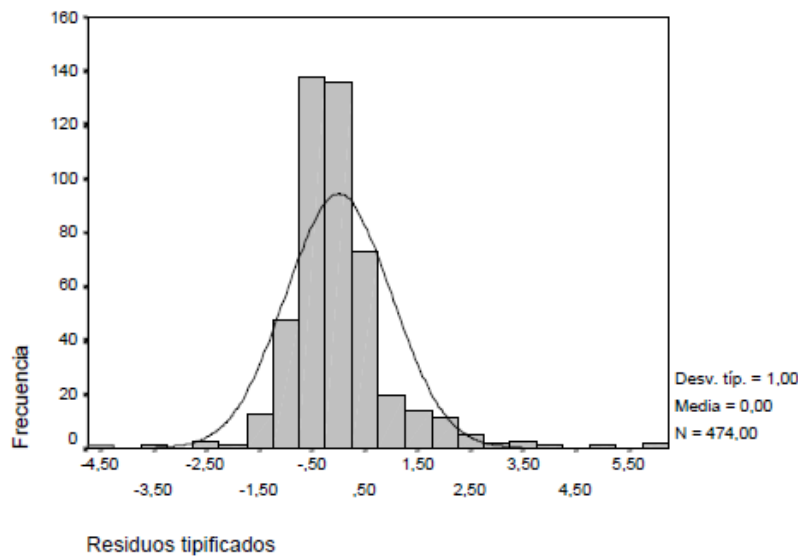


Gráfico 20. Histograma de residuos tipificados

Gráfico de probabilidad normal. Permite obtener un diagrama de probabilidad normal. En el eje de abscisas está representada la probabilidad acumulada que corresponde a cada residuo tipificado. El de ordenadas representa la probabilidad acumulada teórica que corresponde a cada puntuación típica en una curva normal con media 0 y desviación típica 1. Cuando los residuos se distribuyen normalmente, la nube de puntos se encuentra alineada sobre la diagonal del gráfico.

En nuestro caso de ejemplo, los puntos no se encuentran alineados sobre la diagonal del gráfico, lo cual nos está avisando de nuevo del posible incumplimiento del supuesto de normalidad.

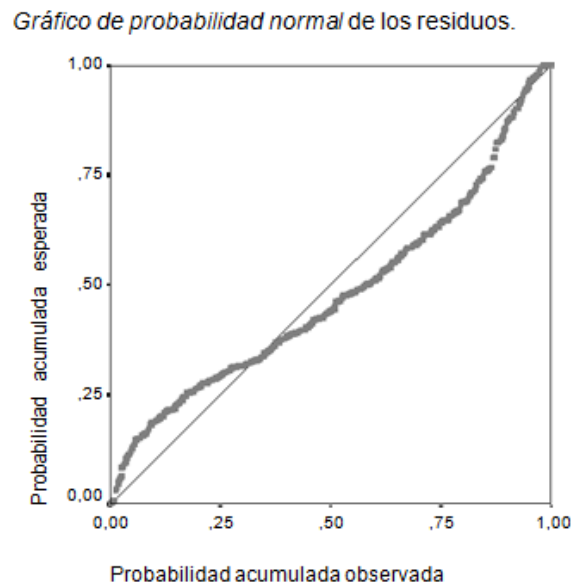


Gráfico 21. Probabilidad normal de los residuos

Linealidad

Los diagramas de regresión parcial permiten formarse una idea rápida sobre la forma que adopta una relación. En el contexto del análisis de regresión, permiten examinar la relación existente entre la variable dependiente y cada una de las variables independientes por separado, tras eliminar de ellas el efecto del resto de las variables independientes incluidas en el análisis.

Estos diagramas son similares a los de dispersión ya estudiados, pero no están basados en las puntuaciones originales de las dos variables representadas, sino en los residuos obtenidos al efectuar un análisis de regresión con el resto de las variables independientes. Por ejemplo, en el diagrama de regresión parcial de salario actual y salario inicial están representados los residuos que resultan de efectuar un análisis de regresión sobre salario actual incluyendo todas las variables independientes excepto salario inicial, y los residuos que resultan de efectuar un análisis de regresión sobre salario inicial incluyendo el resto de variables independientes. La utilidad de estos diagramas está en que, puesto que se controla el efecto del resto de las variables, muestran la relación neta entre las variables representadas. Además, las rectas que mejor se ajustan a la nube de puntos de estos diagramas son las definidas por los correspondientes coeficientes de regresión (es justamente en esa nube de puntos en la que se basan los coeficientes de regresión parcial).

Esta opción genera tantos gráficos parciales como variables independientes se hayan incluido en el análisis. En nuestro ejemplo, por tanto, aparecerán tres de estos gráficos. La siguiente figura muestra uno de ellos. Podemos observar que la relación entre salini (una de las variables independientes) y salario (la variable dependiente), tras eliminar el efecto del resto de variables independientes, es claramente lineal y positiva.

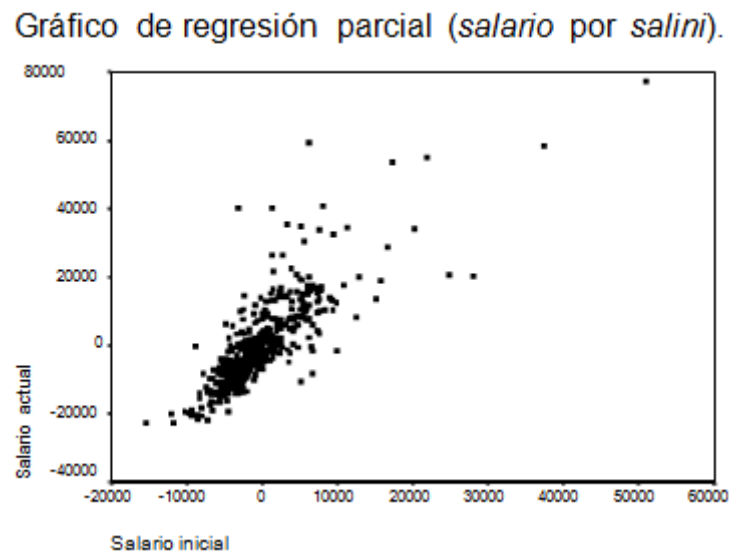


Gráfico 22. Nube de puntos de la regresión parcial

Así pues, los diagramas de regresión parcial permiten formarse una rápida idea sobre el tamaño y el signo de los coeficientes de regresión parcial (los coeficientes de la ecuación de regresión). En estos diagramas, los valores extremos pueden resultar muy informativos.

Colinealidad

Existe colinealidad perfecta cuando una de las variables independientes se relaciona de forma perfectamente lineal con una o más del resto de variables independientes de la ecuación. Esto ocurre, por ejemplo, cuando se utilizan como variables independientes en la misma ecuación las puntuaciones de las subescalas de un test y la puntuación total en el test (que es la suma de las subescalas y, por tanto, una combinación lineal perfecta de las mismas). Hablamos de colinealidad parcial o, simplemente, colinealidad, cuando entre las variables independientes de una ecuación existen correlaciones altas. Se puede dar, por ejemplo, en una investigación de mercados al tomar registros de muchos atributos de un mismo producto; o al utilizar muchos indicadores económicos para construir una ecuación de regresión. En términos generales, cuantas más variables hay

en una ecuación, más fácil es que exista colinealidad (aunque, en principio, bastan dos variables).

La colinealidad es un problema porque, en el caso de colinealidad perfecta, no es posible estimar los coeficientes de la ecuación de regresión; y en el caso de colinealidad parcial, aumenta el tamaño de los residuos tipificados y esto produce coeficientes de regresión muy inestables: pequeños cambios en los datos (añadir o quitar un caso, por ejemplo) produce cambios muy grandes en los coeficientes de regresión. Esta es una de las razones por las que podemos encontrarnos con coeficientes con signo cambiado: correlaciones positivas pueden transformarse en coeficientes de regresión negativos (incluso significativamente negativos). Curiosa-mente, la medida de ajuste R^2 no se altera por la presencia de colinealidad; pero los efectos atribuidos a las variables independientes pueden ser engañosos.

Al evaluar la existencia o no de colinealidad, la dificultad estriba precisamente en determinar cuál es el grado máximo de relación permisible entre las variables independientes. No existe un consenso generalizado sobre esta cuestión, pero puede servirnos de guía la presencia de ciertos indicios que podemos encontrar en los resultados de un análisis de regresión (estos indicios, no obstante, pueden tener su origen en otras causas):

- El estadístico F que evalúa el ajuste general de la ecuación de regresión es significativo, pero no lo es ninguno de los coeficientes de regresión parcial.
- Los coeficientes de regresión parcial estandarizados (los coeficientes beta) están inflados tanto en positivo como en negativo (adoptan, al mismo tiempo, valores mayores que 1 y menores que -1).
- Existen valores de tolerancia pequeños (próximos a 0,01). La tolerancia de una variable independiente es la proporción de varianza de esa variable que no está asociada (que no depende) del resto de variables independientes incluidas en la ecuación. Una variable con una tolerancia de, por ejemplo, 0,01 es una variable que comparte el 99 % de su varianza con el resto de variables independientes, lo cual significa que se trata de una variable redundante casi por completo.
- Los coeficientes de correlación estimados son muy grandes (por encima de 0,90 en valor absoluto).

Las afirmaciones del tipo “inflados”, “próximos a cero”, “muy grandes” se deben al hecho de que no existe un criterio estadístico formal en el que basar nuestras decisiones. Sólo existen recomendaciones basadas en trabajos de simulación.

Al margen de estos indicios, el SPSS ofrece la posibilidad de obtener algunos estadísticos que pueden ayudar a diagnosticar la presencia de colinealidad. Se trata de estadísticos orientativos que, aunque pueden ayudarnos a determinar si existe mayor o menor grado de colinealidad, no permiten tomar una decisión clara sobre la presencia o no de colinealidad.

Esta opción permite obtener los estadísticos de colinealidad que recogen las tablas siguientes. La primera es la tabla de coeficientes de regresión parcial ya vista, pero ahora contiene información adicional sobre los niveles de tolerancia y sus inversos (FIV).

El nivel de tolerancia de una variable se obtiene restando a 1 el coeficiente de determinación (R^2) que resulta al regresar esa variable sobre el resto de variables independientes. Valores de tolerancia muy pequeños indican que esa variable puede ser explicada por una combinación lineal del resto de variables, lo cual significa que existe colinealidad.

Los factores de inflación de la varianza (FIV) son los inversos de los niveles de tolerancia. Reciben ese nombre porque son utilizados en el cálculo de las varianzas de los coeficientes de regresión. Cuanto mayor es el FIV de una variable, mayor es la varianza del correspondiente coeficiente de regresión. De ahí que uno de los problemas de la presencia de colinealidad (tolerancias pequeñas, FIV grandes) sea la inestabilidad de las estimaciones de los coeficientes de regresión.

Coeficientes de regresión parcial y niveles de tolerancia.

	Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.	Estadísticos de colinealidad	
	B	Error típ.	Beta			Tolerancia	FIV
(Constante)	-3661,5	1935,49		-1,892	,059		
Salario inicial	1,749	,060	,806	29,198	,000	,554	1,804
Experiencia previa	-16,730	3,605	-,102	-4,641	,000	,866	1,154
Nivel educativo	735,956	168,689	,124	4,363	,000	,520	1,923

Tabla 47. Coeficientes parciales y niveles de tolerancia

La siguiente tabla muestra la solución resultante de aplicar un análisis de componentes principales a la matriz estandarizada no centrada de productos cruzados de las variables independientes.

Diagnósticos de colinealidad.

Dimensión	Autovalor	Índice de condición	Proporciones de la varianza			
			(Constante)	Salario inicial	Experiencia previa	Nivel educativo
1	3,401	1,000	,00	,01	,02	,00
2	,489	2,638	,00	,01	,79	,00
3	9,663 E-02	5,933	,11	,62	,01	,01
4	1,347 E-02	15,892	,88	,36	,18	,98

Tabla 48. Colinealidad

Los autovalores informan sobre cuántas dimensiones o factores diferentes subyacen en el conjunto de variables independientes utilizadas. La presencia de varios autovalores próximos a cero indica que las variables independientes están muy relacionadas entre sí (colinealidad).

Los índices de condición son la raíz cuadrada del cociente entre el autovalor más grande y cada uno del resto de los autovalores. En condiciones de no-colinealidad, estos índices no deben superar el valor 15. Índices mayores que 15 indican un posible problema. Índices mayores que 30 delatan un serio problema de colinealidad.

Las proporciones de varianza recogen la proporción de varianza de cada coeficiente de regresión parcial que está explicada por cada dimensión o factor. En condiciones de no-colinealidad, cada dimensión suele explicar gran cantidad de varianza de un sólo coeficiente (excepto en lo que se refiere al coeficiente B0 o constante, que siempre aparece asociado a uno de los otros coeficientes; en el ejemplo, el término constante aparece asociado al coeficiente de Nivel educativo). La colinealidad es un problema cuando una dimensión o factor con un índice de condición alto, contribuye a explicar gran cantidad de la varianza de los coeficientes de dos o más variables.

Si se detecta la presencia de colinealidad en un conjunto de datos, hay que aplicar algún tipo de remedio. A continuación proponemos algunos: aumentar el tamaño de la muestra (esta solución puede resultar útil si existen pocos casos en relación al número

de variables); crear indicadores múltiples combinando variables (por ejemplo, promediando variables; o efectuando un análisis de componentes principales para reducir las variables a un conjunto de componentes independientes, y aplicando después el análisis de regresión sobre esos componentes); excluir variables redundantes (es decir, excluir variables que correlacionan muy alto con otras, quedándonos con las que consideremos más importantes); utilizar una técnica de estimación sesgada, tal como la regresión "ridge".

4.4 ANÁLISIS DE REGRESIÓN POR PASOS (REGRESIÓN STEPWISE)

En los apartados previos hemos utilizado un método de regresión en el que el control sobre las variables utilizadas para construir el modelo de regresión recae sobre el propio analista. Es el analista quien decide qué variables independientes desea incluir en la ecuación de regresión seleccionándolas en la lista Independientes.

Sin embargo, no es infrecuente encontrarse con situaciones en las que, existiendo un elevado número de posibles variables independientes, no existe una teoría o un trabajo previo que oriente al analista en la elección de las variables relevantes. Este tipo de situaciones pueden afrontarse utilizando procedimientos diseñados para seleccionar, entre una gran cantidad de variables, sólo un conjunto reducido de las mismas: aquellas que permiten obtener el mejor ajuste posible.

Con estos procedimientos de selección, el control sobre las variables que han de formar parte de la ecuación de regresión pasa de las manos del investigador a una regla de decisión basada en criterios estadísticos.

4.4.1 CRITERIOS DE SELECCIÓN DE VARIABLES

Existen diferentes criterios estadísticos para seleccionar variables en un modelo de regresión. Algunos de estos criterios son: el valor del coeficiente de correlación múltiple R^2 (corregido o sin corregir), el valor del coeficiente de correlación parcial entre cada

variable independiente y la dependiente, el grado de reducción que se obtiene en el error típico de los residuos al incorporar una variable, etc. De una u otra forma, todos ellos coinciden en intentar maximizar el ajuste del modelo de regresión utilizando el mínimo número posible de variables.

Los métodos por pasos que incluye el SPSS basan la selección de variables en dos criterios estadísticos:

1. Criterio de significación. De acuerdo con este criterio, sólo se incorporan al modelo de regresión aquellas variables que contribuyen de forma significativa al ajuste del modelo. La contribución individual de una variable al ajuste del modelo se establece contrastando, a partir del coeficiente de correlación parcial, la hipótesis de independencia entre esa variable y la variable dependiente. Para decidir si se mantiene o rechaza esa hipótesis de independencia, el SPSS incluye dos criterios de significación:

- Probabilidad de F. Una variable pasa a formar parte del modelo de regresión si el nivel crítico asociado a su coeficiente de correlación parcial al contrastar la hipótesis de independencia es menor que 0,05 (probabilidad de entrada). Y queda fuera del modelo de regresión si ese nivel crítico es mayor que 0,10 (probabilidad de salida).
- Valor de F. Una variable pasa a formar parte del modelo de regresión si el valor del estadístico F utilizado para contrastar la hipótesis de independencia es mayor que 3,84 (valor de entrada). Y queda fuera del modelo si el valor del estadístico F es menor que 2,71 (valor de salida).

2. Criterio de tolerancia. Superado el criterio de significación, una variable sólo pasa a formar parte del modelo si su nivel de tolerancia es mayor que el nivel establecido por defecto (este nivel es 0,0001, pero puede cambiarse mediante sintaxis) y si, además, aun correspondiéndole un coeficiente de correlación parcial significativamente distinto de cero, su incorporación al modelo hace que alguna de las variables previamente seleccionadas pase a tener un nivel de tolerancia por debajo del nivel establecido por defecto. El concepto de tolerancia se ha descrito ya en el apartado sobre colinealidad.

Una forma muy intuitiva de comprender y valorar el efecto resultante de aplicar estos criterios de selección consiste en observar el cambio que se va produciendo en el coeficiente de determinación R^2 a medida que se van incorporando (o eliminando) variables al modelo. Podemos definir este cambio como $R^2_{\text{cambio}} = R^2 - R^2_i$, donde R^2_i se

refiere al coeficiente de determinación obtenido con todas las variables independientes excepto la i -ésima. Un cambio grande en R^2 indica que esa variable contribuye de forma importante a explicar lo que ocurre con la variable dependiente.

4.4.2 MÉTODOS DE SELECCIÓN DE VARIABLES

Existen diferentes métodos para seleccionar las variables independientes que debe incluir un modelo de regresión, pero los que mayor aceptación han recibido son los métodos de selección por pasos (stepwise). Con estos métodos, se selecciona en primer lugar la mejor variable (siempre de acuerdo con algún criterio estadístico); a continuación, la mejor de las restantes; y así sucesivamente hasta que ya no quedan variables que cumplan los criterios de selección.

El procedimiento de regresión lineal del SPSS incluye varios de estos métodos de selección de variables. Dos de estos métodos permiten incluir o excluir, en un sólo paso, todas las variables independientes seleccionadas (no son métodos de selección por pasos):

- **Introducir.** Este método construye la ecuación de regresión utilizando todas las variables seleccionadas en la lista Independientes. Es el método utilizado por defecto.
- **Eliminar.** Elimina en un sólo paso todas las variables de la lista Independientes y ofrece los coeficientes de regresión que corresponderían a cada variable en el caso de que pasaran a formar parte de la ecuación de regresión.

El resto de métodos de selección de variables son métodos por pasos, es decir, métodos que van incorporando o eliminando variables paso a paso dependiendo de que éstas cumplan o no los criterios de selección:

- **Hacia adelante (forward selection).** Las variables se incorporan al modelo de regresión una a una. En el primer paso se selecciona la variable independiente que, además de superar los criterios de entrada, más alto correlaciona (positiva o negativamente) con la dependiente. En los siguientes pasos se utiliza como criterio de selección el coeficiente de correlación parcial: van siendo seleccionadas una a una las variables que, además de superar los criterios de entrada, poseen el coeficiente de

correlación parcial más alto en valor absoluto (la relación se parcializa controlando el efecto de las variables independientes previamente seleccionadas).

La selección de variables se detiene cuando no quedan variables que superen el criterio de entrada. (Utilizar como criterio de entrada el tamaño, en valor absoluto, del coeficiente de correlación parcial, es equivalente a seleccionar la variable con menor probabilidad de F o mayor valor de F).

- **Hacia atrás (backward elimination).** Comienza incluyendo en el modelo todas las variables seleccionadas en la lista Independientes y luego procede a eliminarlas una a una. La primera variable eliminada es aquella que, además de cumplir los criterios de salida, posee el coeficiente de regresión más bajo en valor absoluto. En cada paso sucesivo se van eliminando las variables con coeficientes de regresión no significativos, siempre en orden inverso al tamaño de su nivel crítico.

La eliminación de variables se detiene cuando no quedan variables en el modelo que cumplan los criterios de salida.

- **Pasos sucesivos.** Este método es una especie de mezcla de los métodos hacia adelante y hacia atrás. Comienza, al igual que el método hacia adelante, seleccionando, en el primer paso, la variable independiente que, además de superar los criterios de entrada, más alto correlaciona (en valor absoluto) con la variable dependiente. A continuación, selecciona la variable independiente que, además de superar los criterios de entrada, posee el coeficiente de correlación parcial más alto (en valor absoluto). Cada vez que se incorpora una nueva variable al modelo, las variables previamente seleccionadas son, al igual que en el método hacia atrás, evaluadas nuevamente para determinar si siguen cumpliendo o no los criterios de salida. Si alguna variable seleccionada cumple los criterios de salida, es eliminada del modelo.

El proceso se detiene cuando no quedan variables que superen los criterios de entrada y las variables seleccionadas no cumplen los criterios de salida.

4.4.3 REGRESIÓN POR PASOS

Para ilustrar el funcionamiento del análisis de regresión por pasos, vamos a presentar el ejemplo seguido anteriormente con el método de pasos sucesivos. Utilizaremos el salario actual (salario) como variable dependiente y, como variables independientes, la fecha de nacimiento (fechnac), el nivel educativo (educ), el salario inicial (salini), la experiencia previa (expprev), y la clasificación étnica (minoría). El objetivo del análisis es encontrar un modelo de regresión que explique, con el mínimo número posible de variables independientes, la mayor cantidad posible de la varianza de la variable salario.

Variables introducidas/eliminadas.

Modelo	Variables introducidas	Variables eliminadas	Método
1	Salario inicial		Por pasos (criterio: Probabilidad de F para entrar $\leq$,050, Probabilidad de F para salir $\geq$,100).
2	Experiencia previa		Por pasos (criterio: Probabilidad de F para entrar $\leq$,050, Probabilidad de F para salir $\geq$,100).
3	Nivel educativo		Por pasos (criterio: Probabilidad de F para entrar $\leq$,050, Probabilidad de F para salir $\geq$,100).

Tabla 49. Variables seleccionadas

La tabla de arriba ofrece un resumen del modelo final al que se ha llegado. En la columna Modelo se indica en número de pasos dados para construir el modelo de regresión: tres pasos. En el primer paso se ha seleccionado la variable salario inicial, en el segundo, experiencia previa y, en el tercero, nivel educativo. También se indica si en alguno de los pasos se ha eliminado alguna variable previamente seleccionada; en este ejemplo no se han eliminado variables. Por último, se informa sobre el método de selección utilizado (por pasos) y sobre los criterios de entrada y salida: una variable es incorporada al modelo si su coeficiente de regresión parcial es significativamente distinto de cero al 5 % y, una vez seleccionada, sólo es eliminada del modelo si con la incorporación de otra u otras variables en un paso posterior su coeficiente de regresión parcial deja de ser significativamente distinto de cero al 10 %.

Resumen del modelo.

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Estadísticos del cambio				
					Cambio en R cuadrado	Cambio en F	gl1	gl2	Sig. del cambio en F
1	,880	,775	,774	8,119.79	,775	1620,83	1	471	,000
2	,891	,794	,793	7,778.94	,019	43,180	1	470	,000
3	,896	,802	,801	7,631.84	,008	19,293	1	469	,000

Tabla 50. Resumen del modelo

Esta tabla recoge el valor de R^2 en cada paso, el cambio experimentado por R^2 en cada paso, y el estadístico F y su significación. El estadístico F permite contrastar la hipótesis de que el cambio en R^2 vale cero en la población. Al seleccionar la primera variable (Modelo 1), el valor de R^2 es 0,775. Lógicamente, en el primer paso, $R^2_{\text{cambio}} = R^2$. Al contrastar la hipótesis de que el valor poblacional de R^2_{cambio} es cero se obtiene un estadístico F de 1.620,83 que, con 1 y 471 grados de libertad, tiene una probabilidad asociada de 0,000. Puesto que este valor es menor que 0,05, podemos afirmar que la proporción de varianza explicada por la variable salario inicial (la variable seleccionada en el primer paso) es significativamente distinta de cero. En el segundo paso (Modelo 2), el valor de R^2 aumenta hasta 0,794. Esto supone un cambio de 0,019 (aproximadamente un 2 por ciento). La tabla muestra el valor del estadístico F (43,180) obtenido al contrastar la hipótesis de que el valor poblacional de R^2_{cambio} es cero, y su significación (0,000). Aunque se trata de un incremento muy pequeño (un 2 por ciento), el valor del nivel crítico nos permite afirmar que la variable experiencia previa (la variable incorporada al modelo en el segundo paso) contribuye significativamente a explicar lo que ocurre con la variable dependiente.

En el tercer y último paso (Modelo 3), R^2 toma un valor de 0,802, lo cual supone un incremento de 0,008 (aproximadamente un 1 por ciento). De nuevo se trata de un incremento muy pequeño, pero al evaluar su significación se obtiene un estadístico F de 19,293 y un nivel crítico de 0,000, lo cual nos está indicando que la variable nivel educativo (variable incorporada en el tercer paso), también contribuye de forma significativa a explicar el comportamiento de la variable dependiente. Las tres variables seleccionadas en el modelo final consiguen explicar un 80 por ciento ($R^2 = 0,802$) de la variabilidad observada en el salario actual.

Resumen del ANOVA.

Modelo		Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	106862706669,340	1	106862706669,340	1620,826	,000
	Residual	31053506813,535	471	65931012,343		
	Total	137916213482,875	472			
2	Regresión	109475617434,356	2	54737808717,178	904,579	,000
	Residual	28440596048,519	470	60511906,486		
	Total	137916213482,875	472			
3	Regresión	110599332927,334	3	36866444309,111	632,955	,000
	Residual	27316880555,541	469	58244947,880		
	Total	137916213482,875	472			

Tabla 51. Resumen del ANOVA

La tabla resumen del ANOVA contiene el valor del estadístico F obtenido al contrastar la hipótesis de que el valor poblacional de R^2 en cada paso es cero. Ahora no se evalúa el cambio que se va produciendo en el valor de R^2 de un paso a otro, sino el valor de R^2 en cada paso. Lógicamente, si R^2 es significativamente distinta de cero en el primer paso, también lo será en los pasos sucesivos.

Coefficientes de regresión parcial.

Modelo		Coeficientes no estandarizados		Coeficientes estandarizados	t	Sig.
		B	Error típ.	Beta		
1	(Constante)	1929,517	889,168		2,170	,031
	Salario inicial	1,910	,047	,880	40,259	,000
2	(Constante)	3856,955	900,928		4,281	,000
	Salario inicial	1,924	,045	,887	42,279	,000
	Experiencia previa	-22,500	3,424	-,138	-6,571	,000
3	(Constante)	-3708,904	1936,045		-1,916	,056
	Salario inicial	1,748	,060	,806	29,192	,000
	Experiencia previa	-16,752	3,605	-,103	-4,647	,000
	Nivel educativo	741,307	168,772	,125	4,392	,000

Tabla 52. Coeficientes parciales

La tabla anterior contiene los coeficientes de regresión parcial de las variables incluidas en el modelo de regresión; es decir, la información necesaria para construir la ecuación de regresión encada paso (incluyendo el término constante). Las primeras columnas recogen el valor de los coeficientes de regresión parcial (B) y su error típico. A continuación aparecen los coeficientes de regresión parcial estandarizados (Beta), los cuales proporcionan una idea acerca de la importancia relativa de cada variable dentro de la ecuación. Las dos últimas columnas muestran el estadístico t y el nivel crítico

(Sig.) obtenidos al contrastar las hipótesis de que los coeficientes de regresión parcial valen cero en la población. Un nivel crítico por debajo de 0,05 indica que la variable contribuye significativamente a mejorar la calidad del modelo de regresión.

Utilizar el estadístico t para contrastar la hipótesis de que un coeficiente de regresión parcial vale cero es exactamente lo mismo que utilizar el estadístico F para contrastar la hipótesis de que el valor poblacional del cambio observado en R^2 vale cero. De hecho, elevando al cuadrado los valores del estadístico t de la tabla anterior obtenemos los valores del estadístico F de la tabla siguiente. De las dos formas se está intentando evaluar la contribución individual de una variable a la proporción de varianza explicada por el conjunto de variables independientes.

Variables excluidas.

Modelo	Beta dentro	t	Sig.	Correlación parcial	Estadísticos de colinealidad
					Tolerancia
1	Nivel educativo	,173	6,385	,000	,283
	Experiencia previa	-,138	-6,571	,000	-,290
	Clasificación étnica	-,040	-1,809	,071	-,083
	Fecha de nacimiento	,136	6,471	,000	,286
2	Nivel educativo	,125	4,392	,000	,199
	Clasificación étnica	-,019	-,885	,377	-,041
	Fecha de nacimiento	,071	2,020	,044	,093
3	Clasificación étnica	-,020	-,965	,335	-,045
	Fecha de nacimiento	,054	1,556	,120	,072

Tabla 53. Variables excluidas

Por último, esta tabla muestra los coeficientes de regresión parcial de las variables no seleccionadas para formar parte de la ecuación de regresión en cada paso. La información que contiene esta tabla permite conocer en detalle por qué unas variables han sido seleccionadas y otras no. En el primer paso se ha seleccionado la variable salario inicial porque es la que más alto correlaciona, en valor absoluto, con la variable dependiente. En ese primer paso, todavía están fuera del modelo el resto de variables independientes. La columna Beta dentro contiene el valor que tomaría el coeficiente de regresión estandarizado de una variable en el caso de que fuera seleccionada en el siguiente paso. Las columnas t y Sig. nos informan sobre si ese valor que adoptaría el coeficiente de regresión de una variable en el caso de ser incorporada al modelo sería o no significativamente distinto de cero.

Vemos que, en el primer paso, hay tres variables todavía no seleccionadas (nivel educativo, experiencia previa y fecha de nacimiento) cuyos coeficientes de regresión poseen niveles críticos por debajo de 0,05 (criterio de entrada). Entre ellas, la que posee un coeficiente de correlación parcial mayor en valor absoluto (experiencia previa = - 0,290) y, además, un nivel de tolerancia por encima de 0,001 (tolerancia mínima establecida por defecto), es la variable que ha sido seleccionada en el segundo paso. En el segundo paso todavía quedan fuera de la ecuación dos variables cuyos coeficientes de regresión serían significativos en caso de ser seleccionadas: nivel educativo y fecha de nacimiento. En el tercer paso ha sido seleccionada la variable nivel educativo porque, teniendo un nivel de tolerancia por encima de 0,001, es la que posee el coeficiente de correlación parcial más alto. Después del tercer paso ya sólo quedan dos variables fuera de la ecuación: fecha de nacimiento y clasificación étnica. Puesto que ninguna de las dos supera el criterio de entrada ($\text{Sig.} < 0,05$), el proceso se detiene y ambas variables quedan fuera del modelo.

4.4.4 QUÉ VARIABLES DEBE INCLUIR LA ECUACIÓN DE REGRESIÓN

El método de selección por pasos nos ha llevado a construir una ecuación de regresión con tres variables. Las tres variables seleccionadas poseen coeficientes de regresión parcial significativos. Sin embargo, la primera variable explica el 78 % de la varianza de la variable dependiente, la segunda el 2 %, y la tercera el 2 %. Si en lugar del método pasos sucesivos hubiéramos utilizado el método introducir, habríamos obtenido los resultados que muestran a continuación.

Resumen del modelo.

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,902	,813	,811	7,437.80

Coefficientes de regresión parcial.

	Coefficientes no estandarizados		Coefficientes estandarizados	t	Sig.
	B	Error típ.	Beta		
(Constante)	-38768,134	18686,944		-2,075	,039
Código de empleado	-12,426	2,526	-,099	-4,920	,000
Fecha de nacimiento	3,275E-06	,000	,071	2,091	,037
Nivel educativo	630,255	166,453	,106	3,786	,000
Salario inicial	1,760	,059	,811	29,826	,000
Experiencia previa	-7,703	5,596	-,047	-1,376	,169
Clasificación étnica	-994,637	846,844	-,024	-1,175	,241

Tabla 54. Resumen del modelo. Tabla 55. Coeficientes parciales.

Por un lado, la ganancia que obtenemos en R^2 utilizando las 5 variables en lugar de las tres seleccionadas con el método por pasos es extremadamente pequeña ($0,813 - 0,802 = 0,011$). No parece que tenga mucho sentido añadir dos variables para obtener una mejora de once milésimas en la proporción de varianza explicada. Es cierto que R^2 nunca disminuye cuando se van incorporando nuevas variables al modelo de regresión, sino que aumenta o se queda como está. Sin embargo, esto no significa, necesariamente, que la ecuación con más variables se ajuste mejor a los datos poblacionales. Generalmente, conforme va aumentando la calidad del modelo, va disminuyendo el error típico de los residuos (error típico de la estimación).

Pero el incremento que se va produciendo en R^2 al ir añadiendo variables no se corresponde necesariamente con una disminución del error típico de los residuos. Con cada variable nueva que se incorpora al modelo, la suma de cuadrados de la regresión gana un grado de libertad y la suma de cuadrados de los residuos lo pierde. Por tanto, el error típico de los residuos puede aumentar cuando el descenso de la variación residual es demasiado pequeño para compensar la pérdida de un grado de libertad en la suma de cuadrados de los residuos. Estas consideraciones sugieren la conveniencia de utilizar modelos parsimoniosos, es decir, modelos con un número reducido de variables independientes.

Por otro lado, las variables que tienen pesos significativos en la ecuación de regresión obtenida con el método pasos sucesivos no son las mismas que las que tienen pesos significativos en la ecuación obtenida con el método introducir. Esta diferencia entre métodos de selección de variables debe ser tomada muy en cuenta. ¿Cuáles son las variables buenas? Atendiendo a criterios puramente estadísticos, la ecuación de regresión con las tres variables seleccionadas por el método de pasos sucesivos, es la mejor de las posibles con el mínimo número de variables. Pero en la práctica, la decisión sobre cuántas variables debe incluir la ecuación de regresión puede tomarse teniendo en cuenta, además de los criterios estadísticos, otro tipo de consideraciones. Si, por ejemplo, resulta muy costoso (tiempo, dinero, etc.) obtener las unidades de análisis, un modelo con sólo una variable podría resultar lo bastante apropiado. Si las consecuencias de los residuos de los pronósticos fueran muy graves, deberían incluirse en el modelo las tres variables del método pasos sucesivos o las cuatro con pesos significativos del método introducir. Así pues, para decidir con qué modelo de regresión nos quedamos, casi siempre es conveniente tomar en consideración criterios adicionales a los puramente estadísticos.

Por supuesto, los contrastes estadísticos sirven de apoyo para tomar decisiones. Pero, dado que la potencia de un contraste se incrementa conforme lo hace el tamaño de la muestra, debemos ser cautelosos con las conclusiones a las que llegamos. Esto significa que, con muestras grandes, efectos muy pequeños desde el punto de vista de su importancia teórica o práctica pueden resultar estadísticamente significativos. Por el contrario, con muestras pequeñas, para que un efecto resulte significativo, debe tratarse de un efecto importante (con muestras pequeñas, existe mayor grado de coincidencia entre la significación estadística y la importancia práctica). Por esta razón, en la determinación de la ecuación de regresión final, debe tenerse en cuenta, cuando se trabaja con muestras grandes, la conveniencia de considerar elementos de decisión adicionales a la significación estadística.

Puesto que la utilización de los métodos de selección por pasos está bastante generalizada, conviene también alertar sobre el peligro de alcanzar un resultado falsamente positivo (un error de tipo I). Es decir, si examinamos un número de variables lo bastante grande, tarde o temprano una o más pueden resultar significativas sólo por azar. Este riesgo es tanto mayor cuanto más variables se incluyen en el análisis. Para evitar este problema, si la muestra es lo bastante grande, puede dividirse en dos, aplicar

el análisis a una mitad y verificar en la otra mitad si se confirma el resultado obtenido. Si la muestra es pequeña, esta solución es inviable y, por tanto, el riesgo de cometer un error de tipo I permanece.

4.5 EXPLICACIÓN MATEMÁTICA DEL MODELO DE REGRESIÓN LINEAL MÚLTIPLE

El modelo de regresión lineal múltiple es idéntico al modelo de regresión lineal simple, con la única diferencia de que aparecen más variables explicativas:

Modelo de regresión simple:

$$y = b_0 + b_1 \cdot x + u$$

Modelo de regresión múltiple:

$$y = b_0 + b_1 \cdot x_1 + b_2 \cdot x_2 + b_3 \cdot x_3 + \dots + b_k \cdot x_k + u$$

Hipótesis

Para realizar un análisis de regresión lineal múltiple se hacen las siguientes consideraciones sobre los datos:

- a) Linealidad: los valores de la variable dependiente están generados por el siguiente modelo lineal:

$$Y = X * B + U$$

- b) Homocedasticidad: todas las perturbaciones tienen la misma varianza:

$$V(u_i) = \sigma^2$$

- c) Independencia: las perturbaciones aleatorias son independientes entre sí:

$$E(u_i \cdot u_j) = 0, \forall i \neq j$$

- d) Normalidad: la distribución de la perturbación aleatoria tiene distribución normal:

$$U \approx N(0, \sigma^2)$$

- e) Las variables explicativas X_k se obtienen sin errores de medida.

Si admitimos que los datos presentan estas hipótesis entonces el teorema de Gauss-Markov establece que el método de estimación de mínimos cuadrados va a producir estimadores óptimos, en el sentido que los parámetros estimados van a estar centrados y van a ser de mínima varianza.

Estimación de los parámetros por mínimos cuadrados

Vamos a calcular un hiperplano de regresión de forma que se minimice la varianza residual: $\text{Min} \sum (y_j - \hat{y}_j)^2$

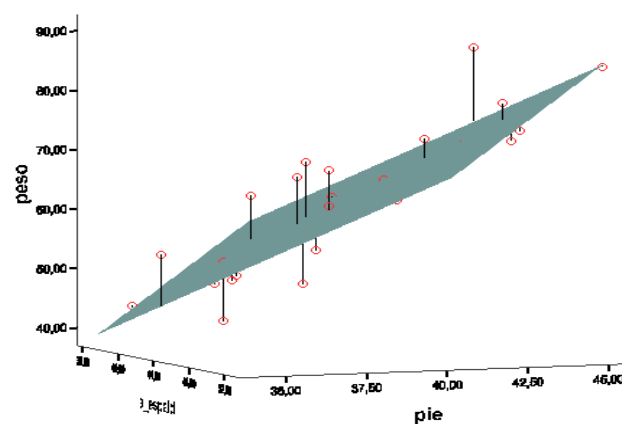


Gráfico 23. Hiperplano

Donde:

$$\hat{y}_j = b_0 + b_1 * x_{1,j} + b_2 * x_{2,j} + \dots b_k * x_{k,j}$$

Utilizando notación matricial:

$$u = \begin{bmatrix} u_1 \\ u_2 \\ . \\ . \\ u_n \end{bmatrix} = \begin{bmatrix} y_1 - \hat{y}_1 \\ y_2 - \hat{y}_2 \\ . \\ . \\ y_n - \hat{y}_n \end{bmatrix} = y - \hat{y}$$

Y teniendo en cuenta la definición de $\hat{y}$:

$$u = \begin{bmatrix} u_1 \\ u_2 \\ . \\ . \\ u_n \end{bmatrix} = \begin{bmatrix} y_1 - b_0 - b_1 * x_{1,1} - b_2 * x_{2,1} - b_3 * x_{3,1} - \dots - b_k * x_{k,1} \\ y_2 - b_0 - b_1 * x_{1,2} - b_2 * x_{2,2} - b_3 * x_{3,2} - \dots - b_k * x_{k,2} \\ . \\ . \\ y_n - b_0 - b_1 * x_{1,n} - b_2 * x_{2,n} - b_3 * x_{3,n} - \dots - b_k * x_{k,n} \end{bmatrix} = y - \hat{y}$$

Por lo tanto:

$$u = \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ y_n \end{bmatrix} - \begin{bmatrix} 1 & x_{1,1} & . & . & x_{k,1} \\ 1 & x_{1,2} & . & . & x_{k,2} \\ . & . & . & . & . \\ . & . & . & . & . \\ 1 & x_{1,n} & . & . & x_{k,n} \end{bmatrix} * \begin{bmatrix} b_0 \\ b_1 \\ . \\ . \\ b_k \end{bmatrix} = y - X * b$$

Por lo tanto la varianza residual se puede expresar de la siguiente forma:

$$n * \sigma^2 = u' * u = (y - X * b)' * (y - X * b)$$

Es decir:

$$\Phi(b) = \sum (y_j - \hat{y}_j)^2 = u' * u$$

Por tanto, la varianza residual es una función del vector de parámetros **b** y la condición para que tenga un mínimo será:

$$\frac{\partial \phi(b)}{\partial b} = 0$$

Antes de derivar vamos a simplificar la expresión de la varianza residual:

$$n * \sigma^2 = u' * u = (y - x * b)' * (y - x * b) = y' * y - y' * x * b - b' * x' * y + b' * x' * x * b$$

Por lo tanto:

$$\Phi(b) = \sum (y_j - \hat{y}_j)^2 = u' * u = y' * y - y' * x * b - b' * x' * y + b' * x' * x * b$$

$$\frac{\partial \phi(b)}{\partial b} = \frac{\partial (y - X * b)' * (y - X * b)}{\partial b} = -2 * X' * Y + 2 * X' * X * B$$

Igualando a cero y despejando:

$$X' * Y = X' * X * B$$

y si $X' * X$ es matriz no singular y por lo tanto tiene inversa, tenemos:

$$X' * Y = X' * X * B$$

Multiplicando por $(X' * X)^{-1}$

$$(X' * X)^{-1} X' * Y = (X' * X)^{-1} X' * X * B$$

$$(X' * X)^{-1} X' * Y = I * B$$

$$B = (X' * X)^{-1} * X' * Y$$

Ésta es la expresión del estimador de parámetros **B**.

Además

$$X' * Y = X' * X * B$$

$$X' * Y - X' * X * B = 0$$

$$X' * (Y - X * B) = 0$$

$$X' * U = 0$$

Es decir, los residuos obtenidos del modelo estimado por mínimos cuadrados no van a estar correlacionados con las variables explicativas.

Nota: Es importante observar que si las variables explicativas X están muy correlacionadas entre sí, la matriz $(X' * X)$ va a tener el determinante con valor cero o muy cercano a cero.

Si hay al menos una variable que puede ser expresada como combinación lineal del resto, el determinante de esta matriz es cero y dicha matriz será singular y por lo tanto no tendrá inversa.

Si no hay variables que sean combinación lineal de las demás, pero están fuertemente correlacionadas, el determinante no será cero pero tendrá un valor muy próximo a cero; este caso va a producir una inestabilidad en la solución del estimador, en general, se va a producir un aumento en su varianza.

En estos casos se impone la utilización de un método de selección de variables explicativas.

A los problemas provocados por la fuerte correlación entre las variables explicativas se les llama multicolinealidad.

Varianza residual

Vamos a descomponer la variabilidad de la variable dependiente Y en dos componentes o fuentes de variabilidad: una componente va a representar la variabilidad explicada por

el modelo de regresión y la otra componente va a representar la variabilidad no explicada por el modelo y, por tanto, atribuida a factores aleatorios.

Consideramos la variabilidad de la variable dependiente como:

$$n \cdot \sigma^2 = \sum (y_i - \bar{Y})^2$$

Es decir, la variabilidad de Y es la suma cuadrática de los valores que toma la variable respecto a la media de la variable.

Sumando y restando el valor pronosticado por el modelo de regresión obtenemos la siguiente expresión:

$$\sum (y_i - \bar{Y})^2 = \sum (\hat{y}_i - \bar{Y})^2 + \sum (y_i - \hat{y}_i)^2$$

Es decir, que la suma de cuadrados de la variable Y respecto a su media se puede descomponer en términos de la varianza residual. De esta expresión se deduce que “la distancia de Y a su media se descompone como la distancia de Y a su estimación más la distancia de su estimación a la media”.

Teniendo en cuenta que el último término representa la varianza no explicada, tenemos:

$$VT = VE + VNE$$

Gráficamente es fácil ver la relación:

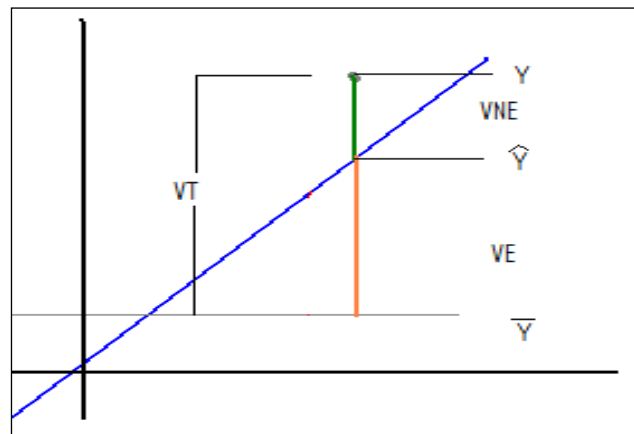


Gráfico 24. Distancias

Dividiendo la variabilidad total entre sus grados de libertad obtenemos la varianza de la variable dependiente Y :

$$S_Y^2 = \frac{VT}{n-1}$$

Dividiendo la variabilidad no explicada entre sus grados de libertad obtenemos la varianza residual de la variable dependiente Y :

$$S_R^2 = \frac{VNE}{n-(k+1)}$$

Tabla resumen

	Suma de cuadrados	Grados de libertad	
VT	$\sum (y - \bar{y})^2$	n-1	$S_Y^2 = \frac{VT}{n-1}$
VE	$\sum (\hat{y} - \bar{y})^2$	k-1	
VNE	$\sum (y - \hat{y})^2$	n-k-1	$S_R^2 = \frac{VNE}{n-k-1}$

Tabla 56. Resumen de VT, VE y VNE

Contraste de regresión

Si sacamos conclusiones de una muestra de un conjunto mucho más amplio de datos, a veces este conjunto será infinito, es obvio que distintas muestras van a dar distintos valores de los parámetros.

Un caso de especial interés es asignar una medida de probabilidad a la siguiente afirmación o hipótesis:

$$H_0 \equiv b_1 = b_2 = \dots = b_k = 0$$

La afirmación contraria sería:

$$H_1 \equiv \exists b_j \neq 0$$

Nota: La hipótesis nula es que todos los coeficientes menos b_0 son nulos y la hipótesis alternativa o complementaria es que existe al menos uno que es distinto de 0, puede haber varios que sean nulos, pero al menos existe uno distinto de cero.

Se denomina contraste de regresión al estudio de la posibilidad de que el modelo de regresión sea nulo, es decir, los valores de las variables explicativas X no van a influir en la variable dependiente Y.

Construcción del contraste

Si los residuos siguen una distribución normal y $b_1 = b_2 = \dots = b_k = 0$, tenemos que:

$$\frac{VT}{\sigma^2} \approx \chi_{n-1}^2$$

$$\frac{VE}{\sigma^2} \approx \chi_1^2$$

$$\frac{VNE}{\sigma^2} \approx \chi_{n-(k+1)}^2$$

Por tanto:

$$\frac{\frac{VE}{1}}{\frac{VNE}{n-(k+1)}} = \frac{VE}{S_R^2} \approx F_{1, n-(k+1)}$$

Es decir, el cociente entre la varianza explicada y la varianza no explicada será aproximadamente 1. Además, al seguir una distribución F, podemos asignar una medida de probabilidad (p-value) a la hipótesis de que la varianza explicada es igual a la varianza no explicada.

En caso contrario la varianza no explicada será muy inferior a la varianza explicada y, por lo tanto, este cociente tendrá un valor muy superior a 1.

Nota: En general si el p-value es menor de 0.05 se acepta que el modelo de regresión es significativo; en caso contrario no podemos hablar de regresión, pues el modelo sería nulo.

Coefficiente de determinación R^2

Vamos a construir un coeficiente (estadístico) que mida la bondad del ajuste del modelo. Si bien la varianza residual (S_R^2) nos indica cómo están de cerca las estimaciones respecto de los puntos, esta varianza está influida por la varianza de la variable dependiente, la cual, a su vez, está influida por su unidad de medida. Por lo tanto, una medida adecuada es la proporción de la varianza explicada (VE) entre la varianza total (VT); de este modo, definimos el coeficiente de determinación R^2 :

$$R^2 = \frac{VE}{VT} = \frac{VT - VNE}{VT} = 1 - \frac{VNE}{VT}$$

Por ser cociente de sumas de cuadrados, este coeficiente será siempre positivo.

Si todos los puntos están sobre la recta de regresión, la varianza no explicada será 0, y por lo tanto:

$$R^2 = \frac{VE}{VT} = 1 - \frac{0}{VT} = 1$$

Este coeficiente es muy importante pues determina qué porcentaje (en tantos por uno) de la varianza de la variable dependiente es explicado por el modelo de regresión.

En general, se pueden clasificar los valores de R^2 de la siguiente manera:

Menor de 0.3	0.3 a 0.4	0.4 a 0.5	0.5 a 0.85	Mayor de 0.85
Muy malo	Malo	Regular	Bueno	Muy Bueno

Tabla 57. Clasificación de los coeficientes de determinación

Además, a diferencia de la varianza residual, este coeficiente es adimensional; esto quiere decir que no está afectado por transformaciones lineales de las variables; por ello, si cambiamos las unidades de medida, el coeficiente de determinación permanecerá invariante.

Diagnosis y validación de un modelo de regresión lineal múltiple

1. Multicolinealidad

Si las variables explicativas se pueden expresar como una combinación lineal:

$$\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_k x_k + \alpha_0 = 0$$

Se dice que tenemos un problema de multicolinealidad.

En general, este problema va a afectar incrementando la varianza de los estimadores. Este problema se detecta fácilmente:

- Solicitando el determinante de la matriz de varianzas-covarianzas, que estará cercano a cero.
- Calculando el cociente entre el primer y último autovalor de la matriz de varianzas-covarianzas que será mayor de 50.
- Calculando para cada variable el coeficiente de determinación (R^2) de dicha variable con el resto.

La solución es eliminar del modelo aquellas variables explicativas que dependen unas de otras. En general, los métodos de selección de variables solucionan automáticamente este problema.

2. Análisis de residuos

Definimos como residuo del i-esimo caso a:

$$u_i = y_i - \hat{y}_i$$

Los residuos son variables aleatorias que siguen una distribución normal. Los residuos tienen unidades de medida y, por tanto no se puede determinar si es grande o pequeño a simple vista.

Para solventar este problema se define el residuo estandarizado como:

$$Zu_i = \frac{u_i}{\hat{S}_R} * \frac{1}{\sqrt{1-h_{ii}}}$$

Se considera que un residuo tiene un valor alto, y por lo tanto puede influir negativamente en el análisis, si su residuo estandarizado es mayor de 3 en valor absoluto.

$$|Zu_i| \geq 3$$

Para evitar la dependencia entre numerador y denominador de la expresión anterior, también se utilizan los residuos estudentizados.

$$SZu_i = \frac{u_i}{\hat{S}(i)_R} * \frac{1}{\sqrt{1-h_{ii}}}$$

Donde $\hat{S}(i)_R$ es la varianza residual calculada sin considerar el i-esimo caso.

El análisis descriptivo y el histograma de los residuos nos indicarán si existen casos que no se adapten bien al modelo lineal.

5. MODELOS ESTADÍSTICOS CON “IBM SPSS STATISTICS”

En los siguientes bloques se exponen los resultados de los diversos modelos estadísticos analizados con el software IBM SPSS STATISTICS 20.

Es muy importante resaltar, como se ha expuesto con anterioridad en la explicación del análisis de regresión lineal múltiple, que cada dato de cada variable introducido en SPSS es su logaritmo en base 10. Esto ocurre en todas las variables excepto en tres de ellas, que son:

- Variable 21: compañías low cost con más pasajeros.
- Variable 22: compañías low cost totales.
- Variable 23: hubs.

A estas tres variables no se les aplica logaritmos debido a que en muchos de sus casos (aeropuertos) el valor del dato introducido es 0, y el logaritmo de 0 no existe, por lo que no tenemos un número que introducir en el modelo. Como posible solución se pensó en dar a los valores 0 un valor muy pequeño próximo a 0, de este modo sus logaritmos tendrían un valor (negativo) y podrían introducirse en SPSS; pero los resultados que se obtenían eran confusos, y la bibliografía consultada lo desaconsejaba. Más adelante habrá que analizar con cautela los resultados de estas tres variables a las que no se les ha aplicado logaritmos (solo afecta al “modelo 6”).

No es objeto de este proyecto hacer una explicación detallada del software utilizado, por lo que a continuación y a modo de familiarizarse con este programa, se muestran algunas de las pantallas de más relevancia para un correcto análisis de regresión. Si fuese deseo del lector, hay diversos manuales en internet para aprender y profundizar a fondo en este software.

MODELO 6 TOTAL.sav [Conjunto_de_datos1] - IBM SPSS Statistics Editor de datos												
Archivo Edición Ver Datos Transformar Analizar Marketing directo Gráficos Utilidades Ventana Ayuda												
Visible: 24 de 24 variables												
	PASAJEROS	POBLACIÓN	RESIDENTES	RENTA	PARO	Í.COMERCIAL	Í.RESTAURACIÓN	Í.TURÍSTICO	Í.ECONÓMICO	Í.INDUSTRIAL	PERNOCTAC	S
1	5,92419500	6,0544	4,0532	4,3388	1,3023	3,4141	3,4393	3,0955	3,3641	3,3927	7	7
2	6,98402455	6,2684	5,3346	4,2326	1,4786	3,6196	3,6420	3,7054	3,5348	3,5061	7	7
3	5,84850564	5,8382	4,7922	4,2150	1,5593	3,0824	3,0941	3,2858	3,0441	3,0116	7	7
4	6,01678522	6,0249	4,2712	4,3137	1,3473	3,3847	3,4824	3,1059	3,3583	3,4365	7	7
5	4,46407204	5,8383	4,1144	4,1758	1,5148	3,1139	3,0414	2,5551	2,9832	2,9595	6	6
6	7,54675024	6,7356	5,1828	4,4248	1,3373	4,1197	4,0997	4,0108	4,1632	4,1086	7	7
7	6,57987205	6,0568	4,1582	4,4492	1,2312	3,3950	3,3808	2,8426	3,5392	3,6487	6	6
8	4,27657668	5,5627	4,1996	4,4256	1,2725	2,8876	2,9294	2,5694	3,0306	3,2605	6	6
9	5,14349257	4,0270	2,9799	4,2731	1,5186	1,3550	1,3150	1,9231	1,1915	,9892	5	5
10	6,62921269	5,0330	3,8468	4,2784	1,5227	2,4362	2,2859	2,9455	2,2503	1,9930	6	6
11	6,43725369	5,8716	4,6277	4,4279	1,3817	3,1443	3,0983	3,4923	3,1664	3,1847	7	7
12	6,98999630	5,9309	4,7445	4,2784	1,5227	3,3339	3,1837	3,8433	3,1481	2,8907	7	7
13	5,80501668	5,9644	4,4694	4,1960	1,5562	3,1772	3,2198	3,2463	3,0338	2,8831	7	7
14	6,75789526	5,1498	4,1810	4,3701	1,3600	2,4318	2,5678	3,1896	2,4077	2,2205	7	7
15	5,90926551	6,0960	4,3419	4,2166	1,6103	3,3473	3,3789	3,4433	3,3243	3,3524	7	7
16	4,38820781	4,3254	3,2789	4,2731	1,5186	1,6540	1,6140	2,2221	1,4904	1,2882	6	6
17	5,90822812	4,9238	3,8768	4,2731	1,5186	2,2519	2,2118	2,8200	2,0883	1,8860	6	6
18	6,72710178	5,1479	3,9614	4,2784	1,5227	2,5508	2,4005	3,0601	2,3650	2,1076	6	6
19	4,48981791	5,6855	4,0115	4,3079	1,3829	2,9939	3,0969	2,7067	3,0052	3,1086	6	6
20	4,02522392	5,4981	4,1632	4,4027	1,3021	2,8739	2,8645	2,4771	2,9325	3,0711	6	6
21	7,59917997	6,8041	5,4678	4,4611	1,3228	4,1685	4,1933	4,0349	4,2232	4,0086	7	7
22	7,11143680	6,2096	5,1582	4,2255	1,5587	3,5495	3,5743	3,8430	3,4765	3,3113	7	7

Figura 17. Vista de datos

MODELO 6 TOTAL.sav [Conjunto_de_datos1] - IBM SPSS Statistics Editor de datos												
Archivo Edición Ver Datos Transformar Analizar Marketing directo Gráficos Utilidades Ventana Ayuda												
Nombre Tipo Anchura Decimales Etiqueta Valores Perdidos Columnas Alineación Medida Rol												
1	PASAJEROS	Númérico	12	8		Ninguna	Ninguna	12	Derecha	Escala	Entrada	
2	POBLACIÓN	Númérico	12	4		Ninguna	Ninguna	12	Derecha	Escala	Entrada	
3	RESIDENTES	Númérico	12	4		Ninguna	Ninguna	12	Derecha	Escala	Entrada	
4	RENTA	Númérico	12	4		Ninguna	Ninguna	12	Derecha	Escala	Entrada	
5	PARO	Númérico	12	4		Ninguna	Ninguna	12	Derecha	Escala	Entrada	
6	Í.COMERCIAL	Númérico	12	4		Ninguna	Ninguna	12	Derecha	Escala	Entrada	
7	Í.RESTAUR...	Númérico	12	4		Ninguna	Ninguna	12	Derecha	Escala	Entrada	
8	Í.TURÍSTICO	Númérico	12	4		Ninguna	Ninguna	12	Derecha	Escala	Entrada	
9	Í.ECONÓMI...	Númérico	12	4		Ninguna	Ninguna	12	Derecha	Escala	Entrada	
10	Í.INDUSTRIAL	Númérico	12	4		Ninguna	Ninguna	12	Derecha	Escala	Entrada	
11	PERNOCTAC...	Númérico	12	4		Ninguna	Ninguna	12	Derecha	Escala	Entrada	
12	EXPORTAC...	Númérico	12	4		Ninguna	Ninguna	12	Derecha	Escala	Entrada	
13	IMPORTACI...	Númérico	12	4		Ninguna	Ninguna	12	Derecha	Escala	Entrada	
14	POTENCIAL	Númérico	12	4		Ninguna	Ninguna	12	Derecha	Escala	Entrada	
15	CUOTA	Númérico	12	4		Ninguna	Ninguna	12	Derecha	Escala	Entrada	
16	COMPAÑÍAS	Númérico	12	4		Ninguna	Ninguna	12	Derecha	Escala	Entrada	
17	DESTINOS	Númérico	12	4		Ninguna	Ninguna	12	Derecha	Escala	Entrada	
18	DESTINOS....	Númérico	12	4		Ninguna	Ninguna	12	Derecha	Escala	Entrada	
19	KMS.DESTI...	Númérico	12	4		Ninguna	Ninguna	12	Derecha	Escala	Entrada	
20	RENTA.PAÍ...	Númérico	12	4		Ninguna	Ninguna	12	Derecha	Escala	Entrada	
21	LOW.COST...	Númérico	12	0		Ninguna	Ninguna	12	Derecha	Nominal	Entrada	
22	LOW.COST...	Númérico	12	0		Ninguna	Ninguna	12	Derecha	Nominal	Entrada	
23	HUBS	Númérico	12	0		Ninguna	Ninguna	12	Derecha	Nominal	Entrada	
24	AEROPUE...	Cadena	17	0		Ninguna	Ninguna	17	Izquierda	Nominal	Entrada	

Figura 18. Vista de variables



Figura 19. Análisis de regresión lineal

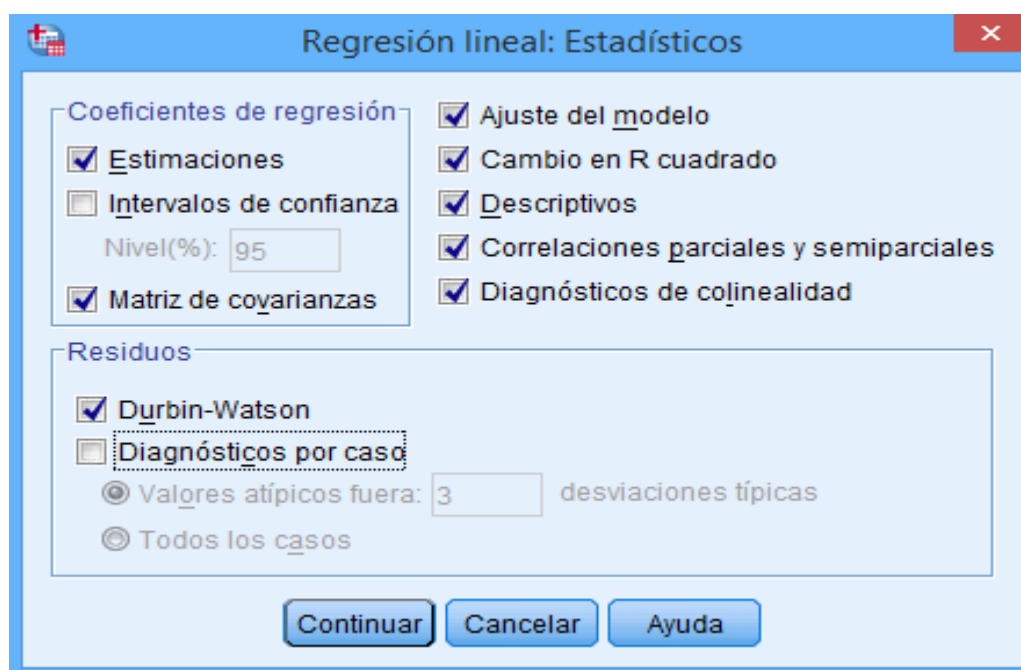


Figura 20. Estadísticos



Figura 21. Generador de gráficos

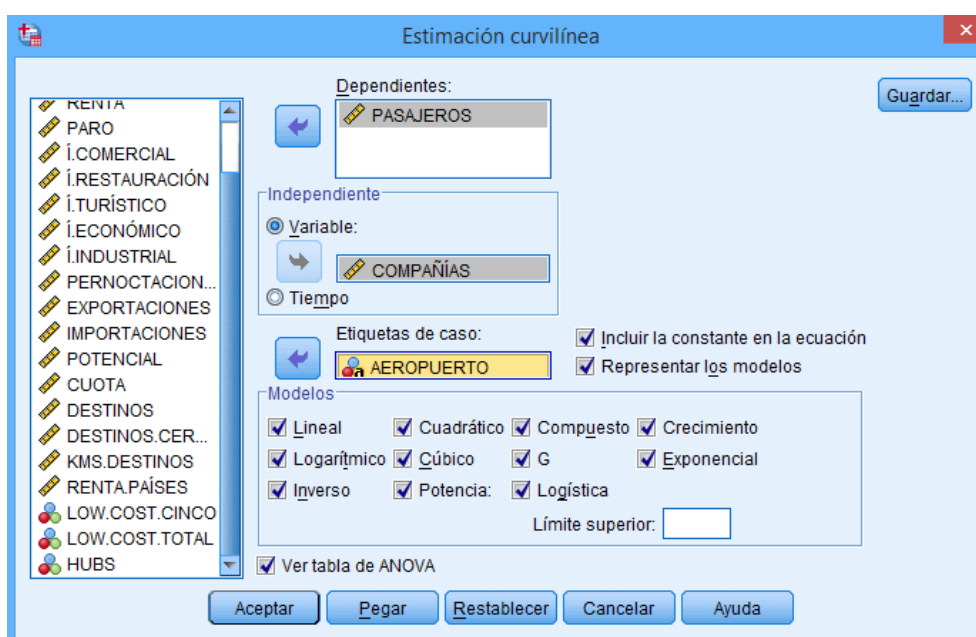


Figura 22. Estimador curvilíneo para modelos no lineales

Como ya se ha comentado, los anteriores cuadros solo muestran algunas de las sendas seguidas para llegar a los resultados que se detallan más adelante. No se incluyen aquí otros pasos como la identificación y edición de casos y variables, barras de herramientas, criterios y tolerancias según método, edición de gráficos y etiquetas de casos, otros estadísticos, exportar resultados, etc...

Los modelos resultantes analizados son:

- Modelos 1, 2, 3, 4 y 5 con las 15 variables socioeconómicas para los años 2013, 2012, 2011, 2010 y 2009 respectivamente. (En el modelo 5 del año 2009 se añade además la variable “población en los entornos aeroportuarios por distancias disuasorias”).
- Modelo 6 con 23 variables, donde se incluyen todas las variables socioeconómicas y aeroportuarias para el año 2013.
- Modelo de única variable para el año 2013 (modelo 6), donde se analiza cada variable independiente por separado respecto de la variable dependiente “pasajeros”.
- Más estadísticos para el año 2013 (modelo 6), donde se muestran algunos de los estadísticos analizados con SPSS, pero que no se han incluido en el resto de modelos por ser demasiado extensos.
- Modelos no lineales para el año 2013 (modelo 6), donde se comparan los coeficientes de determinación (R^2) de múltiples modelos no lineales respecto al modelo lineal.

5.1 RESULTADOS

MODELO 1

AÑO 2013

15 VARIABLES

CORRELACIONES

		PASAJEROS	POBLACIÓN	RESIDENTES
Correlación de Pearson	PASAJEROS	1,000	,459	,627
	POBLACIÓN	,459	1,000	,828
	RESIDENTES	,627	,828	1,000
	RENTA	,057	,181	,138
	PARO	,146	-,214	-,008
	Í.COMERCIAL	,454	,999	,835
	Í.RESTAUACIÓN	,445	,992	,830
	Í.TURÍSTICO	,785	,734	,907
	Í.ECONÓMICO	,400	,982	,807
	Í.INDUSTRIAL	,259	,946	,725
	PERNOCTACIONES	,672	,778	,922
	IMPORTACIONES	,321	,886	,667
	POTENCIAL	,294	,933	,715

		RENTA	PARO	Í.COMERCIAL
Correlación de Pearson	PASAJEROS	,057	,146	,454
	POBLACIÓN	,181	-,214	,999
	RESIDENTES	,138	-,008	,835
	RENTA	1,000	-,858	,198
	PARO	-,858	1,000	-,229
	Í.COMERCIAL	,198	-,229	1,000
	Í.RESTAUACIÓN	,228	-,274	,993
	Í.TURÍSTICO	,098	,078	,739
	Í.ECONÓMICO	,331	-,355	,985
	Í.INDUSTRIAL	,386	-,445	,950
	PERNOCTACIONES	,093	,041	,786
	IMPORTACIONES	,383	-,368	,884
	POTENCIAL	,140	-,184	,930

		Í.RESTAURACIÓN	Í.TURÍSTICO	Í.ECONÓMICO
Correlación de Pearson	PASAJEROS	,445	,785	,400
	POBLACIÓN	,992	,734	,982
	RESIDENTES	,830	,907	,807
	RENTA	,228	,098	,331
	PARO	-,274	,078	-,355
	Í.COMERCIAL	,993	,739	,985
	Í.RESTAURACIÓN	1,000	,735	,987
	Í.TURÍSTICO	,735	1,000	,697
	Í.ECONÓMICO	,987	,697	1,000
	Í.INDUSTRIAL	,956	,585	,982
	PERNOCTACIONES	,777	,949	,745
	IMPORTACIONES	,878	,556	,908
	POTENCIAL	,919	,572	,919

		Í.INDUSTRIAL	PERNOCTACIONES	IMPORTACIONES
Correlación de Pearson	PASAJEROS	,259	,672	,321
	POBLACIÓN	,946	,778	,886
	RESIDENTES	,725	,922	,667
	RENTA	,386	,093	,383
	PARO	-,445	,041	-,368
	Í.COMERCIAL	,950	,786	,884
	Í.RESTAURACIÓN	,956	,777	,878
	Í.TURÍSTICO	,585	,949	,556
	Í.ECONÓMICO	,982	,745	,908
	Í.INDUSTRIAL	1,000	,665	,916
	PERNOCTACIONES	,665	1,000	,586
	IMPORTACIONES	,916	,586	1,000
	POTENCIAL	,900	,625	,860

		POTENCIAL
Correlación de Pearson	PASAJEROS	,294
	POBLACIÓN	,933
	RESIDENTES	,715
	RENTA	,140
	PARO	-,184
	Í.COMERCIAL	,930
	Í.RESTAURACIÓN	,919
	Í.TURÍSTICO	,572
	Í.ECONÓMICO	,919
	Í.INDUSTRIAL	,900
	PERNOCTACIONES	,625
	IMPORTACIONES	,860
	POTENCIAL	1,000

COEFICIENTES

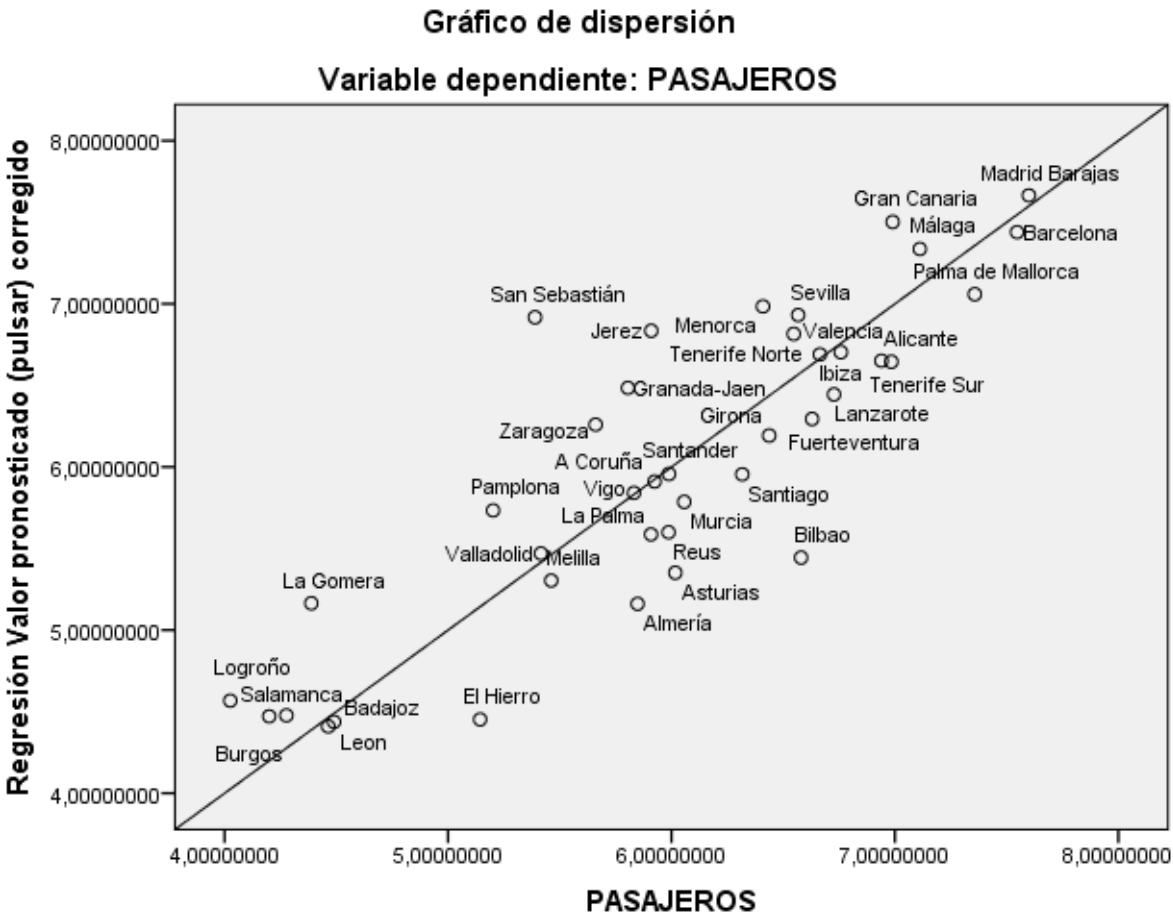
Modelo	Coeficientes no estandarizados		Coeficientes tipificados	t
	B	Error típ.	Beta	
(Constante)	-92,784	18,820		-4,930
POBLACIÓN	27,153	5,714	16,524	4,752
RESIDENTES	,292	,357	,175	,817
RENTA	4,818	2,429	,425	1,983
PARO	-2,032	1,829	-,251	-1,111
Í.COMERCIAL	-28,441	5,893	-17,113	-4,826
Í.RESTAURACIÓN	,484	1,129	,305	,428
Í.TURÍSTICO	,484	,588	,283	,823
Í.ECONÓMICO	-,041	2,287	-,027	-,018
Í.INDUSTRIAL	-2,379	1,255	-1,762	-1,896
PERNOCTACIONES	1,122	,597	,615	1,879
IMPORTACIONES	-,334	,334	-,270	-1,001
POTENCIAL	-,476	,327	-,336	-1,455

RESUMEN DEL MODELO (MÉTODO INTRODUCIR)

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
1	,955 ^a	,912	,860	,35434368334	1,694

a. Variables predictoras: (Constante), CUOTA, RENTA, Í.TURÍSTICO, PARO, IMPORTACIONES, RESIDENTES, POTENCIAL, PERNOCTACIONES, EXPORTACIONES, Í.INDUSTRIAL, Í.RESTAURACIÓN, Í.COMERCIAL, Í.ECONÓMICO, POBLACIÓN

b. Variable dependiente: PASAJEROS



RESUMEN DEL MODELO (MÉTODO HACIA ATRÁS)

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Estadísticos de cambio	
					Cambio en R cuadrado	Cambio en F
1	,955 ^a	,912	,860	,35434368334	,912	17,722
2	,955 ^b	,912	,866	,34718683279	,000	,000
3	,955 ^c	,911	,870	,34176221379	-,001	,194
4	,954 ^d	,909	,873	,33845681968	-,002	,480
5	,953 ^e	,907	,874	,33627950428	-,002	,641
6	,951 ^f	,904	,875	,33558816358	-,003	,881
7	,947 ^g	,897	,870	,34203483136	-,007	2,164
8	,945 ^h	,893	,868	,34416898908	-,005	1,388

a. Variables predictoras: (Constante), CUOTA, RENTA, Í.TURÍSTICO, PARO, IMPORTACIONES, RESIDENTES, POTENCIAL, PERNOCTACIONES, EXPORTACIONES, Í.INDUSTRIAL, Í.RESTAURACIÓN, Í.COMERCIAL, Í.ECONÓMICO, POBLACIÓN

b. Variables predictoras: (Constante), CUOTA, RENTA, Í.TURÍSTICO, PARO, IMPORTACIONES, RESIDENTES, POTENCIAL, PERNOCTACIONES, EXPORTACIONES, Í.INDUSTRIAL, Í.RESTAURACIÓN, Í.COMERCIAL, POBLACIÓN

c. Variables predictoras: (Constante), CUOTA, RENTA, Í.TURÍSTICO, PARO, IMPORTACIONES, RESIDENTES, POTENCIAL, PERNOCTACIONES, EXPORTACIONES, Í.INDUSTRIAL, Í.COMERCIAL, POBLACIÓN

d. Variables predictoras: (Constante), CUOTA, RENTA, Í.TURÍSTICO, PARO, IMPORTACIONES, RESIDENTES, POTENCIAL, PERNOCTACIONES, Í.INDUSTRIAL, Í.COMERCIAL, POBLACIÓN

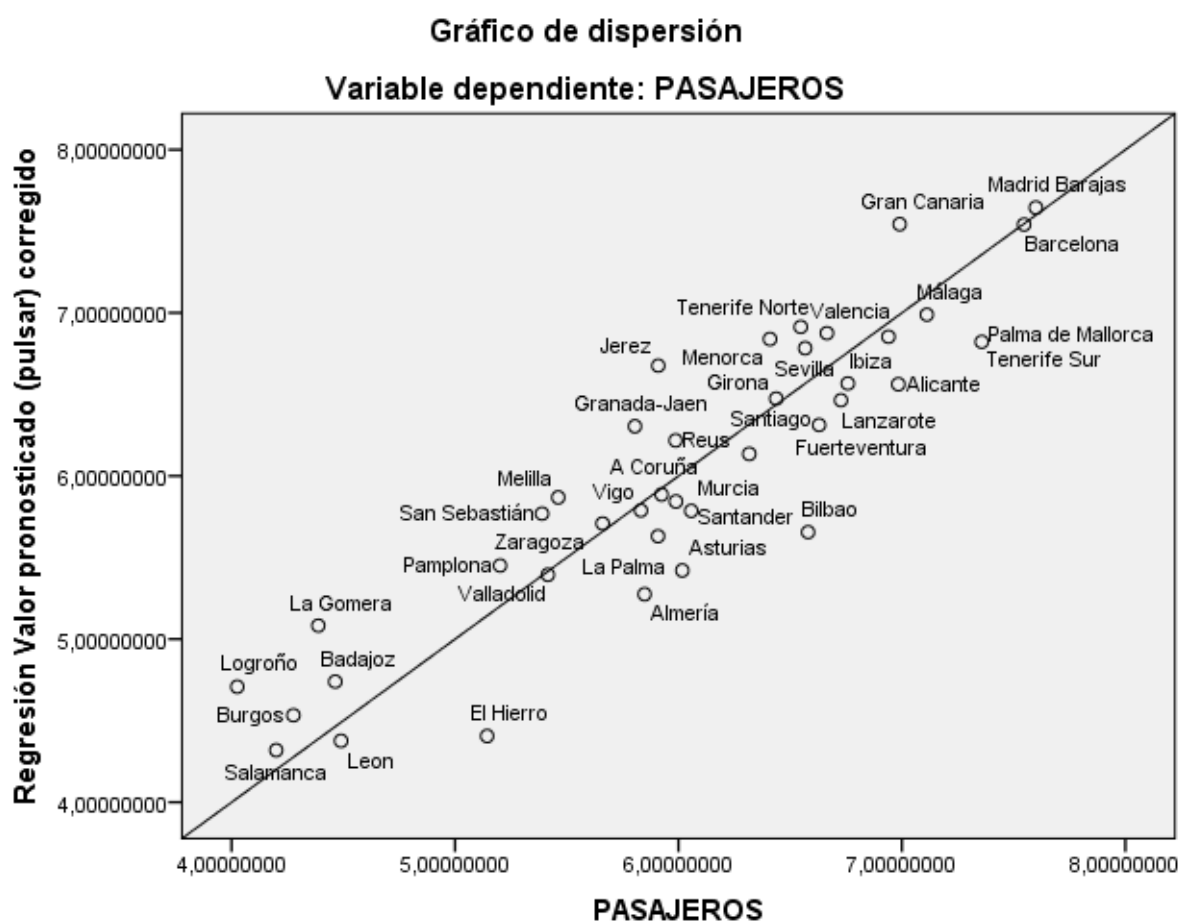
e. Variables predictoras: (Constante), CUOTA, RENTA, Í.TURÍSTICO, PARO, IMPORTACIONES, POTENCIAL, PERNOCTACIONES, Í.INDUSTRIAL, Í.COMERCIAL, POBLACIÓN

f. Variables predictoras: (Constante), CUOTA, RENTA, Í.TURÍSTICO, PARO, POTENCIAL, PERNOCTACIONES, Í.INDUSTRIAL, Í.COMERCIAL, POBLACIÓN

g. Variables predictoras: (Constante), CUOTA, RENTA, PARO, POTENCIAL, PERNOCTACIONES, Í.INDUSTRIAL, Í.COMERCIAL, POBLACIÓN

h. Variables predictoras: (Constante), CUOTA, RENTA, POTENCIAL, PERNOCTACIONES, Í.INDUSTRIAL, Í.COMERCIAL, POBLACIÓN

i. Variable dependiente: PASAJEROS



RESUMEN DEL MODELO (MÉTODO HACIA ADELANTE)

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Estadísticos de cambio	
					Cambio en R cuadrado	Cambio en F
1	,785 ^a	,617	,607	,59470928740	,617	59,600
2	,823 ^b	,678	,660	,55271558786	,061	6,836
3	,859 ^c	,737	,715	,50662000816	,059	7,849

a. Variables predictoras: (Constante), Í.TURÍSTICO

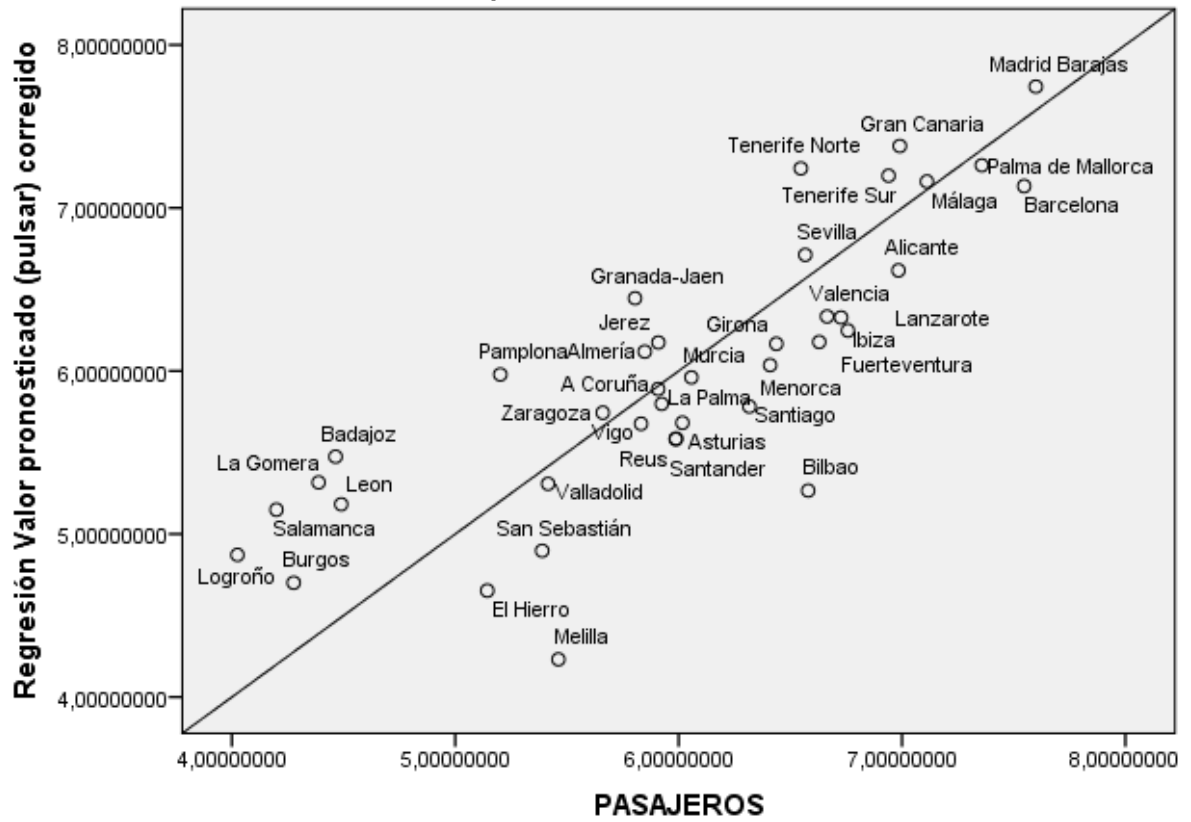
b. Variables predictoras: (Constante), Í.TURÍSTICO, Í.INDUSTRIAL

c. Variables predictoras: (Constante), Í.TURÍSTICO, Í.INDUSTRIAL, Í.ECONÓMICO

d. Variable dependiente: PASAJEROS

Gráfico de dispersión

Variable dependiente: PASAJEROS



5.2 RESULTADOS

MODELO 2

AÑO 2012

15 VARIABLES

CORRELACIONES

		PASAJEROS	POBLACIÓN	RESIDENTES
Correlación de Pearson	PASAJEROS	1,000	,485	,610
	POBLACIÓN	,485	1,000	,884
	RESIDENTES	,610	,884	1,000
	RENTA	,048	,176	,140
	PARO	,168	-,198	-,033
	Í.COMERCIAL	,483	,999	,891
	Í.RESTAUACIÓN	,475	,992	,891
	Í.TURÍSTICO	,792	,733	,896
	Í.ECONÓMICO	,429	,984	,868
	Í.INDUSTRIAL	,288	,948	,799
	PERNOCTACIONES	,684	,777	,918
	IMPORTACIONES	,358	,886	,725
	POTENCIAL	,393	,957	,821

		RENTA	PARO	Í.COMERCIAL
Correlación de Pearson	PASAJEROS	,048	,168	,483
	POBLACIÓN	,176	-,198	,999
	RESIDENTES	,140	-,033	,891
	RENTA	1,000	-,820	,191
	PARO	-,820	1,000	-,208
	Í.COMERCIAL	,191	-,208	1,000
	Í.RESTAUACIÓN	,218	-,259	,993
	Í.TURÍSTICO	,084	,105	,741
	Í.ECONÓMICO	,320	-,327	,986
	Í.INDUSTRIAL	,379	-,425	,951
	PERNOCTACIONES	,083	,077	,786
	IMPORTACIONES	,387	-,384	,883
	POTENCIAL	,226	-,260	,956

		Í.RESTAURACIÓN	Í.TURÍSTICO	Í.ECONÓMICO
Correlación de Pearson	PASAJEROS	,475	,792	,429
	POBLACIÓN	,992	,733	,984
	RESIDENTES	,891	,896	,868
	RENTA	,218	,084	,320
	PARO	-,259	,105	-,327
	Í.COMERCIAL	,993	,741	,986
	Í.RESTAURACIÓN	1,000	,738	,986
	Í.TURÍSTICO	,738	1,000	,699
	Í.ECONÓMICO	,986	,699	1,000
	Í.INDUSTRIAL	,955	,588	,982
	PERNOCTACIONES	,779	,950	,746
	IMPORTACIONES	,882	,556	,911
	POTENCIAL	,959	,633	,951

		Í.INDUSTRIAL	PERNOCTACIONES	IMPORTACIONES
Correlación de Pearson	PASAJEROS	,288	,684	,358
	POBLACIÓN	,948	,777	,886
	RESIDENTES	,799	,918	,725
	RENTA	,379	,083	,387
	PARO	-,425	,077	-,384
	Í.COMERCIAL	,951	,786	,883
	Í.RESTAURACIÓN	,955	,779	,882
	Í.TURÍSTICO	,588	,950	,556
	Í.ECONÓMICO	,982	,746	,911
	Í.INDUSTRIAL	1,000	,668	,917
	PERNOCTACIONES	,668	1,000	,592
	IMPORTACIONES	,917	,592	1,000
	POTENCIAL	,927	,687	,859

		POTENCIAL
Correlación de Pearson	PASAJEROS	,393
	POBLACIÓN	,957
	RESIDENTES	,821
	RENTA	,226
	PARO	-,260
	Í.COMERCIAL	,956
	Í.RESTAURACIÓN	,959
	Í.TURÍSTICO	,633
	Í.ECONÓMICO	,951
	Í.INDUSTRIAL	,927
	PERNOCTACIONES	,687
	IMPORTACIONES	,859
	POTENCIAL	1,000

COEFICIENTES

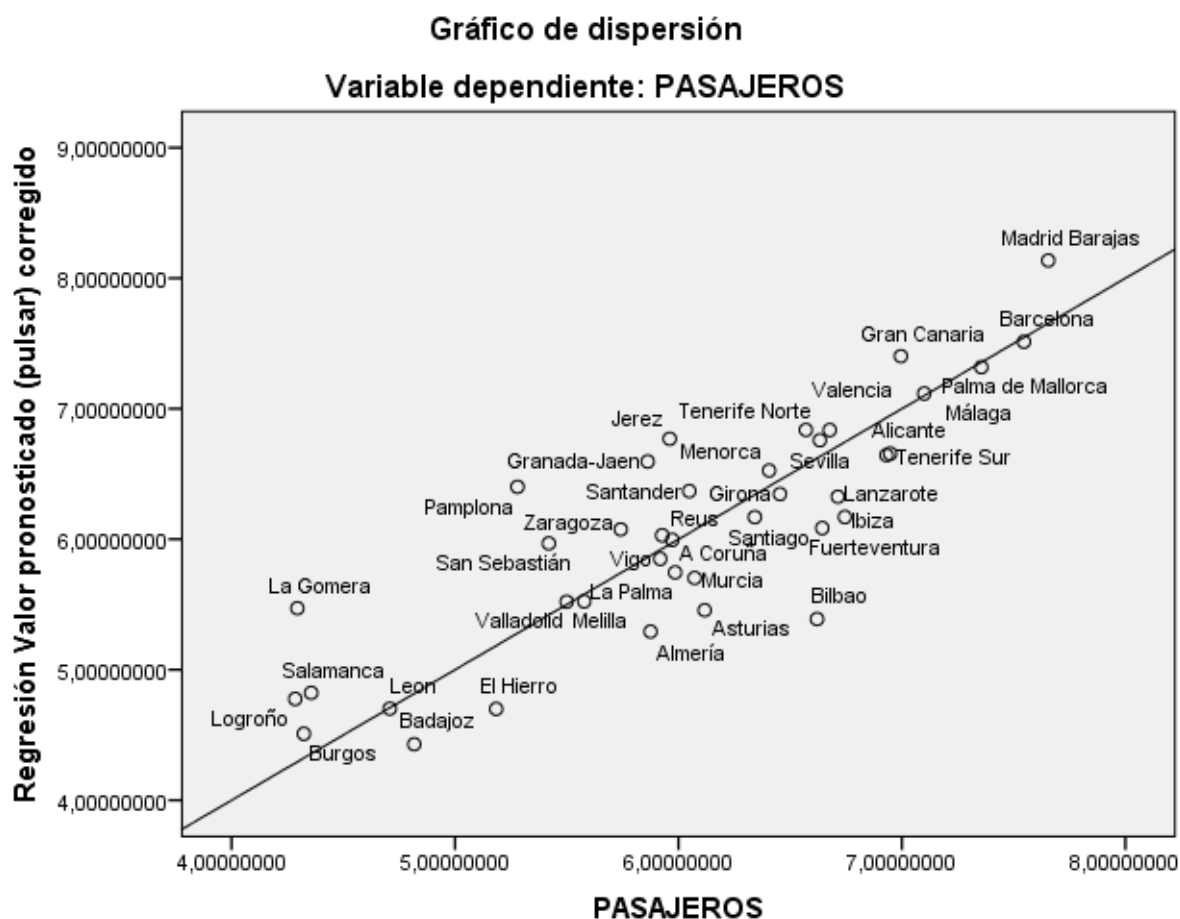
Modelo	Coeficientes no estandarizados		Coeficientes tipificados	t
	B	Error típ.	Beta	
(Constante)	-78,219	22,050		-3,547
POBLACIÓN	20,814	5,858	13,330	3,553
RESIDENTES	,236	,460	,148	,513
RENTA	4,950	2,457	,454	2,014
PARO	-1,540	1,676	-,208	-,919
Í.COMERCIAL	-22,187	6,258	-14,033	-3,546
Í.RESTAURACIÓN	1,282	1,442	,850	,889
Í.TURÍSTICO	-,028	,643	-,017	-,043
Í.ECONÓMICO	1,673	2,174	1,168	,769
Í.INDUSTRIAL	-3,980	1,287	-3,092	-3,094
PERNOCTACIONES	1,412	,607	,813	2,328
IMPORTACIONES	-,273	,303	-,231	-,902
POTENCIAL	-,644	,462	-,425	-1,393

RESUMEN DEL MODELO (MÉTODO INTRODUCIR)

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
1	,943 ^a	,890	,826	,37653862467	1,822

a. Variables predictoras: (Constante), CUOTA, RENTA, Í.TURÍSTICO, PARO, IMPORTACIONES, RESIDENTES, PERNOCTACIONES, POTENCIAL, EXPORTACIONES, Í.INDUSTRIAL, Í.COMERCIAL, Í.RESTAURACIÓN, Í.ECONÓMICO, POBLACIÓN

b. Variable dependiente: PASAJEROS



RESUMEN DEL MODELO (MÉTODO HACIA ATRÁS)

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Estadísticos de cambio	
					Cambio en R cuadrado	Cambio en F
1	,943 ^a	,890	,826	,37653862467	,890	13,863
2	,943 ^b	,890	,833	,36894518501	,000	,002
3	,943 ^c	,889	,837	,36375984265	-,001	,274
4	,941 ^d	,886	,839	,36159639322	-,003	,680
5	,940 ^e	,884	,842	,35856962592	-,002	,533
6	,939 ^f	,881	,844	,35644321093	-,003	,657
7	,938 ^g	,880	,847	,35238522058	-,001	,321

a. Variables predictoras: (Constante), CUOTA, RENTA, Í.TURÍSTICO, PARO, IMPORTACIONES, RESIDENTES, PERNOCTACIONES, POTENCIAL, EXPORTACIONES, Í.INDUSTRIAL, Í.COMERCIAL, Í.RESTAURACIÓN, Í.ECONÓMICO, POBLACIÓN

b. Variables predictoras: (Constante), CUOTA, RENTA, PARO, IMPORTACIONES, RESIDENTES, PERNOCTACIONES, POTENCIAL, EXPORTACIONES, Í.INDUSTRIAL, Í.COMERCIAL, Í.RESTAURACIÓN, Í.ECONÓMICO, POBLACIÓN

c. Variables predictoras: (Constante), CUOTA, RENTA, PARO, IMPORTACIONES, PERNOCTACIONES, POTENCIAL, EXPORTACIONES, Í.INDUSTRIAL, Í.COMERCIAL, Í.RESTAURACIÓN, Í.ECONÓMICO, POBLACIÓN

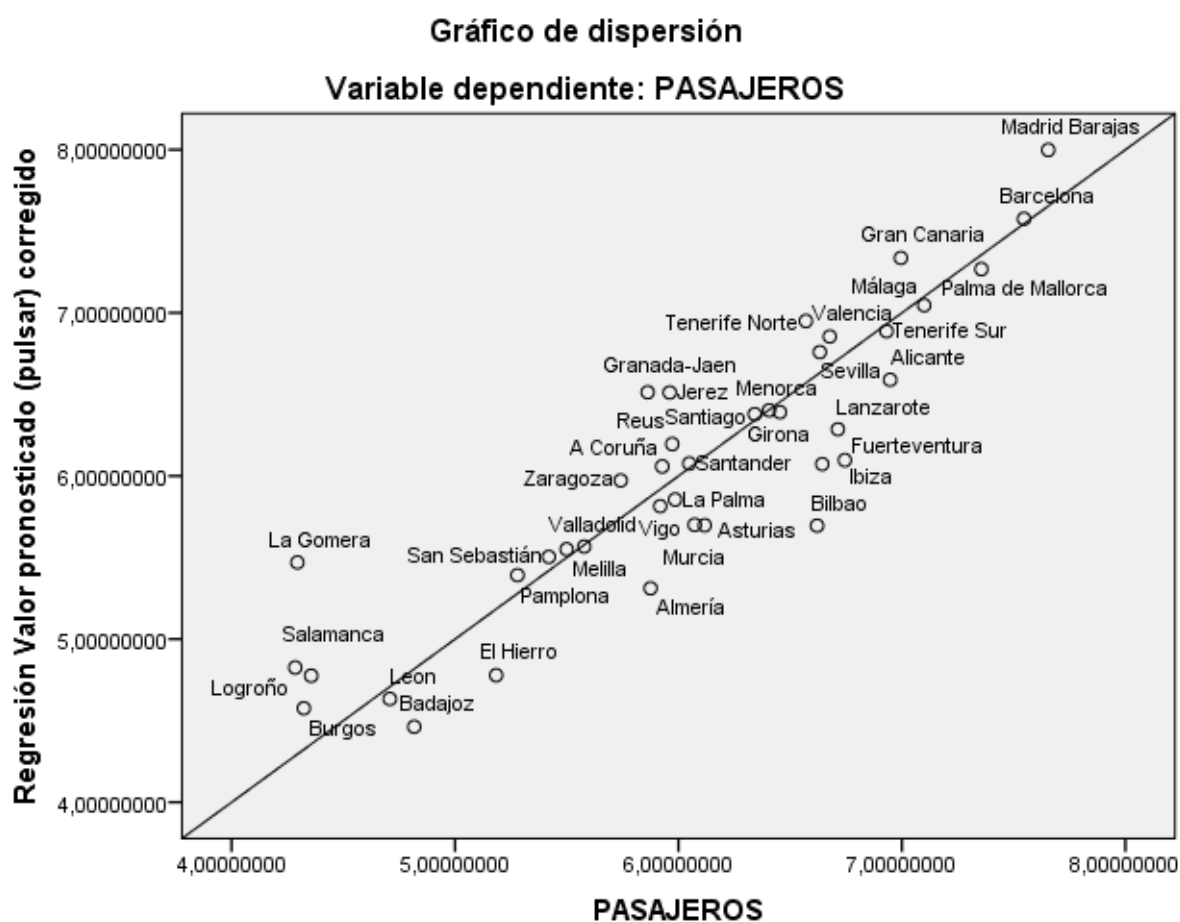
d. Variables predictoras: (Constante), CUOTA, RENTA, IMPORTACIONES, PERNOCTACIONES, POTENCIAL, EXPORTACIONES, Í.INDUSTRIAL, Í.COMERCIAL, Í.RESTAURACIÓN, Í.ECONÓMICO, POBLACIÓN

e. Variables predictoras: (Constante), CUOTA, RENTA, IMPORTACIONES, PERNOCTACIONES, POTENCIAL, EXPORTACIONES, Í.INDUSTRIAL, Í.COMERCIAL, Í.RESTAURACIÓN, POBLACIÓN

f. Variables predictoras: (Constante), CUOTA, RENTA, IMPORTACIONES, PERNOCTACIONES, POTENCIAL, Í.INDUSTRIAL, Í.COMERCIAL, Í.RESTAURACIÓN, POBLACIÓN

g. Variables predictoras: (Constante), CUOTA, RENTA, PERNOCTACIONES, POTENCIAL, Í.INDUSTRIAL, Í.COMERCIAL, Í.RESTAURACIÓN, POBLACIÓN

h. Variable dependiente: PASAJEROS



RESUMEN DEL MODELO (MÉTODO HACIA ADELANTE)

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Estadísticos de cambio	
					Cambio en R cuadrado	Cambio en F
1	,792 ^a	,627	,617	,55847449071	,627	62,136
2	,823 ^b	,677	,659	,52642215313	,051	5,643

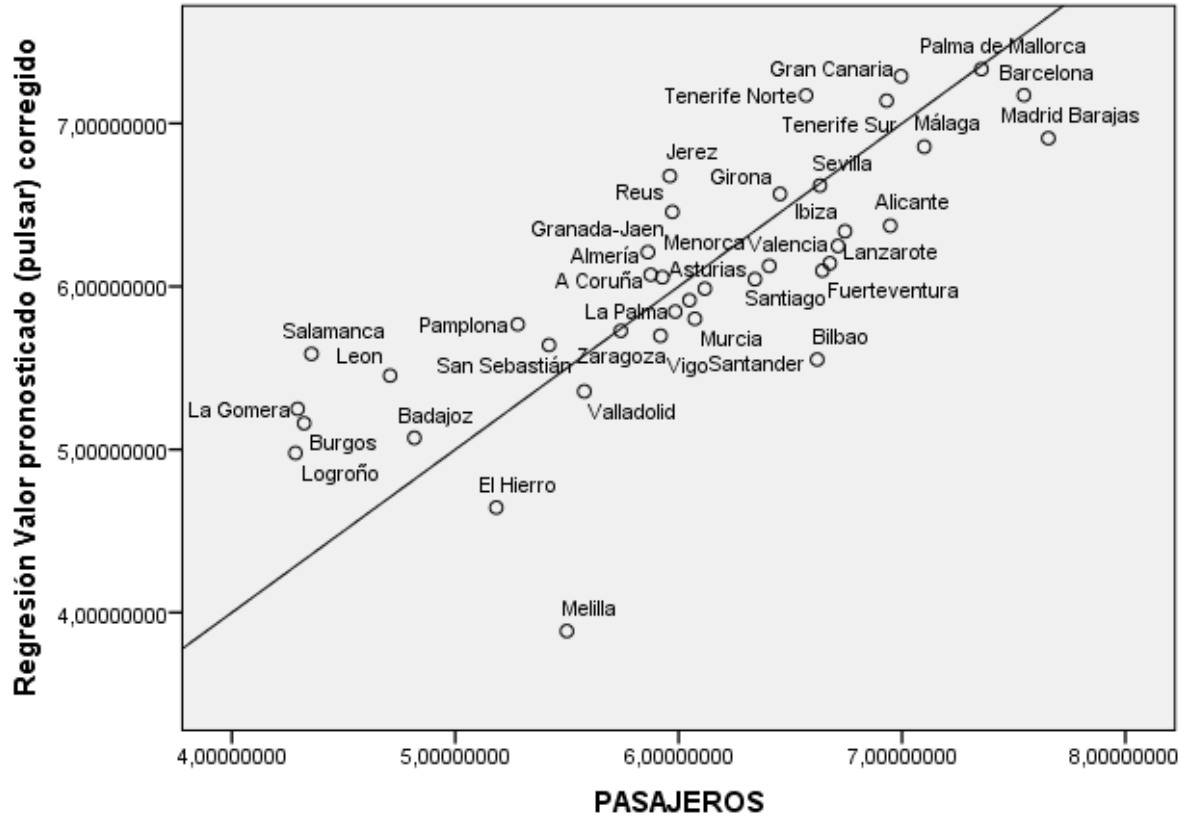
a. Variables predictoras: (Constante), Í.TURÍSTICO

b. Variables predictoras: (Constante), Í.TURÍSTICO, RESIDENTES

c. Variable dependiente: PASAJEROS

Gráfico de dispersión

Variable dependiente: PASAJEROS



5.3 RESULTADOS MODELO 3 AÑO 2011 15 VARIABLES

CORRELACIONES

		PASAJEROS	POBLACIÓN	RESIDENTES
Correlación de Pearson	PASAJEROS	1,000	,491	,607
	POBLACIÓN	,491	1,000	,894
	RESIDENTES	,607	,894	1,000
	RENTA	,049	,168	,147
	PARO	,296	-,201	,016
	Í.COMERCIAL	,487	,999	,898
	Í.RESTAURACIÓN	,482	,992	,897
	Í.TURÍSTICO	,804	,739	,886
	Í.ECONÓMICO	,440	,985	,879
	Í.INDUSTRIAL	,306	,950	,817
	PERNOCTACIONES	,695	,780	,914
	IMPORTACIONES	,374	,895	,749
	POTENCIAL	,407	,958	,836

		RENTA	PARO	Í.COMERCIAL
Correlación de Pearson	PASAJEROS	,049	,296	,487
	POBLACIÓN	,168	-,201	,999
	RESIDENTES	,147	,016	,898
	RENTA	1,000	-,757	,185
	PARO	-,757	1,000	-,211
	Í.COMERCIAL	,185	-,211	1,000
	Í.RESTAURACIÓN	,205	-,236	,993
	Í.TURÍSTICO	,056	,232	,744
	Í.ECONÓMICO	,309	-,319	,987
	Í.INDUSTRIAL	,377	-,432	,953
	PERNOCTACIONES	,067	,185	,787
	IMPORTACIONES	,404	-,407	,893
	POTENCIAL	,216	-,242	,958

		Í.RESTAURACIÓN	Í.TURÍSTICO	Í.ECONÓMICO
Correlación de Pearson	PASAJEROS	,482	,804	,440
	POBLACIÓN	,992	,739	,985
	RESIDENTES	,897	,886	,879
	RENTA	,205	,056	,309
	PARO	-,236	,232	-,319
	Í.COMERCIAL	,993	,744	,987
	Í.RESTAURACIÓN	1,000	,736	,986
	Í.TURÍSTICO	,736	1,000	,702
	Í.ECONÓMICO	,986	,702	1,000
	Í.INDUSTRIAL	,956	,596	,983
	PERNOCTACIONES	,777	,953	,748
	IMPORTACIONES	,894	,569	,925
	POTENCIAL	,959	,641	,953

		Í.INDUSTRIAL	PERNOCTACIONES	IMPORTACIONES
Correlación de Pearson	PASAJEROS	,306	,695	,374
	POBLACIÓN	,950	,780	,895
	RESIDENTES	,817	,914	,749
	RENTA	,377	,067	,404
	PARO	-,432	,185	-,407
	Í.COMERCIAL	,953	,787	,893
	Í.RESTAURACIÓN	,956	,777	,894
	Í.TURÍSTICO	,596	,953	,569
	Í.ECONÓMICO	,983	,748	,925
	Í.INDUSTRIAL	1,000	,674	,932
	PERNOCTACIONES	,674	1,000	,608
	IMPORTACIONES	,932	,608	1,000
	POTENCIAL	,926	,696	,865

		POTENCIAL
Correlación de Pearson	PASAJEROS	,407
	POBLACIÓN	,958
	RESIDENTES	,836
	RENTA	,216
	PARO	-,242
	Í.COMERCIAL	,958
	Í.RESTAURACIÓN	,959
	Í.TURÍSTICO	,641
	Í.ECONÓMICO	,953
	Í.INDUSTRIAL	,926
	PERNOCTACIONES	,696
	IMPORTACIONES	,865
	POTENCIAL	1,000

COEFICIENTES

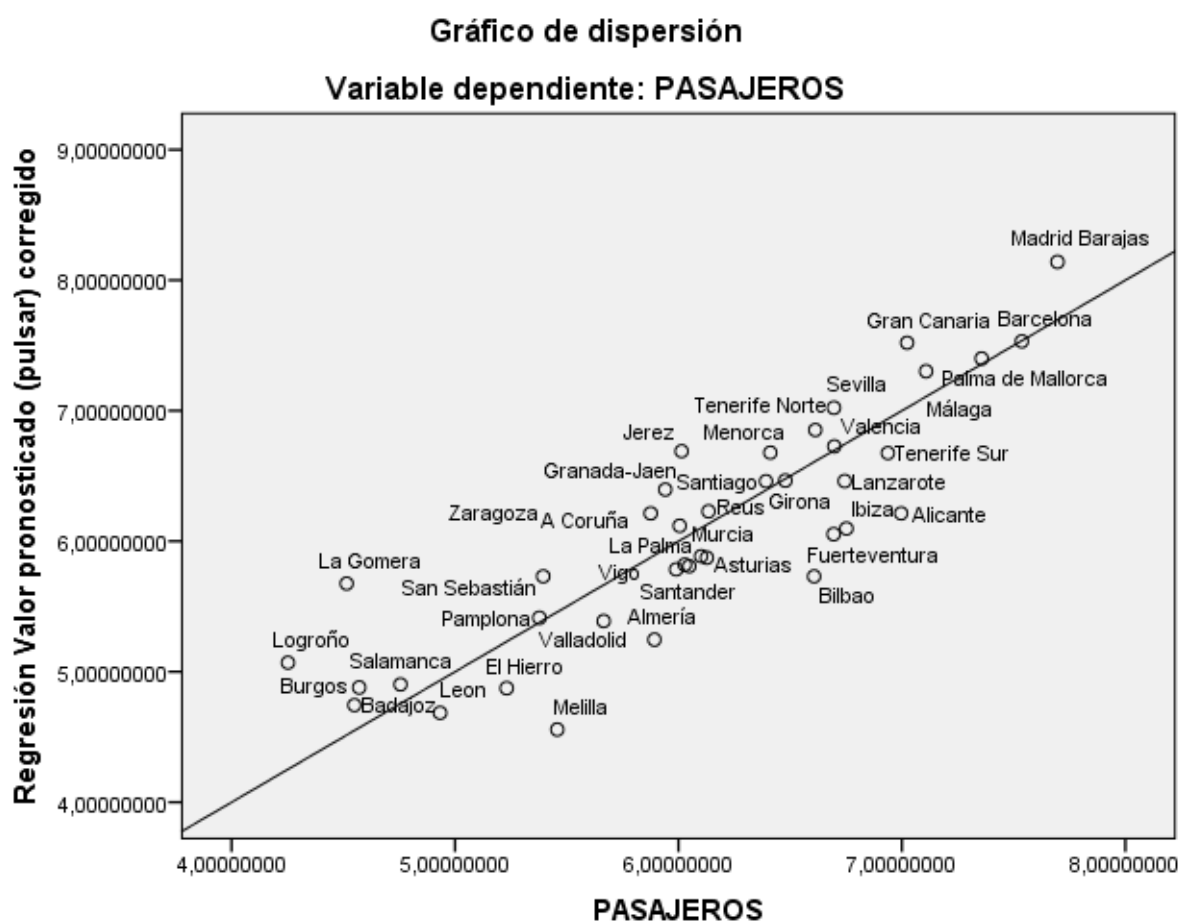
Modelo	Coeficientes no estandarizados		Coeficientes tipificados	t
	B	Error típ.	Beta	
(Constante)	-79,238	19,023		-4,165
POBLACIÓN	15,034	5,048	10,008	2,978
RESIDENTES	-,428	,399	-,283	-1,072
RENTA	8,618	2,214	,800	3,892
PARO	1,747	1,516	,265	1,152
Í.COMERCIAL	-16,131	5,206	-10,598	-3,099
Í.RESTAURACIÓN	1,871	1,012	1,292	1,848
Í.TURÍSTICO	,421	,614	,266	,685
Í.ECONÓMICO	-,549	2,162	-,397	-,254
Í.INDUSTRIAL	-1,957	1,318	-1,561	-1,485
PERNOCTACIONES	,730	,645	,436	1,132
IMPORTACIONES	-,176	,284	-,156	-,620
POTENCIAL	-,745	,470	-,509	-1,587

RESUMEN DEL MODELO (MÉTODO INTRODUCIR)

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
1	,946 ^a	,895	,833	,35451572390	1,927

a. Variables predictoras: (Constante), CUOTA, RENTA, Í.TURÍSTICO, PARO, IMPORTACIONES, EXPORTACIONES, RESIDENTES, PERNOCTACIONES, POTENCIAL, Í.RESTAURACIÓN, Í.INDUSTRIAL, Í.COMERCIAL, Í.ECONÓMICO, POBLACIÓN

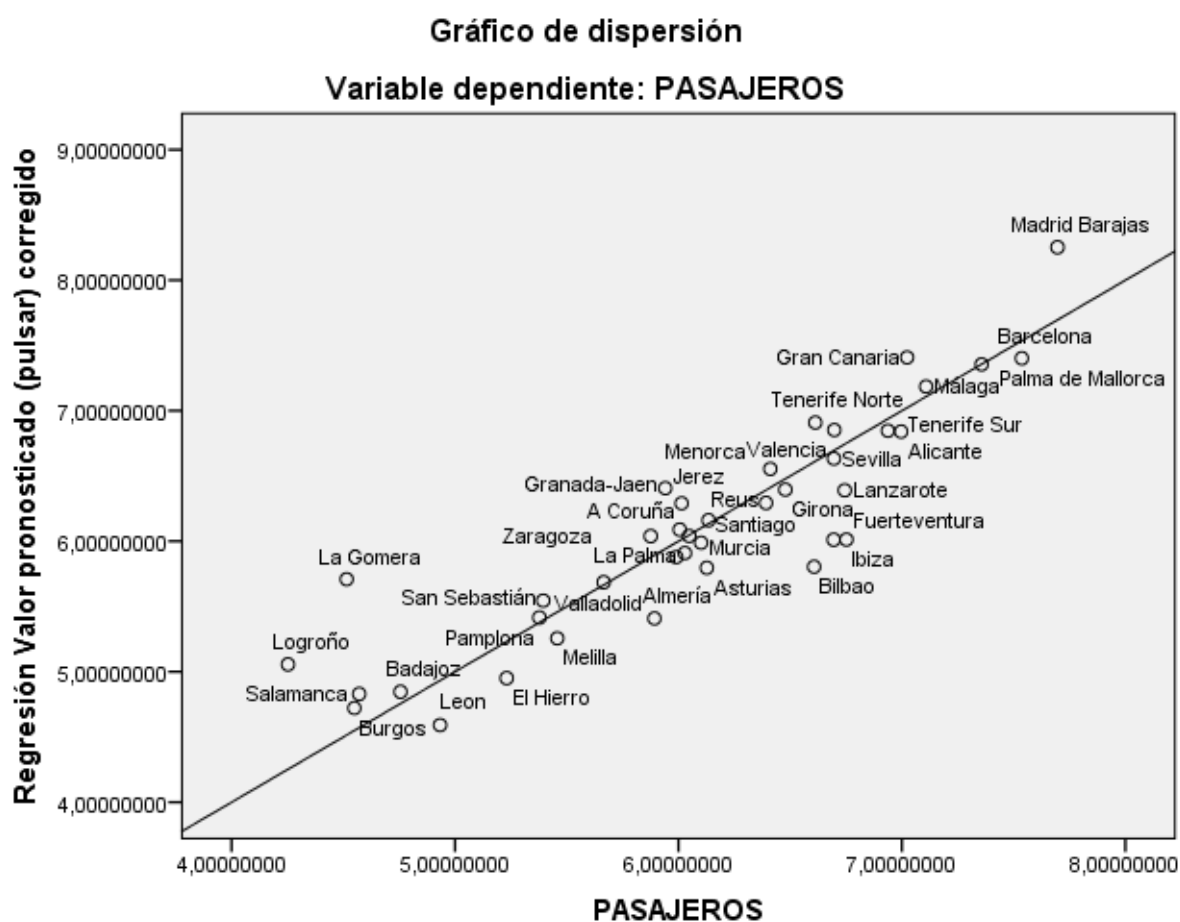
b. Variable dependiente: PASAJEROS



RESUMEN DEL MODELO (MÉTODO HACIA ATRÁS)

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Estadísticos de cambio	
					Cambio en R cuadrado	Cambio en F
1	,946 ^a	,895	,833	,35451572390	,895	14,584
2	,946 ^b	,895	,840	,34778692032	,000	,060
3	,946 ^c	,894	,845	,34153774667	,000	,074
4	,945 ^d	,893	,849	,33757604064	-,002	,377
5	,944 ^e	,891	,852	,33466545393	-,002	,519
6	,942 ^f	,887	,852	,33460312167	-,004	,989
7	,940 ^g	,883	,852	,33410631910	-,004	,911

- a. Variables predictorias: (Constante), CUOTA, RENTA, Í.TURÍSTICO, PARO, IMPORTACIONES, EXPORTACIONES, RESIDENTES, PERNOCTACIONES, POTENCIAL, Í.RESTAUACIÓN, Í.INDUSTRIAL, Í.COMERCIAL, Í.ECONÓMICO, POBLACIÓN
- b. Variables predictorias: (Constante), CUOTA, RENTA, Í.TURÍSTICO, PARO, IMPORTACIONES, RESIDENTES, PERNOCTACIONES, POTENCIAL, Í.RESTAUACIÓN, Í.INDUSTRIAL, Í.COMERCIAL, Í.ECONÓMICO, POBLACIÓN
- c. Variables predictorias: (Constante), CUOTA, RENTA, Í.TURÍSTICO, PARO, IMPORTACIONES, RESIDENTES, PERNOCTACIONES, POTENCIAL, Í.RESTAUACIÓN, Í.INDUSTRIAL, Í.COMERCIAL, POBLACIÓN
- d. Variables predictorias: (Constante), CUOTA, RENTA, PARO, IMPORTACIONES, RESIDENTES, PERNOCTACIONES, POTENCIAL, Í.RESTAUACIÓN, Í.INDUSTRIAL, Í.COMERCIAL, POBLACIÓN
- e. Variables predictorias: (Constante), CUOTA, RENTA, PARO, RESIDENTES, PERNOCTACIONES, POTENCIAL, Í.RESTAUACIÓN, Í.INDUSTRIAL, Í.COMERCIAL, POBLACIÓN
- f. Variables predictorias: (Constante), CUOTA, RENTA, PARO, PERNOCTACIONES, POTENCIAL, Í.RESTAUACIÓN, Í.INDUSTRIAL, Í.COMERCIAL, POBLACIÓN
- g. Variables predictorias: (Constante), CUOTA, RENTA, PERNOCTACIONES, POTENCIAL, Í.RESTAUACIÓN, Í.INDUSTRIAL, Í.COMERCIAL, POBLACIÓN
- h. Variable dependiente: PASAJEROS



RESUMEN DEL MODELO (MÉTODO HACIA ADELANTE)

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Estadísticos de cambio	
					Cambio en R cuadrado	Cambio en F
1	,804 ^a	,647	,637	,52338864552	,647	67,689
2	,837 ^b	,701	,685	,48772636586	,055	6,609

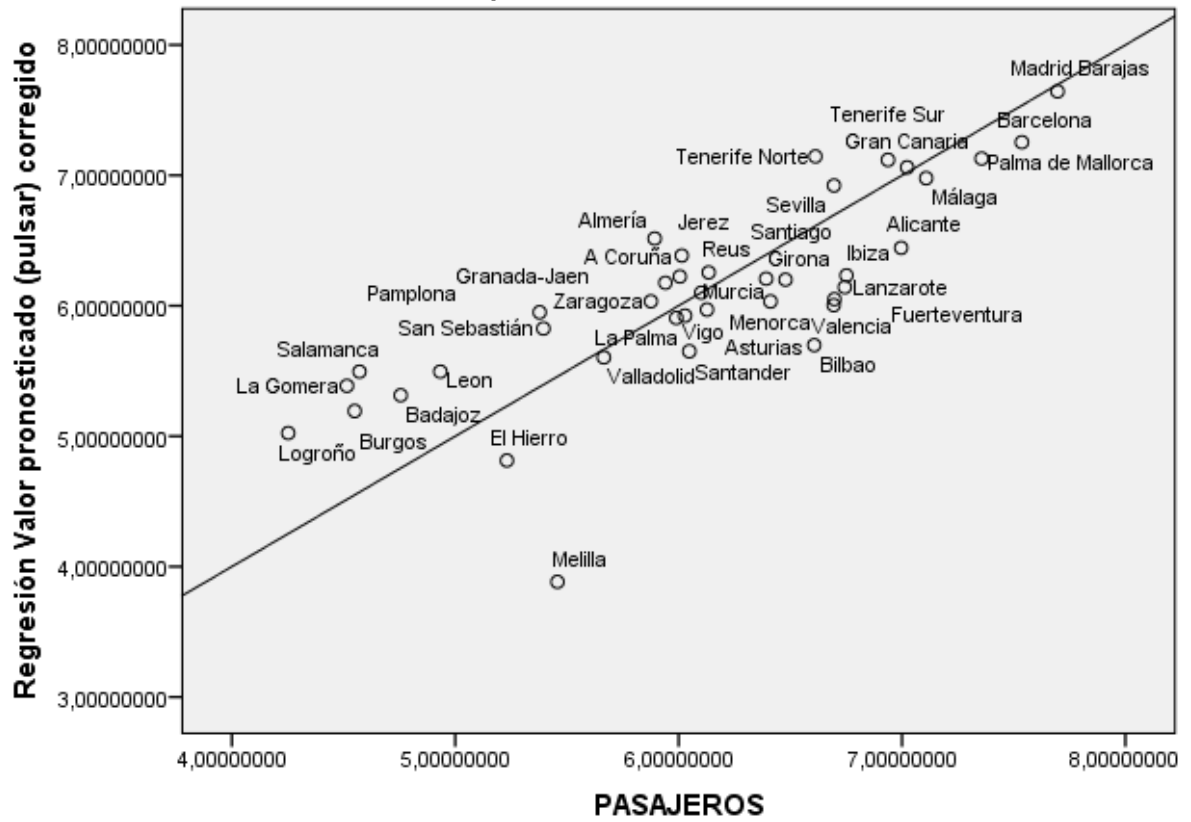
a. Variables predictoras: (Constante), Í.TURÍSTICO

b. Variables predictoras: (Constante), Í.TURÍSTICO, PERNOCTACIONES

c. Variable dependiente: PASAJEROS

Gráfico de dispersión

Variable dependiente: PASAJEROS



5.4 RESULTADOS

MODELO 4

AÑO 2010

15 VARIABLES

CORRELACIONES

		PASAJEROS	POBLACIÓN	RESIDENTES
Correlación de Pearson	PASAJEROS	1,000	,500	,616
	POBLACIÓN	,500	1,000	,893
	RESIDENTES	,616	,893	1,000
	RENTA	,081	,183	,172
	PARO	,278	-,225	-,008
	Í.COMERCIAL	,501	,999	,899
	Í.RESTAURACIÓN	,498	,993	,901
	Í.TURÍSTICO	,807	,744	,886
	Í.ECONÓMICO	,464	,986	,883
	Í.INDUSTRIAL	,325	,952	,822
	PERNOCTACIONES	,715	,758	,905
	IMPORTACIONES	,402	,901	,756
	POTENCIAL	,421	,954	,837

		RENTA	PARO	Í.COMERCIAL
Correlación de Pearson	PASAJEROS	,081	,278	,501
	POBLACIÓN	,183	-,225	,999
	RESIDENTES	,172	-,008	,899
	RENTA	1,000	-,766	,198
	PARO	-,766	1,000	-,232
	Í.COMERCIAL	,198	-,232	1,000
	Í.RESTAURACIÓN	,213	-,252	,994
	Í.TURÍSTICO	,060	,223	,751
	Í.ECONÓMICO	,322	-,338	,988
	Í.INDUSTRIAL	,387	-,458	,955
	PERNOCTACIONES	,039	,212	,767
	IMPORTACIONES	,416	-,424	,899
	POTENCIAL	,225	-,259	,954

		Í.RESTAURACIÓN	Í.TURÍSTICO	Í.ECONÓMICO
Correlación de Pearson	PASAJEROS	,498	,807	,464
	POBLACIÓN	,993	,744	,986
	RESIDENTES	,901	,886	,883
	RENTA	,213	,060	,322
	PARO	-,252	,223	-,338
	Í.COMERCIAL	,994	,751	,988
	Í.RESTAURACIÓN	1,000	,751	,987
	Í.TURÍSTICO	,751	1,000	,715
	Í.ECONÓMICO	,987	,715	1,000
	Í.INDUSTRIAL	,955	,608	,982
	PERNOCTACIONES	,763	,957	,727
	IMPORTACIONES	,899	,592	,926
	POTENCIAL	,955	,653	,950

		Í.INDUSTRIAL	PERNOCTACIONES	IMPORTACIONES
Correlación de Pearson	PASAJEROS	,325	,715	,402
	POBLACIÓN	,952	,758	,901
	RESIDENTES	,822	,905	,756
	RENTA	,387	,039	,416
	PARO	-,458	,212	-,424
	Í.COMERCIAL	,955	,767	,899
	Í.RESTAURACIÓN	,955	,763	,899
	Í.TURÍSTICO	,608	,957	,592
	Í.ECONÓMICO	,982	,727	,926
	Í.INDUSTRIAL	1,000	,646	,931
	PERNOCTACIONES	,646	1,000	,594
	IMPORTACIONES	,931	,594	1,000
	POTENCIAL	,922	,677	,866

		POTENCIAL
Correlación de Pearson	PASAJEROS	,421
	POBLACIÓN	,954
	RESIDENTES	,837
	RENTA	,225
	PARO	-,259
	Í.COMERCIAL	,954
	Í.RESTAURACIÓN	,955
	Í.TURÍSTICO	,653
	Í.ECONÓMICO	,950
	Í.INDUSTRIAL	,922
	PERNOCTACIONES	,677
	IMPORTACIONES	,866
	POTENCIAL	1,000

COEFICIENTES

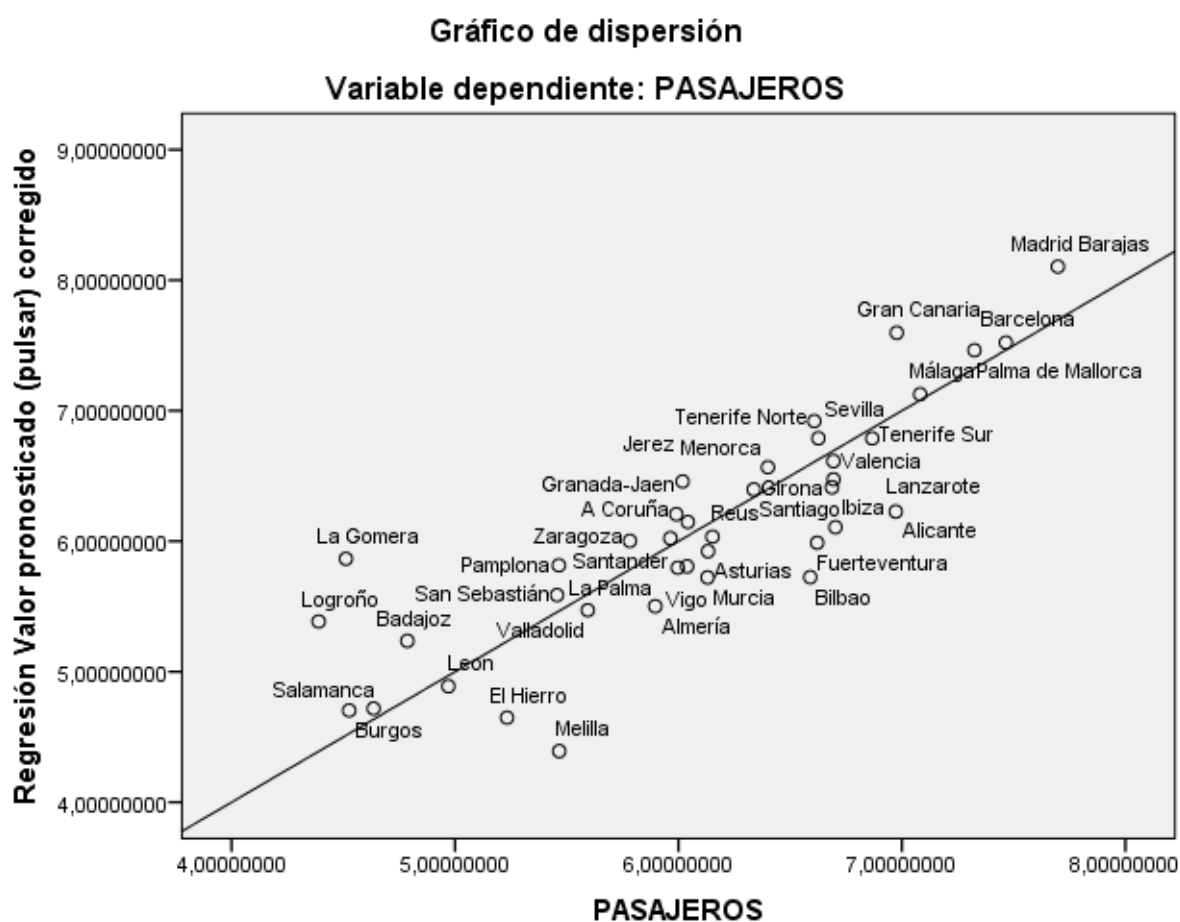
Modelo	Coeficientes no estandarizados		Coeficientes tipificados	t
	B	Error típ.	Beta	
(Constante)	-64,852	20,227		-3,206
POBLACIÓN	10,386	4,720	7,117	2,201
RESIDENTES	-,591	,422	-,404	-1,401
RENTA	8,147	2,822	,746	2,887
PARO	1,204	1,823	,192	,660
Í.COMERCIAL	-11,326	5,039	-7,668	-2,248
Í.RESTAURACIÓN	1,446	1,182	1,021	1,223
Í.TURÍSTICO	,364	,611	,236	,596
Í.ECONÓMICO	,194	2,362	,144	,082
Í.INDUSTRIAL	-2,466	1,277	-2,014	-1,931
PERNOCTACIONES	,905	,678	,555	1,335
IMPORTACIONES	-,231	,317	-,207	-,727
POTENCIAL	-,604	,471	-,419	-1,282

RESUMEN DEL MODELO (MÉTODO INTRODUCIR)

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
1	,941 ^a	,885	,818	,36027683811	1,959

a. Variables predictoras: (Constante), CUOTA, RENTA, Í.TURÍSTICO, PARO, IMPORTACIONES, EXPORTACIONES, RESIDENTES, PERNOCTACIONES, POTENCIAL, Í.INDUSTRIAL, Í.RESTAURACIÓN, Í.COMERCIAL, Í.ECONÓMICO, POBLACIÓN

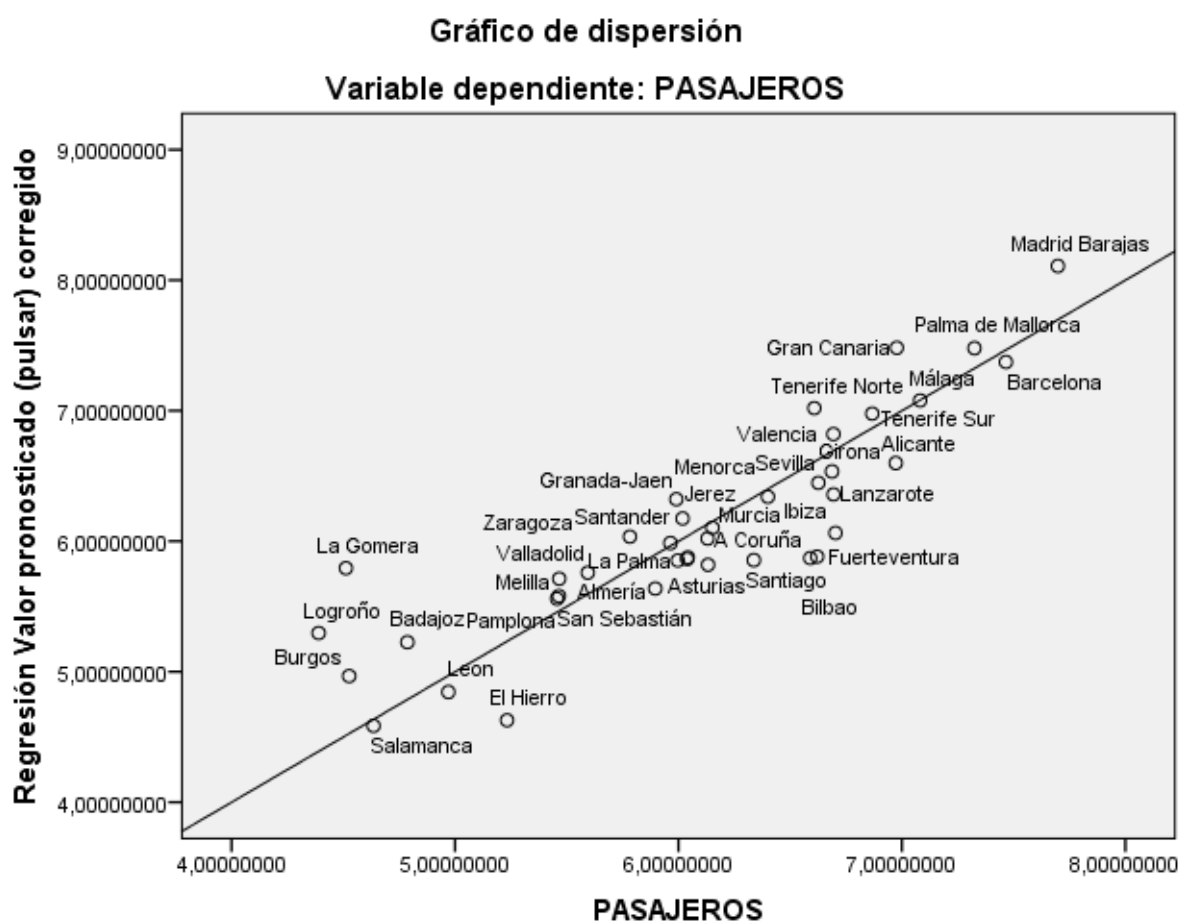
b. Variable dependiente: PASAJEROS



RESUMEN DEL MODELO (MÉTODO HACIA ATRÁS)

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Estadísticos de cambio	
					Cambio en R cuadrado	Cambio en F
1	,941 ^a	,885	,818	,36027683811	,885	13,198
2	,941 ^b	,885	,825	,35304742152	,000	,007
3	,941 ^c	,885	,832	,34651064936	,000	,046
4	,939 ^d	,882	,834	,34421738253	-,003	,644
5	,937 ^e	,879	,836	,34239970432	-,003	,705
6	,935 ^f	,875	,836	,34223536724	-,004	,972
7	,933 ^g	,870	,836	,34204391622	-,004	,966
8	,929 ^h	,863	,832	,34606425509	-,007	1,733
9	,924 ⁱ	,853	,825	,35285148475	-,010	2,268

- a. Variables predictorias: (Constante), CUOTA, RENTA, Í.TURÍSTICO, PARO, IMPORTACIONES, EXPORTACIONES, RESIDENTES, PERNOCTACIONES, POTENCIAL, Í.INDUSTRIAL, Í.RESTAUACIÓN, Í.COMERCIAL, Í.ECONÓMICO, POBLACIÓN
- b. Variables predictorias: (Constante), CUOTA, RENTA, Í.TURÍSTICO, PARO, IMPORTACIONES, EXPORTACIONES, RESIDENTES, PERNOCTACIONES, POTENCIAL, Í.INDUSTRIAL, Í.RESTAUACIÓN, Í.COMERCIAL, POBLACIÓN
- c. Variables predictorias: (Constante), CUOTA, RENTA, Í.TURÍSTICO, PARO, IMPORTACIONES, RESIDENTES, PERNOCTACIONES, POTENCIAL, Í.INDUSTRIAL, Í.RESTAUACIÓN, Í.COMERCIAL, POBLACIÓN
- d. Variables predictorias: (Constante), CUOTA, RENTA, PARO, IMPORTACIONES, RESIDENTES, PERNOCTACIONES, POTENCIAL, Í.INDUSTRIAL, Í.RESTAUACIÓN, Í.COMERCIAL, POBLACIÓN
- e. Variables predictorias: (Constante), CUOTA, RENTA, PARO, RESIDENTES, PERNOCTACIONES, POTENCIAL, Í.INDUSTRIAL, Í.RESTAUACIÓN, Í.COMERCIAL, POBLACIÓN
- f. Variables predictorias: (Constante), CUOTA, RENTA, RESIDENTES, PERNOCTACIONES, POTENCIAL, Í.INDUSTRIAL, Í.RESTAUACIÓN, Í.COMERCIAL, POBLACIÓN
- g. Variables predictorias: (Constante), CUOTA, RENTA, PERNOCTACIONES, POTENCIAL, Í.INDUSTRIAL, Í.RESTAUACIÓN, Í.COMERCIAL, POBLACIÓN
- h. Variables predictorias: (Constante), CUOTA, RENTA, PERNOCTACIONES, POTENCIAL, Í.INDUSTRIAL, Í.COMERCIAL, POBLACIÓN
- i. Variables predictorias: (Constante), CUOTA, RENTA, PERNOCTACIONES, Í.INDUSTRIAL, Í.COMERCIAL, POBLACIÓN
- j. Variable dependiente: PASAJEROS



RESUMEN DEL MODELO (MÉTODO HACIA ADELANTE)

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Estadísticos de cambio	
					Cambio en R cuadrado	Cambio en F
1	,807 ^a	,651	,641	,50576918698	,651	68,937
2	,834 ^b	,696	,679	,47834320595	,045	5,364

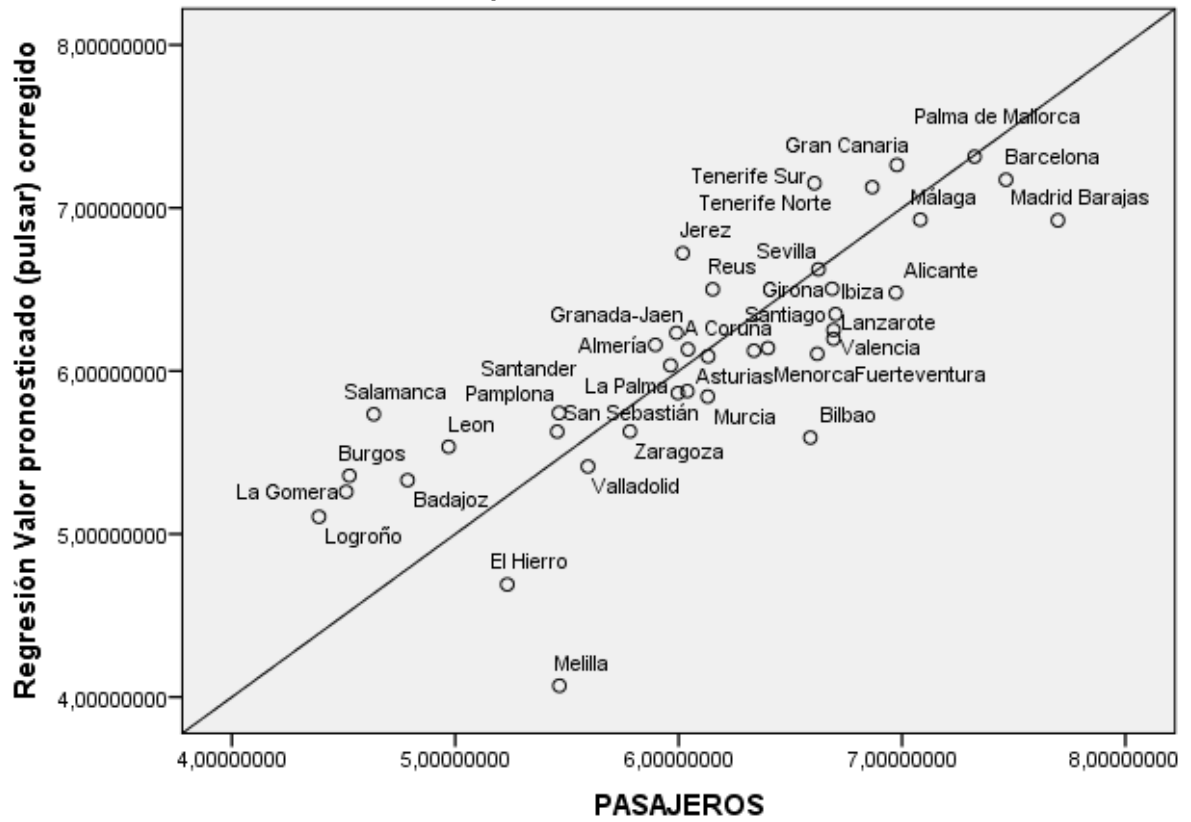
a. Variables predictoras: (Constante), Í.TURÍSTICO

b. Variables predictoras: (Constante), Í.TURÍSTICO, RESIDENTES

c. Variable dependiente: PASAJEROS

Gráfico de dispersión

Variable dependiente: PASAJEROS



5.5 RESULTADOS MODELO 5 AÑO 2009 15 VARIABLES

CORRELACIONES

		PASAJEROS	POBLACIÓN	RESIDENTES
Correlación de Pearson	PASAJEROS	1,000	,505	,620
	POBLACIÓN	,505	1,000	,892
	RESIDENTES	,620	,892	1,000
	RENTA	,099	,206	,198
	PARO	,144	-,246	-,092
	Í.COMERCIAL	,508	,999	,899
	Í.RESTAURACIÓN	,500	,993	,899
	Í.TURÍSTICO	,813	,736	,877
	Í.ECONÓMICO	,477	,988	,886
	Í.INDUSTRIAL	,342	,956	,829
	PERNOCTACIONES	,726	,757	,903
	IMPORTACIONES	,392	,907	,759
	POTENCIAL	,442	,955	,848

		RENTA	PARO	Í.COMERCIAL
Correlación de Pearson	PASAJEROS	,099	,144	,508
	POBLACIÓN	,206	-,246	,999
	RESIDENTES	,198	-,092	,899
	RENTA	1,000	-,801	,220
	PARO	-,801	1,000	-,255
	Í.COMERCIAL	,220	-,255	1,000
	Í.RESTAURACIÓN	,234	-,293	,994
	Í.TURÍSTICO	,075	,115	,745
	Í.ECONÓMICO	,338	-,350	,990
	Í.INDUSTRIAL	,401	-,453	,958
	PERNOCTACIONES	,051	,098	,768
	IMPORTACIONES	,409	-,434	,905
	POTENCIAL	,252	-,286	,956

		Í.RESTAURACIÓN	Í.TURÍSTICO	Í.ECONÓMICO
Correlación de Pearson	PASAJEROS	,500	,813	,477
	POBLACIÓN	,993	,736	,988
	RESIDENTES	,899	,877	,886
	RENTA	,234	,075	,338
	PARO	-,293	,115	-,350
	Í.COMERCIAL	,994	,745	,990
	Í.RESTAURACIÓN	1,000	,743	,988
	Í.TURÍSTICO	,743	1,000	,716
	Í.ECONÓMICO	,988	,716	1,000
	Í.INDUSTRIAL	,959	,611	,982
	PERNOCTACIONES	,762	,957	,733
	IMPORTACIONES	,903	,571	,929
	POTENCIAL	,958	,666	,951

		Í.INDUSTRIAL	PERNOCTACIONES	IMPORTACIONES
Correlación de Pearson	PASAJEROS	,342	,726	,392
	POBLACIÓN	,956	,757	,907
	RESIDENTES	,829	,903	,759
	RENTA	,401	,051	,409
	PARO	-,453	,098	-,434
	Í.COMERCIAL	,958	,768	,905
	Í.RESTAURACIÓN	,959	,762	,903
	Í.TURÍSTICO	,611	,957	,571
	Í.ECONÓMICO	,982	,733	,929
	Í.INDUSTRIAL	1,000	,653	,937
	PERNOCTACIONES	,653	1,000	,581
	IMPORTACIONES	,937	,581	1,000
	POTENCIAL	,924	,694	,877

		POTENCIAL
Correlación de Pearson	PASAJEROS	,442
	POBLACIÓN	,955
	RESIDENTES	,848
	RENTA	,252
	PARO	-,286
	Í.COMERCIAL	,956
	Í.RESTAURACIÓN	,958
	Í.TURÍSTICO	,666
	Í.ECONÓMICO	,951
	Í.INDUSTRIAL	,924
	PERNOCTACIONES	,694
	IMPORTACIONES	,877
	POTENCIAL	1,000

COEFICIENTES

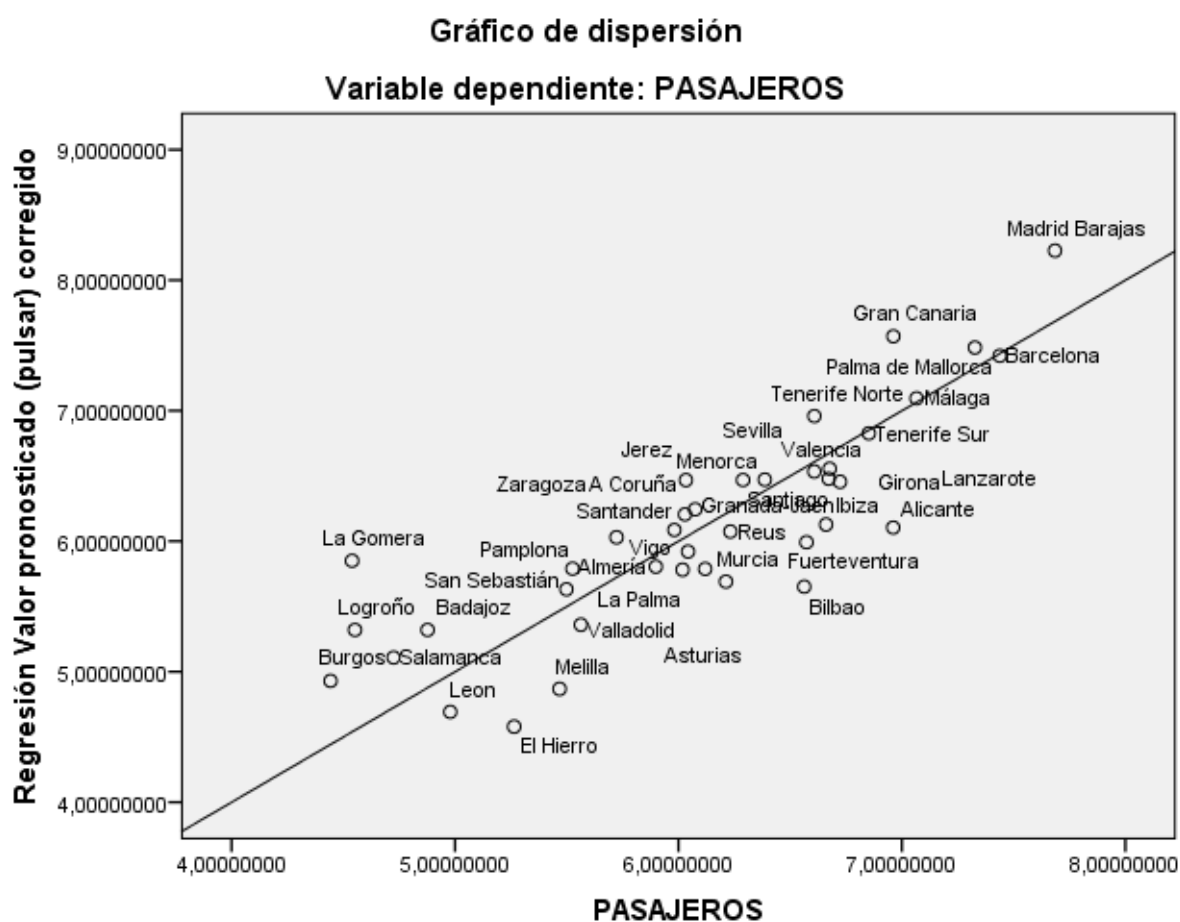
Modelo	Coeficientes no estandarizados		Coeficientes tipificados	t
	B	Error típ.	Beta	
(Constante)	-53,445	20,587		-2,596
POBLACIÓN	7,999	4,384	5,621	1,825
RESIDENTES	-,442	,400	-,313	-1,106
RENTA	7,217	3,044	,652	2,371
PARO	,380	1,205	,072	,315
Í.COMERCIAL	-8,694	4,820	-6,051	-1,804
Í.RESTAUACIÓN	1,157	1,387	,838	,834
Í.TURÍSTICO	,606	,557	,402	1,087
Í.ECONÓMICO	-,129	2,389	-,098	-,054
Í.INDUSTRIAL	-2,659	,994	-2,197	-2,675
PERNOCTACIONES	,848	,586	,534	1,448
IMPORTACIONES	-,020	,311	-,018	-,064
POTENCIAL	-,453	,427	-,317	-1,062

RESUMEN DEL MODELO (MÉTODO INTRODUCIR)

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
1	,938 ^a	,879	,808	,36047043521	2,207

a. Variables predictoras: (Constante), CUOTA, RENTA, Í.TURÍSTICO, PARO, IMPORTACIONES, RESIDENTES, POTENCIAL, EXPORTACIONES, PERNOCTACIONES, Í.INDUSTRIAL, Í.RESTAUACIÓN, Í.COMERCIAL, Í.ECONÓMICO, POBLACIÓN

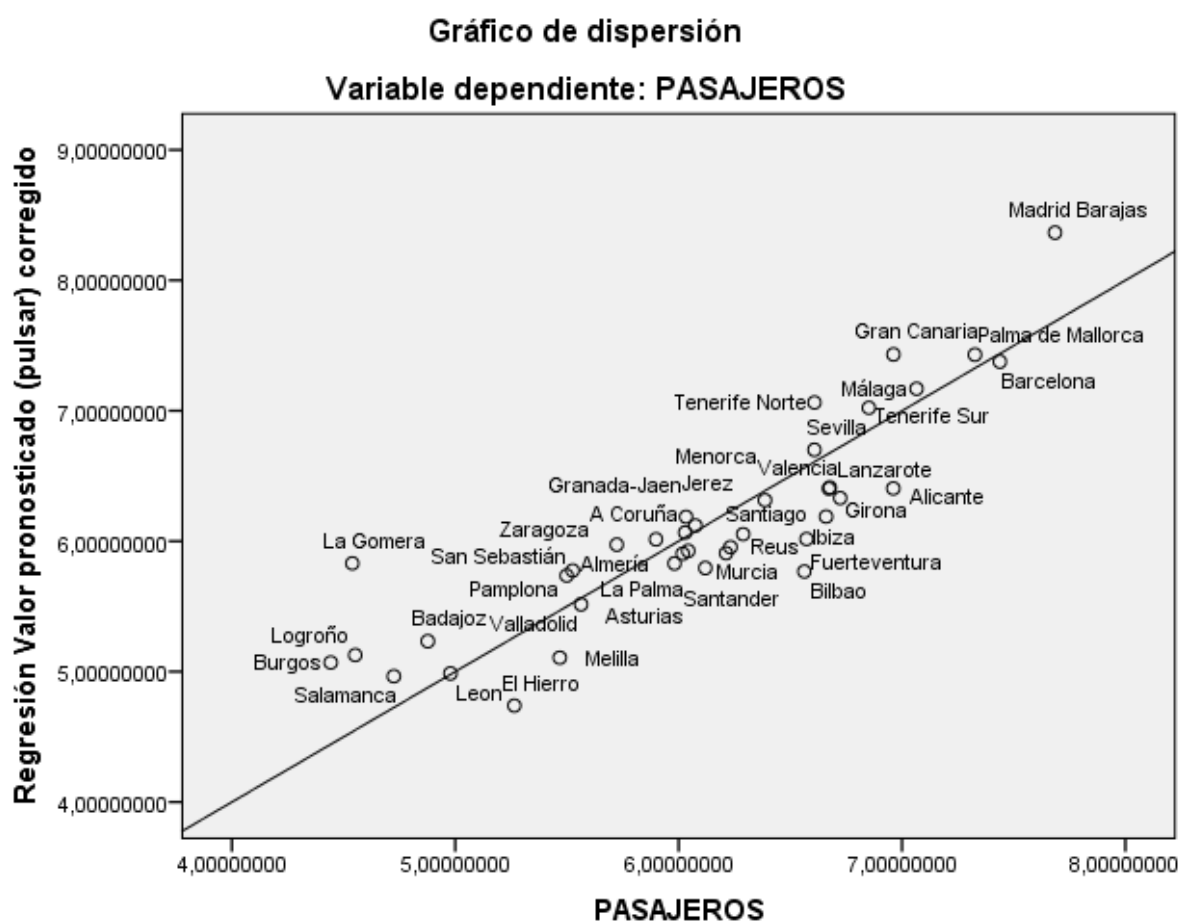
b. Variable dependiente: PASAJEROS



RESUMEN DEL MODELO (MÉTODO HACIA ATRÁS)

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Estadísticos de cambio	
					Cambio en R cuadrado	Cambio en F
1	,938 ^a	,879	,808	,36047043521	,879	12,450
2	,938 ^b	,879	,816	,35320901460	,000	,003
3	,938 ^c	,879	,823	,34637012362	,000	,003
4	,937 ^d	,878	,829	,34063188435	-,001	,113
5	,937 ^e	,877	,833	,33635565499	-,001	,301
6	,934 ^f	,873	,833	,33604863590	-,004	,947
7	,932 ^g	,869	,834	,33583028188	-,004	,961
8	,929 ^h	,864	,833	,33679089686	-,005	1,178
9	,926 ⁱ	,858	,832	,33794894947	-,005	1,220

- a. Variables predictoras: (Constante), CUOTA, RENTA, Í.TURÍSTICO, PARO, IMPORTACIONES, RESIDENTES, POTENCIAL, EXPORTACIONES, PERNOCTACIONES, Í.INDUSTRIAL, Í.RESTAUACIÓN, Í.COMERCIAL, Í.ECONÓMICO, POBLACIÓN
- b. Variables predictoras: (Constante), CUOTA, RENTA, Í.TURÍSTICO, PARO, IMPORTACIONES, RESIDENTES, POTENCIAL, EXPORTACIONES, PERNOCTACIONES, Í.INDUSTRIAL, Í.RESTAUACIÓN, Í.COMERCIAL, POBLACIÓN
- c. Variables predictoras: (Constante), CUOTA, RENTA, Í.TURÍSTICO, PARO, RESIDENTES, POTENCIAL, EXPORTACIONES, PERNOCTACIONES, Í.INDUSTRIAL, Í.RESTAUACIÓN, Í.COMERCIAL, POBLACIÓN
- d. Variables predictoras: (Constante), CUOTA, RENTA, Í.TURÍSTICO, RESIDENTES, POTENCIAL, EXPORTACIONES, PERNOCTACIONES, Í.INDUSTRIAL, Í.RESTAUACIÓN, Í.COMERCIAL, POBLACIÓN
- e. Variables predictoras: (Constante), CUOTA, RENTA, Í.TURÍSTICO, RESIDENTES, POTENCIAL, PERNOCTACIONES, Í.INDUSTRIAL, Í.RESTAUACIÓN, Í.COMERCIAL, POBLACIÓN
- f. Variables predictoras: (Constante), CUOTA, RENTA, Í.TURÍSTICO, RESIDENTES, POTENCIAL, PERNOCTACIONES, Í.INDUSTRIAL, Í.COMERCIAL, POBLACIÓN
- g. Variables predictoras: (Constante), CUOTA, RENTA, Í.TURÍSTICO, POTENCIAL, PERNOCTACIONES, Í.INDUSTRIAL, Í.COMERCIAL, POBLACIÓN
- h. Variables predictoras: (Constante), CUOTA, RENTA, Í.TURÍSTICO, POTENCIAL, Í.INDUSTRIAL, Í.COMERCIAL, POBLACIÓN
- i. Variables predictoras: (Constante), CUOTA, RENTA, Í.TURÍSTICO, Í.INDUSTRIAL, Í.COMERCIAL, POBLACIÓN
- j. Variable dependiente: PASAJEROS



RESUMEN DEL MODELO (MÉTODO HACIA ADELANTE)

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Estadísticos de cambio	
					Cambio en R cuadrado	Cambio en F
1	,813 ^a	,661	,652	,48586257978	,661	72,155
2	,836 ^b	,699	,683	,46397752971	,038	4,573
3	,873 ^c	,761	,741	,41920053015	,062	9,101
4	,897 ^d	,804	,781	,38538193773	,043	7,412
5	,917 ^e	,841	,817	,35206607996	,037	7,739

a. Variables predictoras: (Constante), Í.TURÍSTICO

b. Variables predictoras: (Constante), Í.TURÍSTICO, Í.INDUSTRIAL

c. Variables predictoras: (Constante), Í.TURÍSTICO, Í.INDUSTRIAL, Í.ECONÓMICO

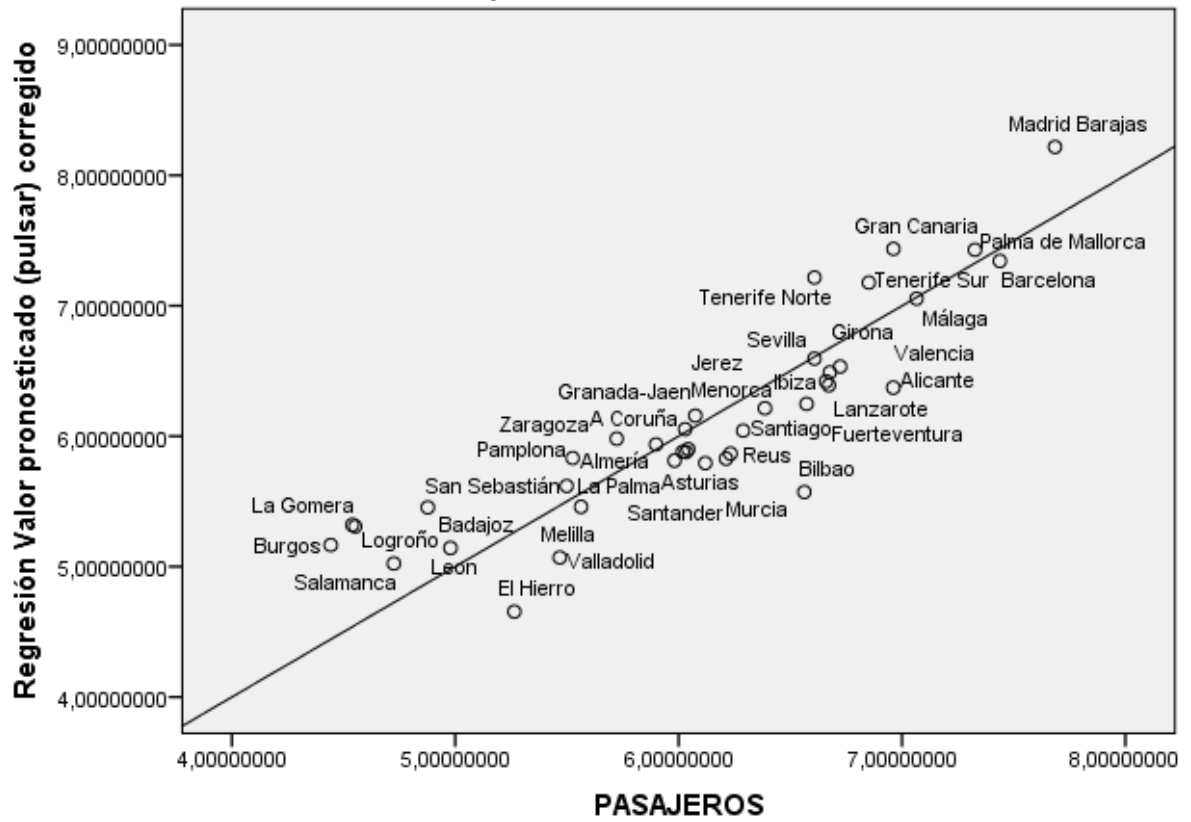
d. Variables predictoras: (Constante), Í.TURÍSTICO, Í.INDUSTRIAL, Í.ECONÓMICO, RENTA

e. Variables predictoras: (Constante), Í.TURÍSTICO, Í.INDUSTRIAL, Í.ECONÓMICO, RENTA, Í.COMERCIAL

f. Variable dependiente: PASAJEROS

Gráfico de dispersión

Variable dependiente: PASAJEROS



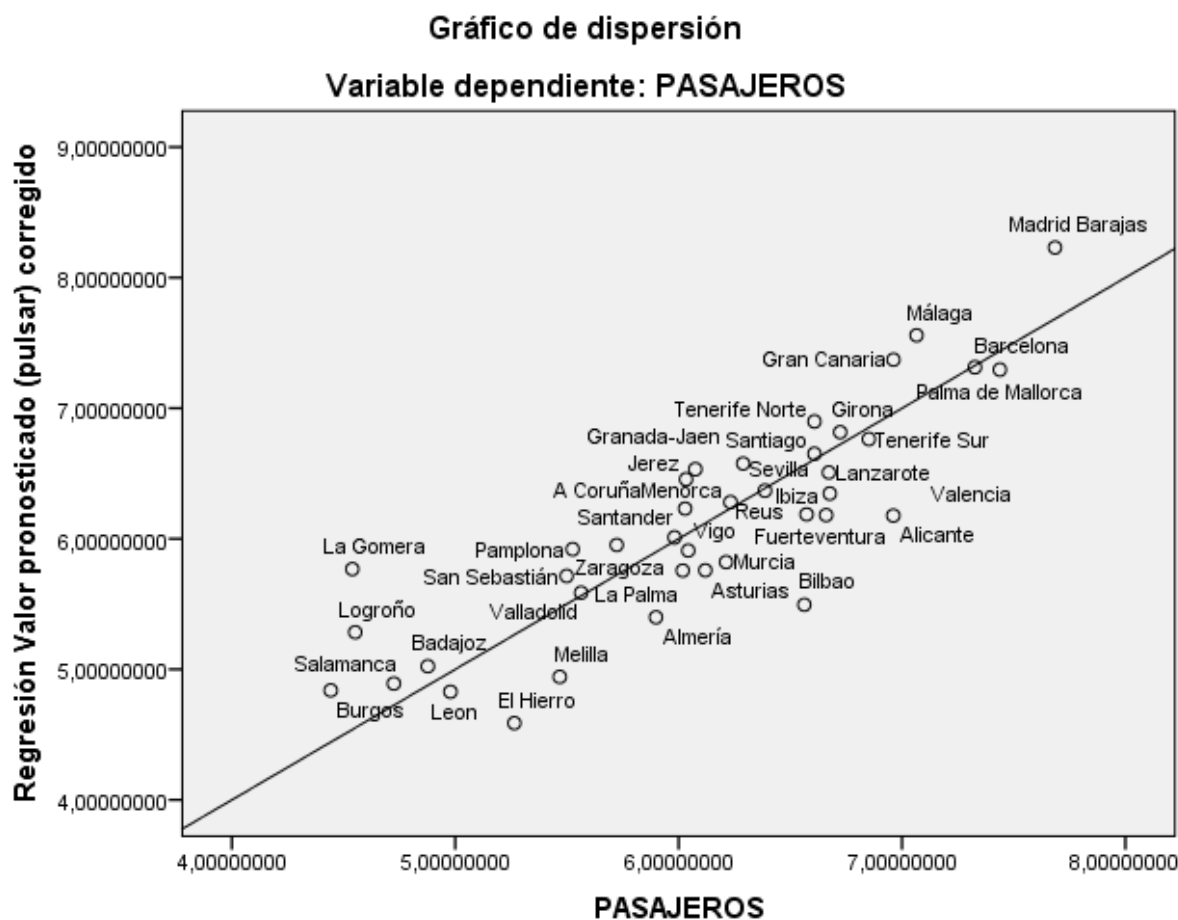
RESULTADOS DEL MODELO 5 AÑADIENDO LA **VARIABLE “POBLACIÓN EN LOS ENTORNOS** **AEROPORTUARIOS POR DISTANCIAS DISUASORIAS”**

RESUMEN DEL MODELO (MÉTODO INTRODUCIR)

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
1	,946 ^a	,894	,825	,34431389376	2,018

a. Variables predictoras: (Constante), ÁREA.DISUASORIA, PARO, Í.TURÍSTICO, RENTA, POTENCIAL, IMPORTACIONES, RESIDENTES, EXPORTACIONES, PERNOCTACIONES, Í.COMERCIAL, Í.INDUSTRIAL, Í.RESTAURACIÓN, POBLACIÓN, Í.ECONÓMICO, CUOTA

b. Variable dependiente: PASAJEROS



5.6 RESULTADOS

MODELO 6

AÑO 2013

23 VARIABLES

CORRELACIONES

		PASAJEROS	POBLACIÓN	RESIDENTES
Correlación de Pearson	PASAJEROS	1,000	,459	,627
	POBLACIÓN	,459	1,000	,828
	RESIDENTES	,627	,828	1,000
	RENTA	,057	,181	,138
	PARO	,146	-,214	-,008
	Í.COMERCIAL	,500	,358	,452
	Í.RESTAUACIÓN	,445	,992	,830
	Í.TURÍSTICO	,785	,734	,907
	Í.ECONÓMICO	,400	,982	,807
	Í.INDUSTRIAL	,259	,946	,725
	PERNOCTACIONES	,672	,778	,922
	IMPORTACIONES	,321	,886	,667
	POTENCIAL	,294	,933	,715
	KMS.DESTINOS	-,641	-,390	-,662
	RENTA.PAÍSES	,584	,253	,220
	LOW.COST.CINCO	,870	,391	,541
	LOW.COST.TOTAL	,816	,272	,575

		RENTA	PARO	Í.COMERCIAL
Correlación de Pearson	PASAJEROS	,057	,146	,500
	POBLACIÓN	,181	-,214	,358
	RESIDENTES	,138	-,008	,452
	RENTA	1,000	-,858	,255
	PARO	-,858	1,000	-,084
	Í.COMERCIAL	,255	-,084	1,000
	Í.RESTAUACIÓN	,228	-,274	,362
	Í.TURÍSTICO	,098	,078	,522
	Í.ECONÓMICO	,331	-,355	,354
	Í.INDUSTRIAL	,386	-,445	,254
	PERNOCTACIONES	,093	,041	,435
	IMPORTACIONES	,383	-,368	,290
	POTENCIAL	,140	-,184	,349
	KMS.DESTINOS	,028	-,202	-,732
	RENTA.PAÍSES	-,001	,048	,163
	LOW.COST.CINCO	-,040	,156	,412
	LOW.COST.TOTAL	-,069	,242	,520

		Í.RESTAUACIÓN	Í.TURÍSTICO	Í.ECONÓMICO
Correlación de Pearson	PASAJEROS	,445	,785	,400
	POBLACIÓN	,992	,734	,982
	RESIDENTES	,830	,907	,807
	RENTA	,228	,098	,331
	PARO	-,274	,078	-,355
	Í.COMERCIAL	,362	,522	,354
	Í.RESTAUACIÓN	1,000	,735	,987
	Í.TURÍSTICO	,735	1,000	,697
	Í.ECONÓMICO	,987	,697	1,000
	Í.INDUSTRIAL	,956	,585	,982
	PERNOCTACIONES	,777	,949	,745
	IMPORTACIONES	,878	,556	,908
	POTENCIAL	,919	,572	,919
	KMS.DESTINOS	-,391	-,662	-,352
	RENTA.PAÍSES	,252	,417	,221
	LOW.COST.CINCO	,396	,698	,346
	LOW.COST.TOTAL	,278	,710	,229

		Í.INDUSTRIAL	PERNOCTACIONES	IMPORTACIONES
Correlación de Pearson	PASAJEROS	,259	,672	,321
	POBLACIÓN	,946	,778	,886
	RESIDENTES	,725	,922	,667
	RENTA	,386	,093	,383
	PARO	-,445	,041	-,368
	Í.COMERCIAL	,254	,435	,290
	Í.RESTAUACIÓN	,956	,777	,878
	Í.TURÍSTICO	,585	,949	,556
	Í.ECONÓMICO	,982	,745	,908
	Í.INDUSTRIAL	1,000	,665	,916
	PERNOCTACIONES	,665	1,000	,586
	IMPORTACIONES	,916	,586	1,000
	POTENCIAL	,900	,625	,860
	KMS.DESTINOS	-,236	-,627	-,198
	RENTA.PAÍSES	,172	,358	,217
	LOW.COST.CINCO	,217	,587	,223
	LOW.COST.TOTAL	,097	,635	,077

		POTENCIAL	KMS.DESTINOS	RENTA.PAÍSES
Correlación de Pearson	PASAJEROS	,294	-,641	,584
	POBLACIÓN	,933	-,390	,253
	RESIDENTES	,715	-,662	,220
	RENTA	,140	,028	-,001
	PARO	-,184	-,202	,048
	Í.COMERCIAL	,349	-,732	,163
	Í.RESTAURACIÓN	,919	-,391	,252
	Í.TURÍSTICO	,572	-,662	,417
	Í.ECONÓMICO	,919	-,352	,221
	Í.INDUSTRIAL	,900	-,236	,172
	PERNOCTACIONES	,625	-,627	,358
	IMPORTACIONES	,860	-,198	,217
	POTENCIAL	1,000	-,319	,077
	KMS.DESTINOS	-,319	1,000	-,163
	RENTA.PAÍSES	,077	-,163	1,000
	LOW.COST.CINCO	,211	-,552	,642
	LOW.COST.TOTAL	,111	-,756	,471

		LOW.COST.CINCO	LOW.COST.TOTAL
Correlación de Pearson	PASAJEROS	,870	,816
	POBLACIÓN	,391	,272
	RESIDENTES	,541	,575
	RENTA	-,040	-,069
	PARO	,156	,242
	Í.COMERCIAL	,412	,520
	Í.RESTAURACIÓN	,396	,278
	Í.TURÍSTICO	,698	,710
	Í.ECONÓMICO	,346	,229
	Í.INDUSTRIAL	,217	,097
	PERNOCTACIONES	,587	,635
	IMPORTACIONES	,223	,077
	POTENCIAL	,211	,111
	KMS.DESTINOS	-,552	-,756
	RENTA.PAÍSES	,642	,471
	LOW.COST.CINCO	1,000	,847
	LOW.COST.TOTAL	,847	1,000

COEFICIENTES

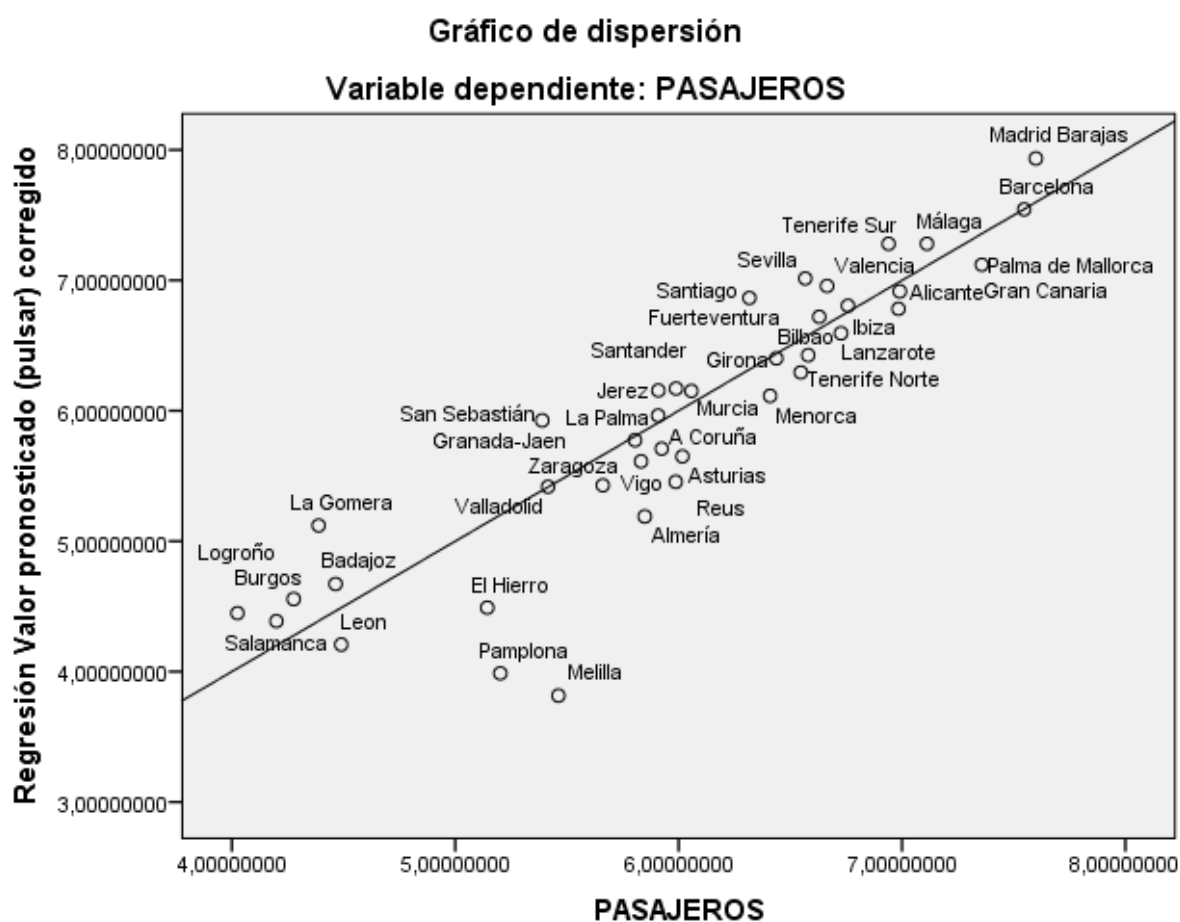
Modelo	Coeficientes no estandarizados		Coeficientes tipificados	t
	B	Error típ.	Beta	
(Constante)	-13,875	11,212		-1,238
POBLACIÓN	1,751	1,615	1,065	1,084
RESIDENTES	-,251	,297	-,150	-,843
RENTA	3,447	1,992	,304	1,731
PARO	-,803	1,293	-,099	-,621
Í.COMERCIAL	-,851	,268	-,276	-3,174
Í.RESTAURACIÓN	1,748	1,070	1,103	1,634
Í.TURÍSTICO	,108	,419	,063	,259
Í.ECONÓMICO	1,260	1,698	,839	,742
Í.INDUSTRIAL	-2,229	,996	-1,651	-2,238
PERNOCTACIONES	-,147	,426	-,080	-,344
IMPORTACIONES	,211	,260	,171	,813
POTENCIAL	,436	,320	,308	1,361
KMS.DESTINOS	-,310	,107	-,316	-2,887
RENTA.PAÍSES	,323	,802	,031	,403
LOW.COST.CINCO	,171	,071	,357	2,410
LOW.COST.TOTAL	,002	,025	,015	,096

RESUMEN DEL MODELO (MÉTODO INTRODUCIR)

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
1	,986 ^a	,973	,940	,23313992209	1,606

a. Variables predictoras: (Constante), HUBS, PARO, RENTA.PAÍSES, POTENCIAL, LOW.COST.TOTAL, DESTINOS.CERCA, PERNOCTACIONES, DESTINOS, KMS.DESTINOS, RENTA, LOW.COST.CINCO, IMPORTACIONES, RESIDENTES, COMPAÑÍAS, Í.TURÍSTICO, EXPORTACIONES, Í.RESTAURACIÓN, Í.INDUSTRIAL, Í.COMERCIAL, POBLACIÓN, Í.ECONÓMICO

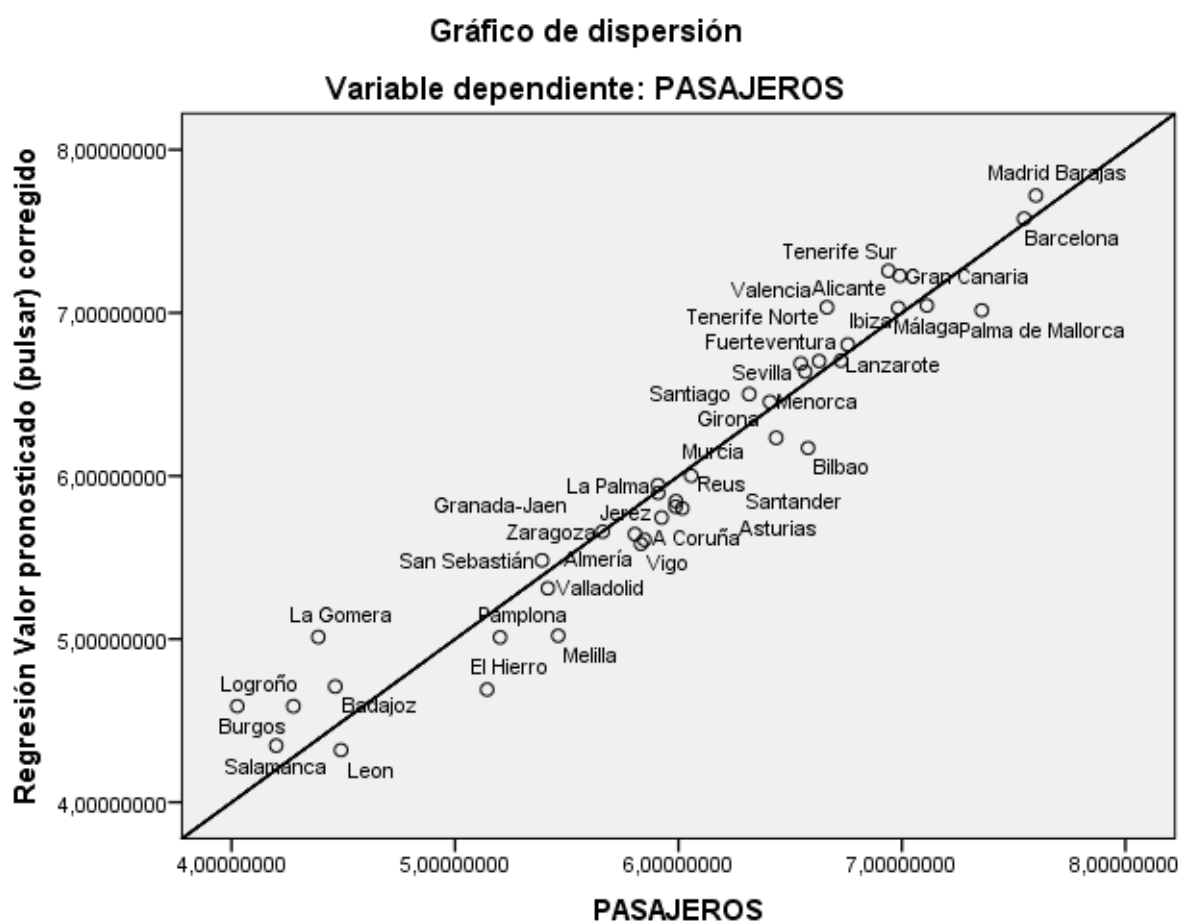
b. Variable dependiente: PASAJEROS



RESUMEN DEL MODELO (MÉTODO HACIA ATRÁS)

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Estadísticos de cambio	
					Cambio en R cuadrado	Cambio en F
1	,986 ^a	,973	,940	,23313992209	,973	29,122
2	,986 ^b	,973	,943	,22663306696	,000	,009
3	,986 ^c	,973	,946	,22095355238	,000	,060
4	,986 ^d	,973	,948	,21599199874	,000	,112
5	,986 ^e	,973	,950	,21129872488	,000	,097
6	,986 ^f	,972	,952	,20782702269	,000	,283
7	,986 ^g	,971	,953	,20632587952	-,001	,669
8	,985 ^h	,970	,953	,20559010797	-,001	,829
9	,984 ⁱ	,969	,953	,20565308888	-,001	1,015
10	,984 ^j	,967	,952	,20711602355	-,002	1,371
11	,982 ^k	,964	,949	,21365780897	-,003	2,733
12	,981 ^l	,963	,950	,21187357132	-,001	,534
13	,980 ^m	,960	,947	,21787815445	-,004	2,667

- a. Variables predictorias: (Constante), HUBS, PARO, RENTA.PAÍSES, POTENCIAL, LOW.COST.TOTAL, DESTINOS.CERCA, PERNOCTACIONES, DESTINOS, KMS.DESTINOS, RENTA, LOW.COST.CINCO, IMPORTACIONES, RESIDENTES, COMPAÑÍAS, Í.TURÍSTICO, EXPORTACIONES, Í.RESTauración, Í.INDUSTRIAL, Í.COMERCIAL, POBLACIÓN, Í.ECONÓMICO
- b. Variables predictorias: (Constante), HUBS, PARO, RENTA.PAÍSES, POTENCIAL, DESTINOS.CERCA, PERNOCTACIONES, DESTINOS, KMS.DESTINOS, RENTA, LOW.COST.CINCO, IMPORTACIONES, RESIDENTES, COMPAÑÍAS, Í.TURÍSTICO, EXPORTACIONES, Í.RESTauración, Í.INDUSTRIAL, Í.COMERCIAL, POBLACIÓN, Í.ECONÓMICO
- c. Variables predictorias: (Constante), HUBS, PARO, RENTA.PAÍSES, POTENCIAL, DESTINOS.CERCA, PERNOCTACIONES, DESTINOS, KMS.DESTINOS, RENTA, LOW.COST.CINCO, IMPORTACIONES, RESIDENTES, Í.TURÍSTICO, EXPORTACIONES, Í.RESTauración, Í.INDUSTRIAL, Í.COMERCIAL, POBLACIÓN, Í.ECONÓMICO
- d. Variables predictorias: (Constante), HUBS, PARO, RENTA.PAÍSES, POTENCIAL, DESTINOS.CERCA, PERNOCTACIONES, DESTINOS, KMS.DESTINOS, RENTA, LOW.COST.CINCO, IMPORTACIONES, RESIDENTES, EXPORTACIONES, Í.RESTauración, Í.INDUSTRIAL, Í.COMERCIAL, POBLACIÓN, Í.ECONÓMICO
- e. Variables predictorias: (Constante), HUBS, PARO, RENTA.PAÍSES, POTENCIAL, DESTINOS.CERCA, DESTINOS, KMS.DESTINOS, RENTA, LOW.COST.CINCO, IMPORTACIONES, RESIDENTES, EXPORTACIONES, Í.RESTauración, Í.INDUSTRIAL, Í.COMERCIAL, POBLACIÓN, Í.ECONÓMICO
- f. Variables predictorias: (Constante), HUBS, PARO, POTENCIAL, DESTINOS.CERCA, DESTINOS, KMS.DESTINOS, RENTA, LOW.COST.CINCO, IMPORTACIONES, RESIDENTES, EXPORTACIONES, Í.RESTauración, Í.INDUSTRIAL, Í.COMERCIAL, POBLACIÓN, Í.ECONÓMICO
- g. Variables predictorias: (Constante), HUBS, PARO, POTENCIAL, DESTINOS.CERCA, DESTINOS, KMS.DESTINOS, RENTA, LOW.COST.CINCO, IMPORTACIONES, RESIDENTES, Í.RESTauración, Í.INDUSTRIAL, Í.COMERCIAL, POBLACIÓN, Í.ECONÓMICO
- h. Variables predictorias: (Constante), HUBS, POTENCIAL, DESTINOS.CERCA, DESTINOS, KMS.DESTINOS, RENTA, LOW.COST.CINCO, IMPORTACIONES, RESIDENTES, Í.RESTauración, Í.INDUSTRIAL, Í.COMERCIAL, POBLACIÓN, Í.ECONÓMICO
- i. Variables predictorias: (Constante), HUBS, POTENCIAL, DESTINOS.CERCA, DESTINOS, KMS.DESTINOS, RENTA, LOW.COST.CINCO, IMPORTACIONES, RESIDENTES, Í.RESTauración, Í.INDUSTRIAL, Í.COMERCIAL, POBLACIÓN
- j. Variables predictorias: (Constante), HUBS, DESTINOS.CERCA, DESTINOS, KMS.DESTINOS, RENTA, LOW.COST.CINCO, IMPORTACIONES, RESIDENTES, Í.RESTauración, Í.INDUSTRIAL, Í.COMERCIAL, POBLACIÓN
- k. Variables predictorias: (Constante), HUBS, DESTINOS.CERCA, DESTINOS, KMS.DESTINOS, RENTA, LOW.COST.CINCO, IMPORTACIONES, RESIDENTES, Í.RESTauración, Í.INDUSTRIAL, Í.COMERCIAL
- l. Variables predictorias: (Constante), HUBS, DESTINOS.CERCA, DESTINOS, KMS.DESTINOS, RENTA, LOW.COST.CINCO, IMPORTACIONES, RESIDENTES, Í.RESTauración, Í.INDUSTRIAL
- m. Variables predictorias: (Constante), HUBS, DESTINOS.CERCA, DESTINOS, KMS.DESTINOS, RENTA, LOW.COST.CINCO, IMPORTACIONES, Í.RESTauración, Í.INDUSTRIAL
- n. Variable dependiente: PASAJEROS



RESUMEN DEL MODELO (MÉTODO HACIA ADELANTE)

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Estadísticos de cambio	
					Cambio en R cuadrado	Cambio en F
1	,915 ^a	,836	,832	,38874707575	,836	189,074
2	,939 ^b	,881	,874	,33601541651	,045	13,524

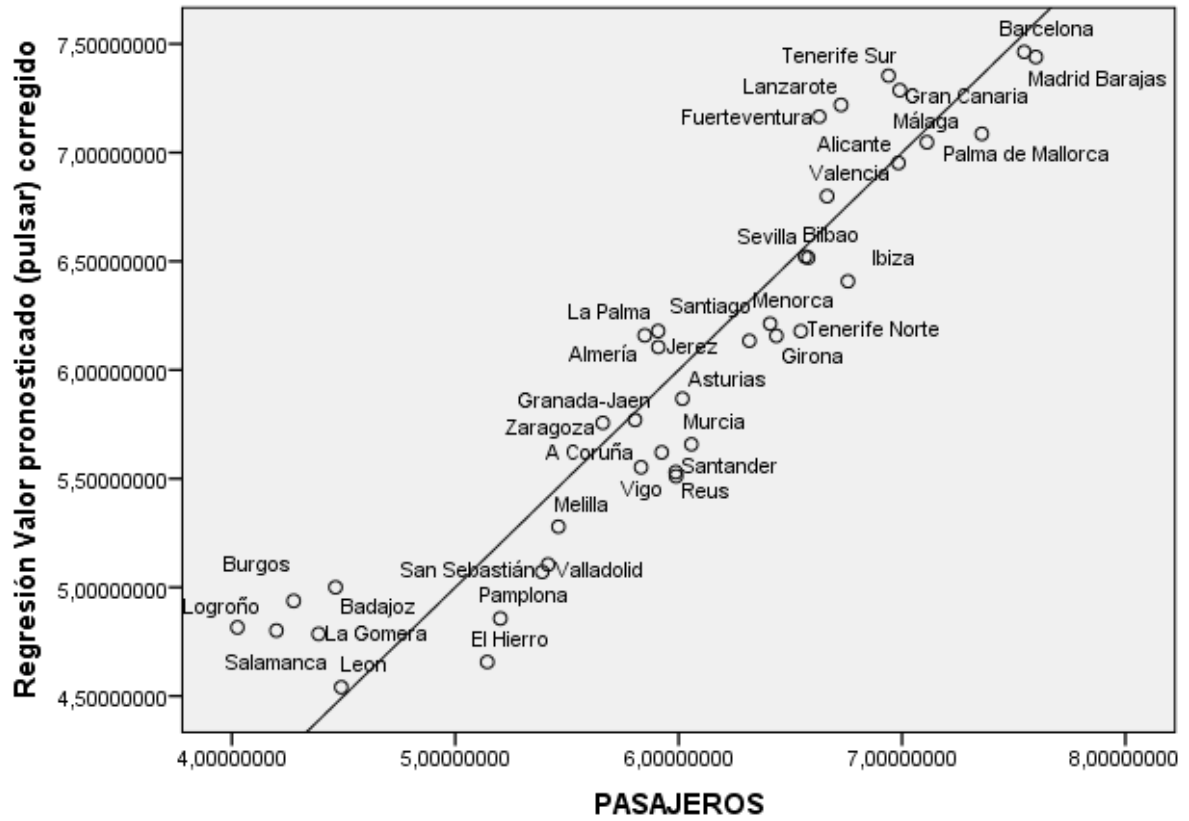
a. Variables predictoras: (Constante), DESTINOS

b. Variables predictoras: (Constante), DESTINOS, COMPAÑÍAS

c. Variable dependiente: PASAJEROS

Gráfico de dispersión

Variable dependiente: PASAJEROS



5.7 RESULTADOS MODELO DE ÚNICA VARIABLE AÑO 2013

POBLACIÓN

Resumen del modelo

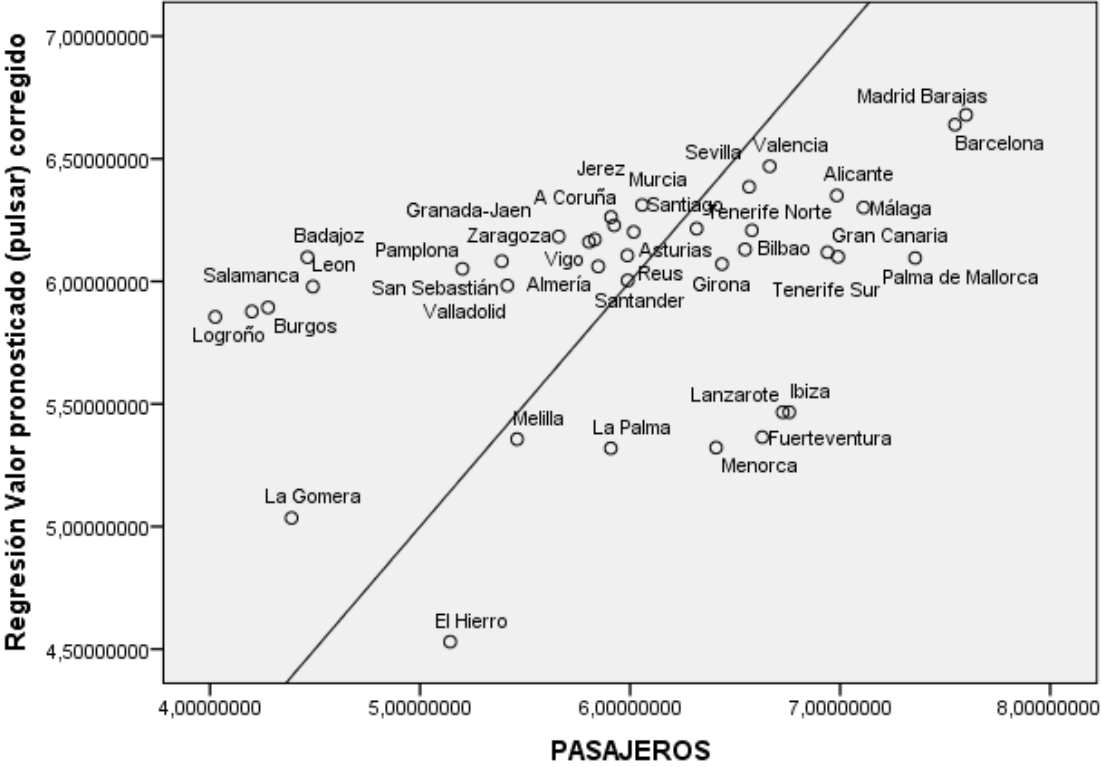
Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
2	,459	,211	,189	,85379657526	2,044

Coefficientes

Modelo	Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
	B	Error típ.	Beta		
2	(Constante)	1,653	1,387	1,192	,241
	POBLACIÓN	,754	,240	,459	,003

Gráfico de dispersión

Variable dependiente: PASAJEROS



POBLACIÓN RESIDENTE DE LA U.E.

Resumen del modelo

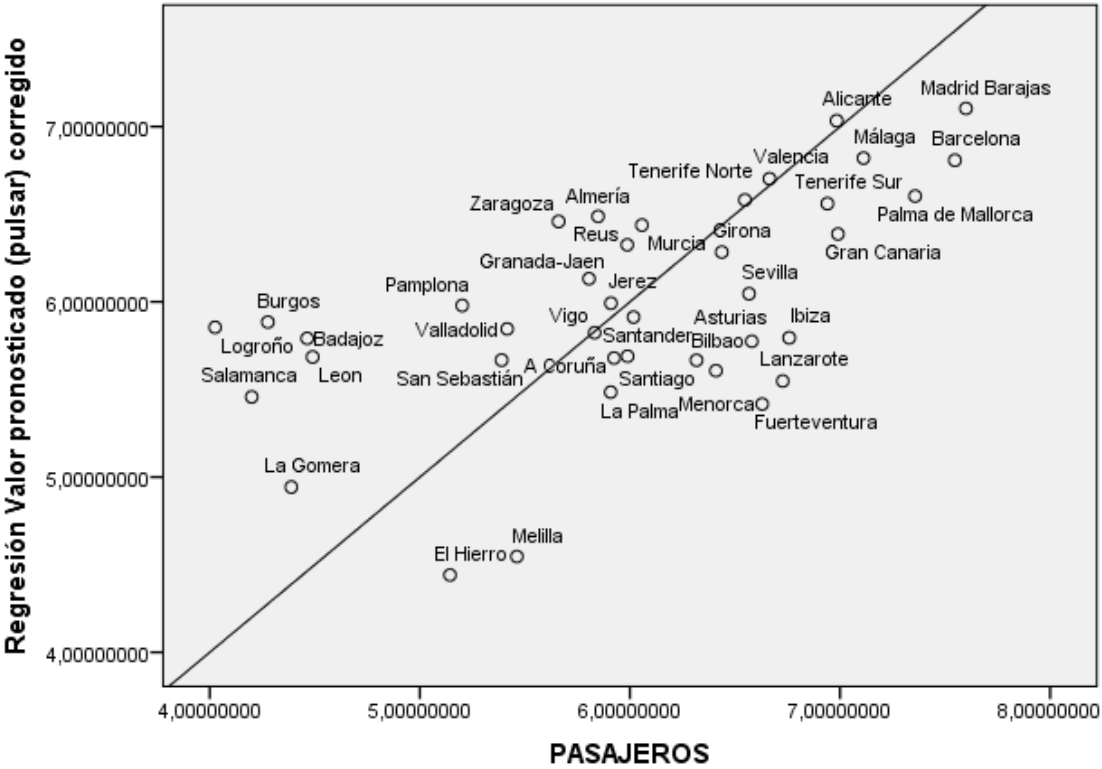
Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
3	,627	,393	,377	,74836907340	2,047

Coefficientes

Modelo	Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
	B	Error típ.	Beta		
3 (Constante)	1,448	,935		1,549	,130
3 RESIDENTES	1,046	,214	,627	4,899	,000

Gráfico de dispersión

Variable dependiente: PASAJEROS



RENTA PER CÁPITA

Resumen del modelo

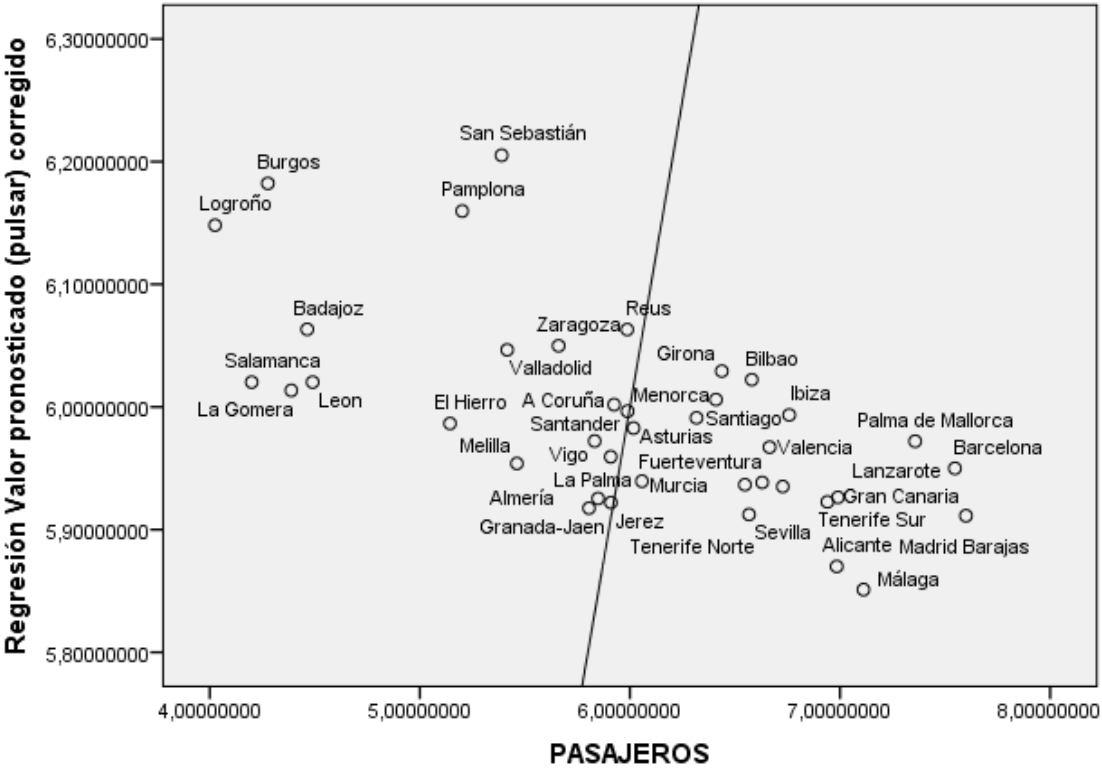
Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
4	,057	,003	-,024	,95938190462	1,949

Coefficientes

Modelo		Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
		B	Error típ.	Beta		
4	(Constante)	3,210	8,039		,399	,692
	RENTA	,643	1,860	,057	,346	,732

Gráfico de dispersión

Variable dependiente: PASAJEROS



TASA DE PARO

Resumen del modelo

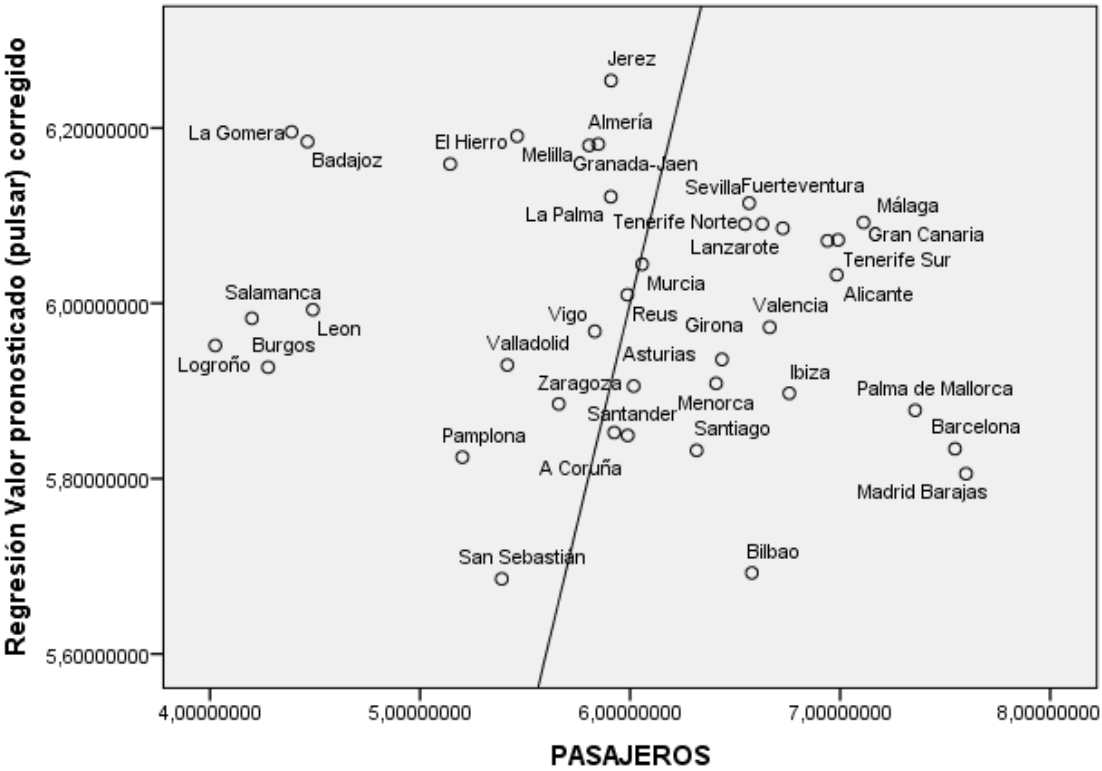
Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
5	,146	,021	-,005	,95063703183	2,084

Coefficientes

Modelo		Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
		B	Error típ.	Beta		
5	(Constante)	4,318	1,867		2,313	,026
	PARO	1,181	1,316	,146	,898	,375

Gráfico de dispersión

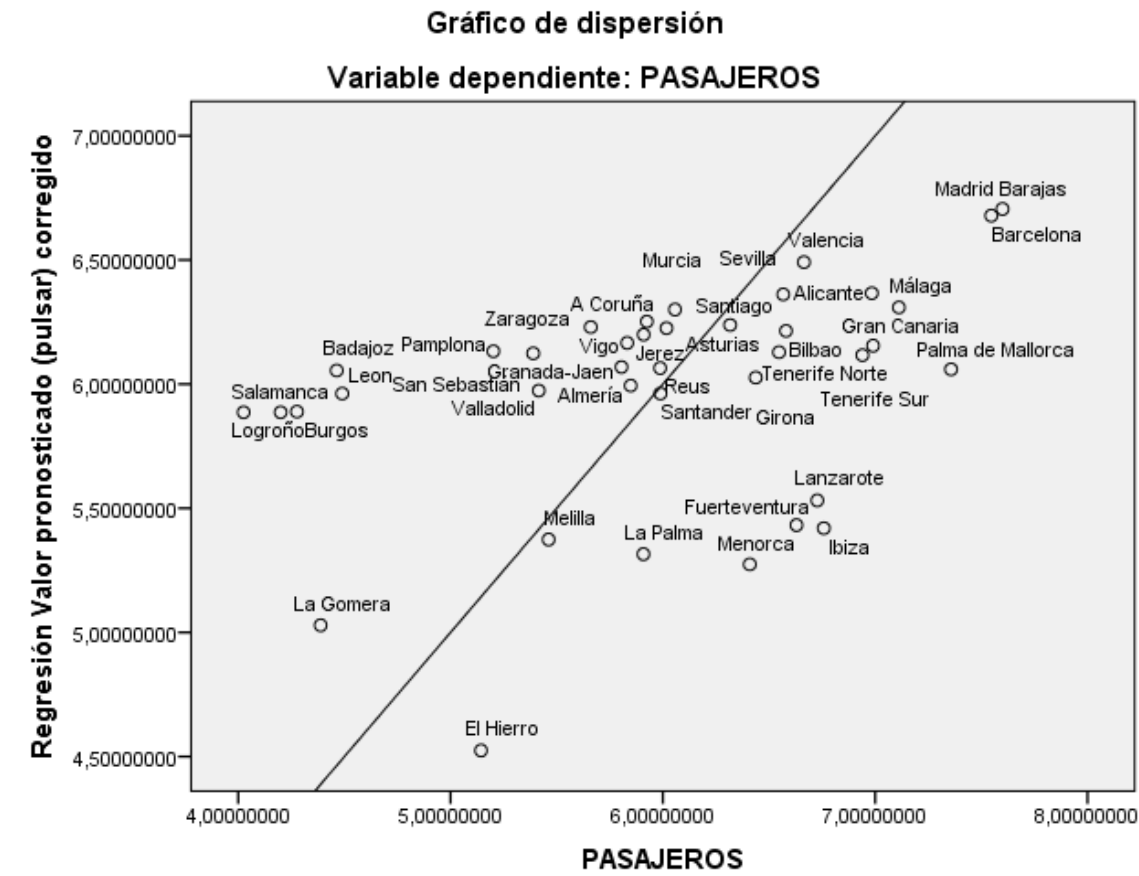
Variable dependiente: PASAJEROS



ÍNDICE COMERCIAL

Resumen del modelo					
Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
6	,462	,214	,193	,85204930440	2,020

Coeficientes					
Modelo	Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
	B	Error típ.	Beta		
6	(Constante)	3,660	,747	4,902	,000
	Í.COMERCIAL	,756	,238	,462	,003



ÍNDICE DE RESTAURACIÓN Y BARES

Resumen del modelo

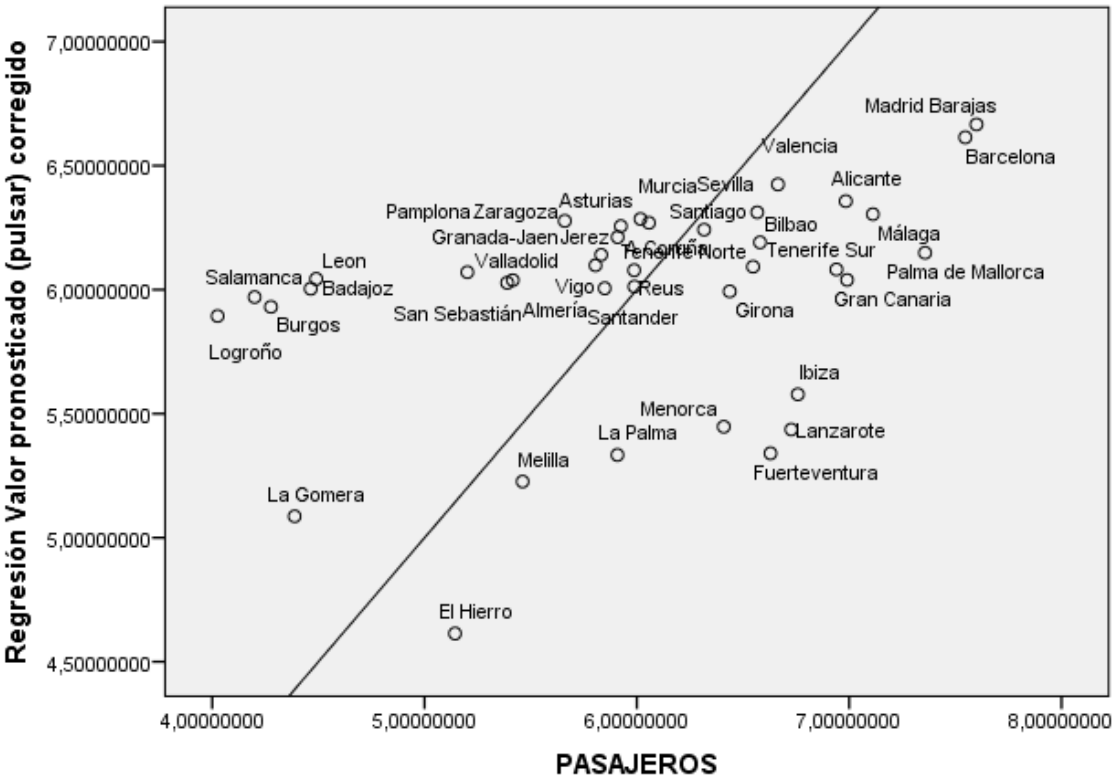
Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
7	,445	,198	,176	,86075792914	1,972

Coefficientes

Modelo		Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
		B	Error típ.	Beta		
7	(Constante)	3,822	,731		5,229	,000
	Í.RESTAURACIÓN	,704	,233	,445	3,019	,005

Gráfico de dispersión

Variable dependiente: PASAJEROS



ÍNDICE TURÍSTICO

Resumen del modelo

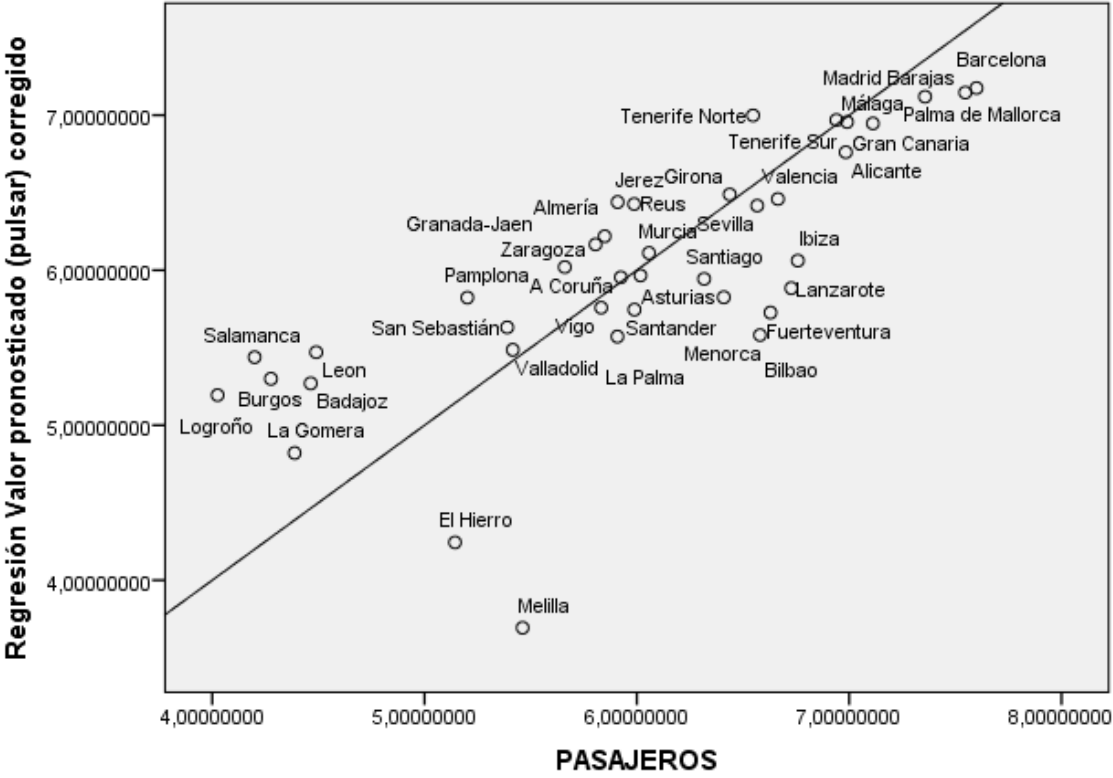
Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
8	,785	,617	,607	,59470928732	1,931

Coefficientes

Modelo		Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
		B	Error típ.	Beta		
8	(Constante)	1,794	,552		3,251	,002
	Í.TURÍSTICO	1,344	,174	,785	7,720	,000

Gráfico de dispersión

Variable dependiente: PASAJEROS



ÍNDICE DE ACTIVIDAD ECONÓMICA

Resumen del modelo

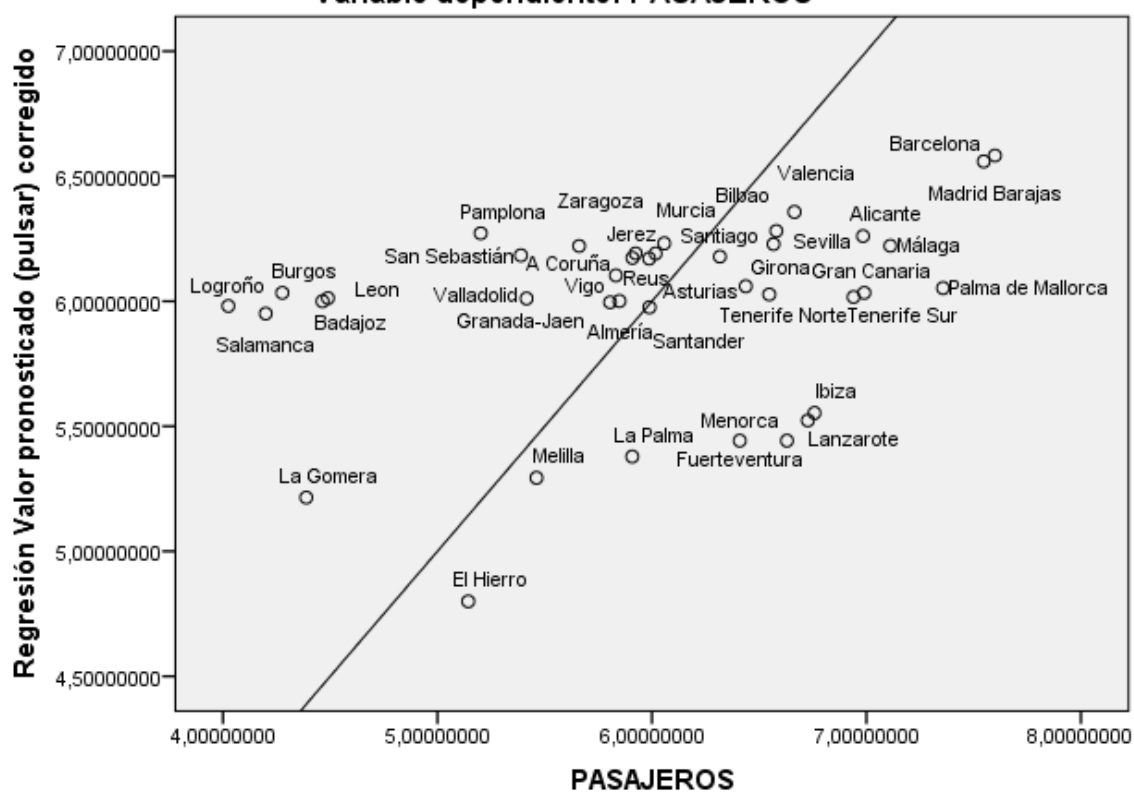
Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
9	,400	,160	,137	,88085367278	1,961

Coefficientes

Modelo	Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
	B	Error típ.	Beta		
9 (Constante)	4,170	,700		5,957	,000
Í.ECONÓMICO	,600	,226	,400	2,652	,012

Gráfico de dispersión

Variable dependiente: PASAJEROS



ÍNDICE INDUSTRIAL

Resumen del modelo

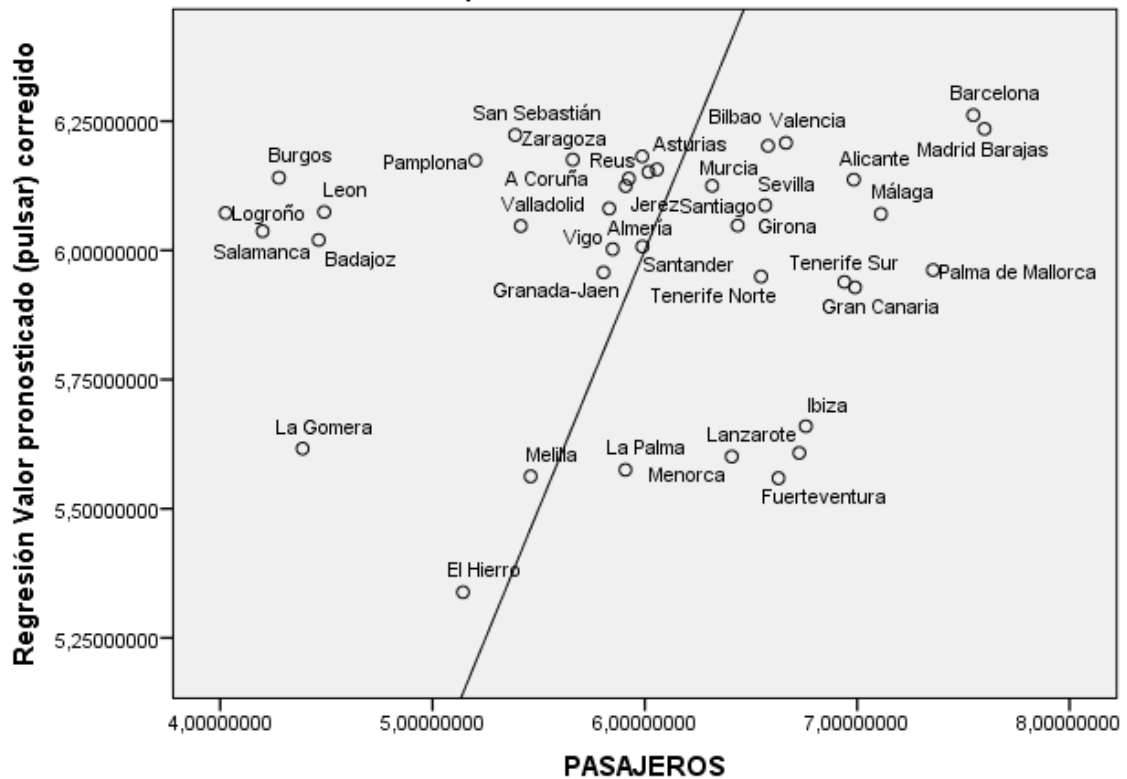
Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
10	,259	,067	,042	,92823861903	1,965

Coeficientes

Modelo		Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
		B	Error típ.	Beta		
10	(Constante)	4,947	,657		7,530	,000
	Í.INDUSTRIAL	,349	,214	,259	1,629	,112

Gráfico de dispersión

Variable dependiente: PASAJEROS



PERNOCTACIONES

Resumen del modelo

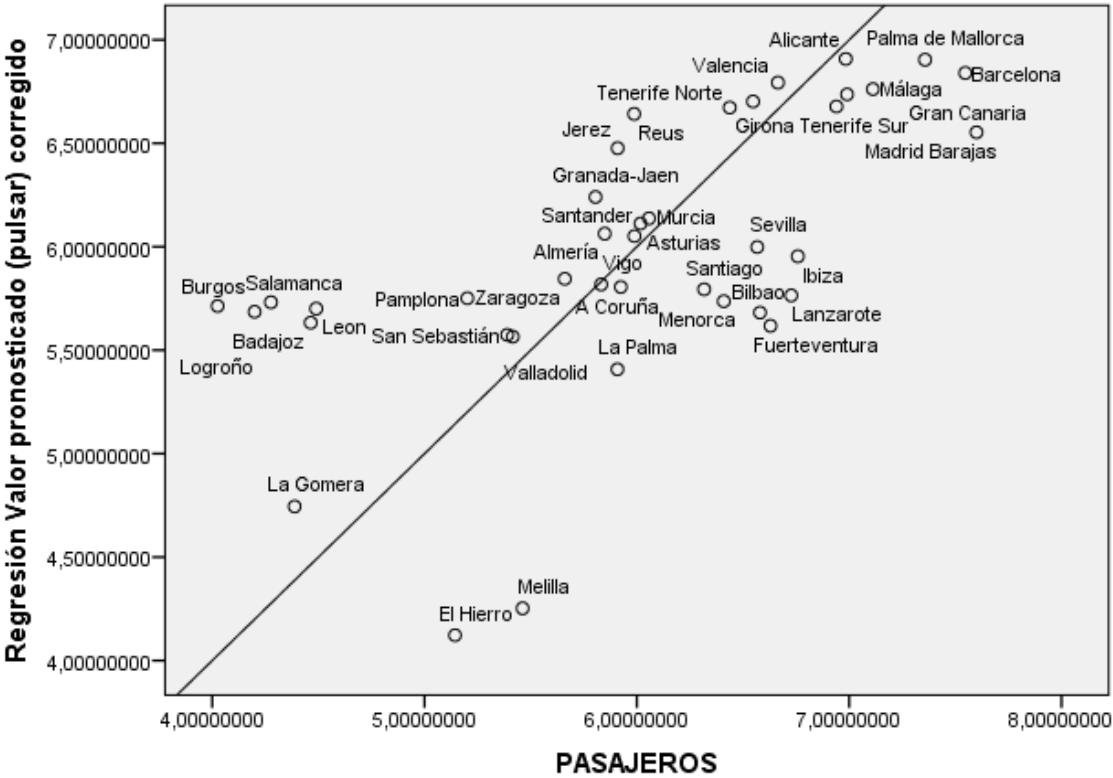
Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
11	,672	,452	,437	,71154273699	2,050

Coefficientes

Modelo		Coeficientes no estandarizados		Coeficientes tipificados	t
		B	Error típ.	Beta	
11	(Constante)	-2,792	1,594		-1,751
	PERNOCTACIONES	1,226	,222	,672	5,521

Gráfico de dispersión

Variable dependiente: PASAJEROS



EXPORTACIONES

Resumen del modelo

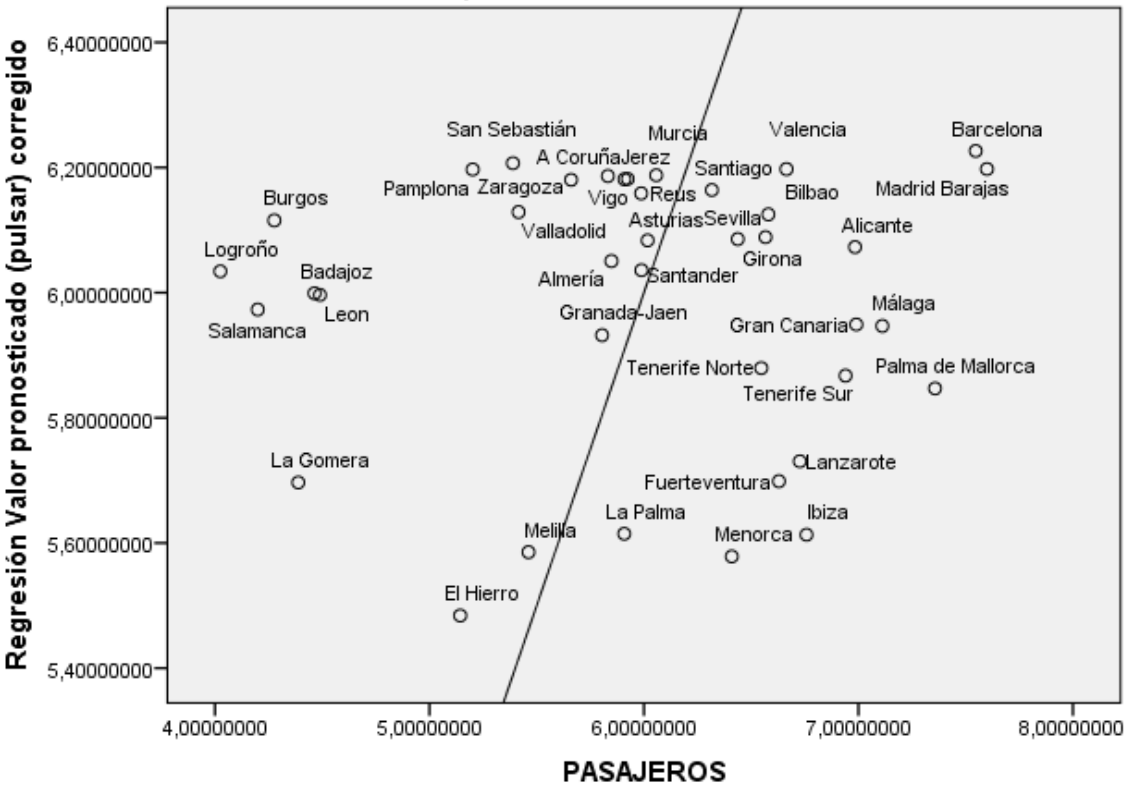
Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
12	,235	,055	,030	,93409700925	1,976

Coefficientes

Modelo		Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
		B	Error típ.	Beta		
12	(Constante)	5,186	,567		9,147	,000
	EXPORTACIONES	,251	,171	,235	1,468	,150

Gráfico de dispersión

Variable dependiente: PASAJEROS



IMPORTACIONES

Resumen del modelo

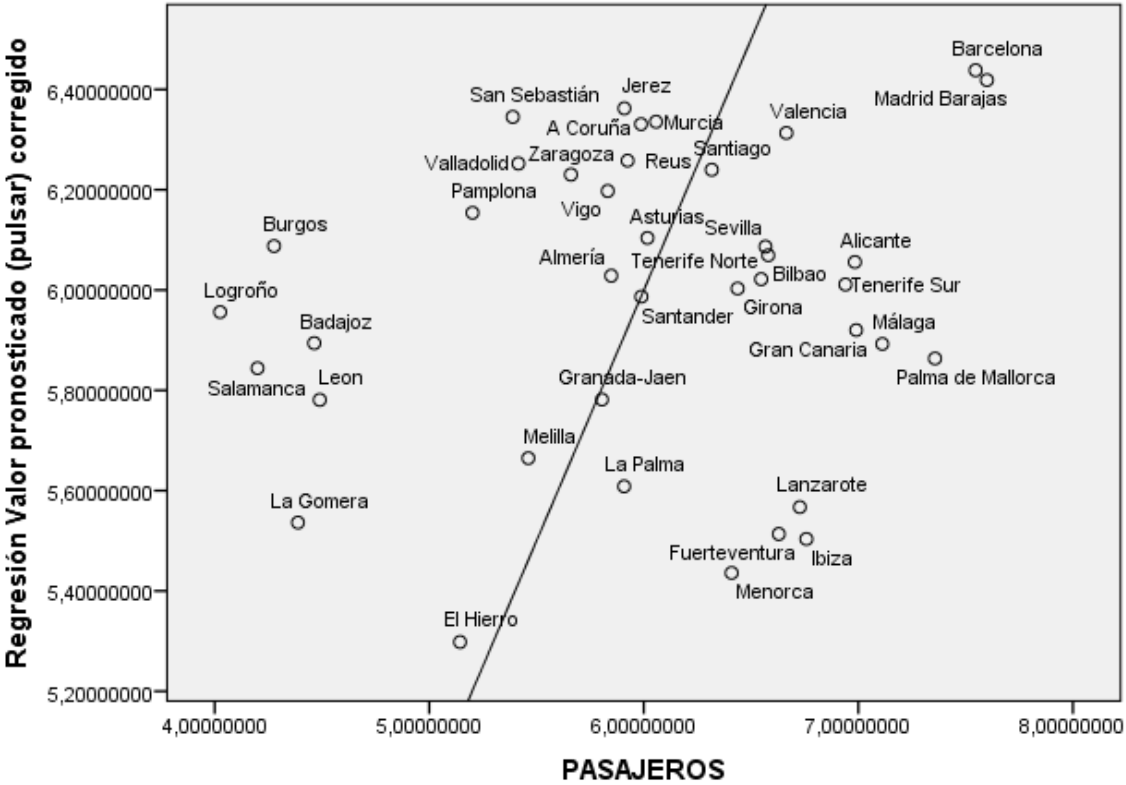
Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
13	,321	,103	,079	,91005584882	1,967

Coeficientes

Modelo		Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
		B	Error típ.	Beta		
13	(Constante)	4,703	,640		7,347	,000
	IMPORTACIONES	,397	,192	,321	2,062	,046

Gráfico de dispersión

Variable dependiente: PASAJEROS



MERCADO POTENCIAL TOTAL

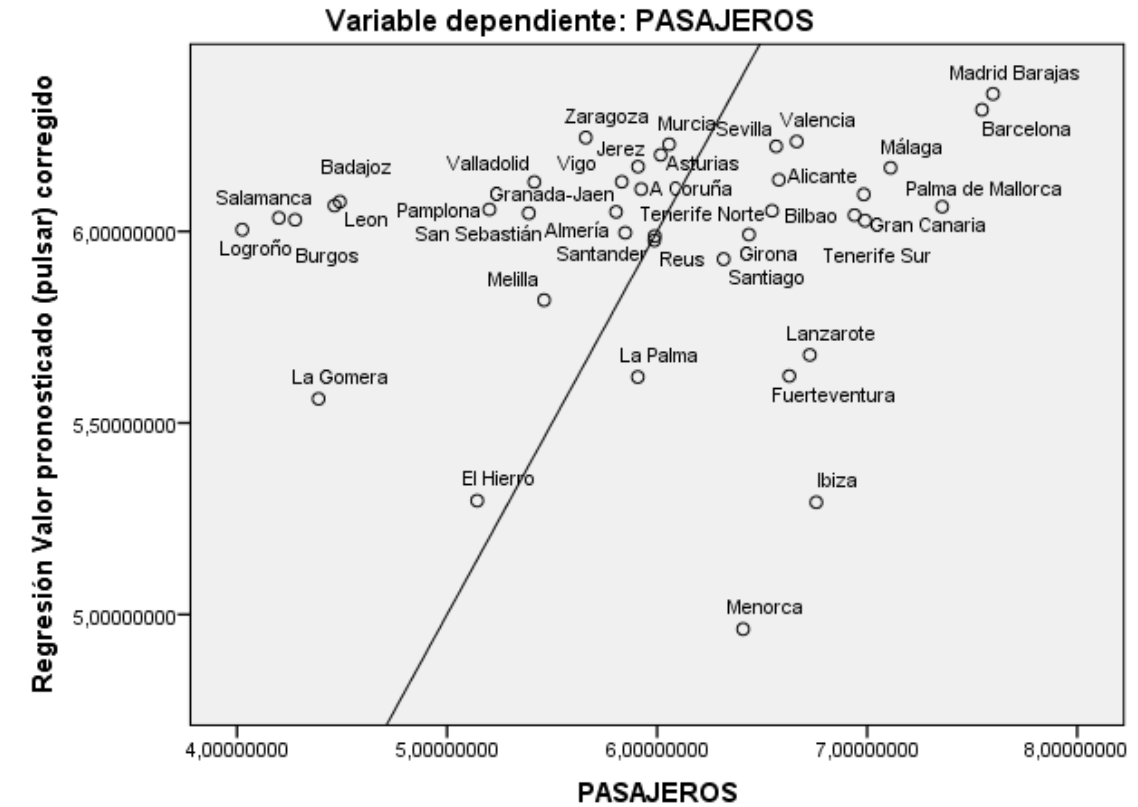
Resumen del modelo

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
14	,294	,086	,062	,91847941408	2,069

Coeficientes

Modelo		Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
		B	Error típ.	Beta		
14	(Constante)	4,761	,673		7,079	,000
	POTENCIAL	,416	,222	,294	1,871	,069

Gráfico de dispersión



CUOTA DE MERCADO

Resumen del modelo

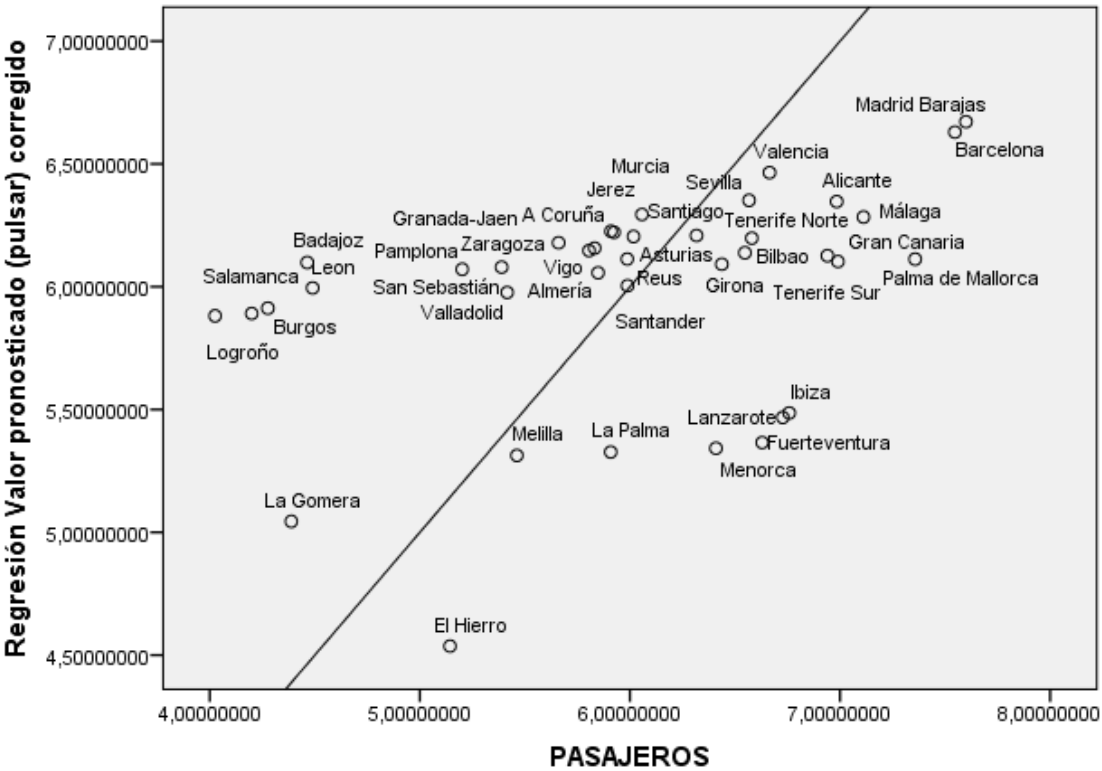
Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
15	,454	,206	,185	,85604728801	2,033

Coefficientes

Modelo		Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
		B	Error típ.	Beta		
15	(Constante)	3,658	,764		4,790	,000
	CUOTA	,755	,243	,454	3,102	,004

Gráfico de dispersión

Variable dependiente: PASAJEROS



COMPañÍAS AÉREAS

Resumen del modelo

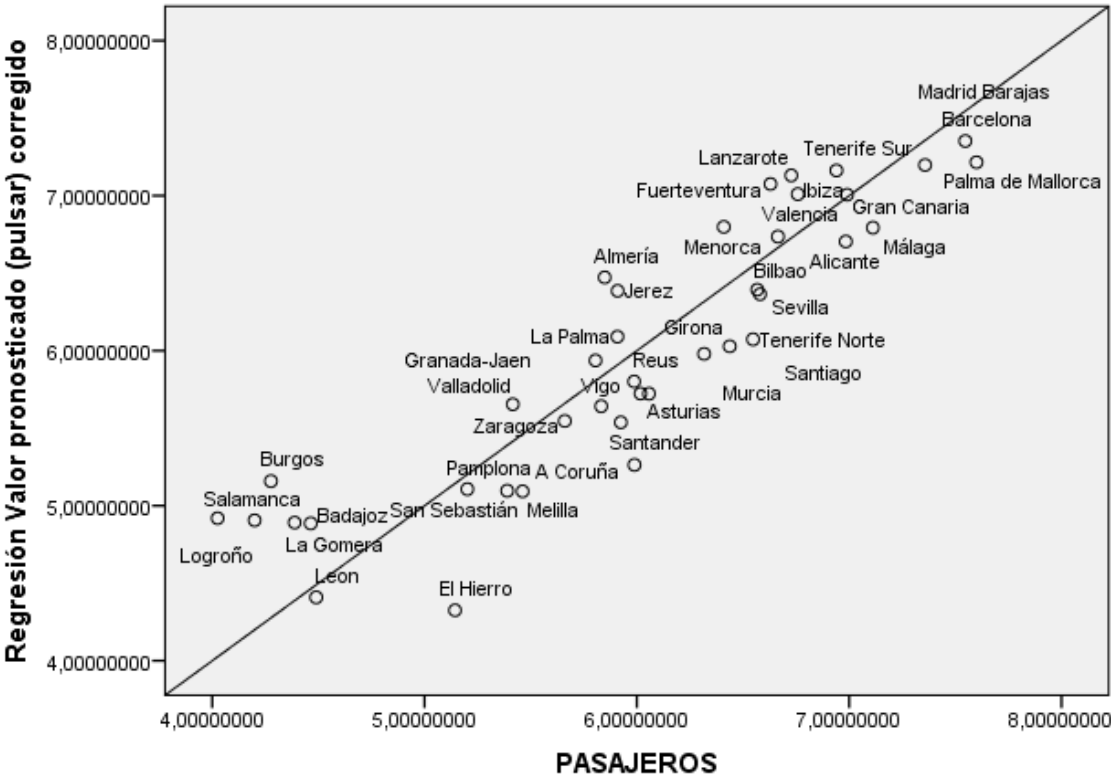
Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
16	,909	,826	,821	,40080566625	2,412

Coefficientes

Modelo		Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
		B	Error típ.	Beta		
16	(Constante)	4,418	,135		32,781	,000
	COMPañÍAS	1,455	,110	,909	13,254	,000

Gráfico de dispersión

Variable dependiente: PASAJEROS



DESTINOS

Resumen del modelo

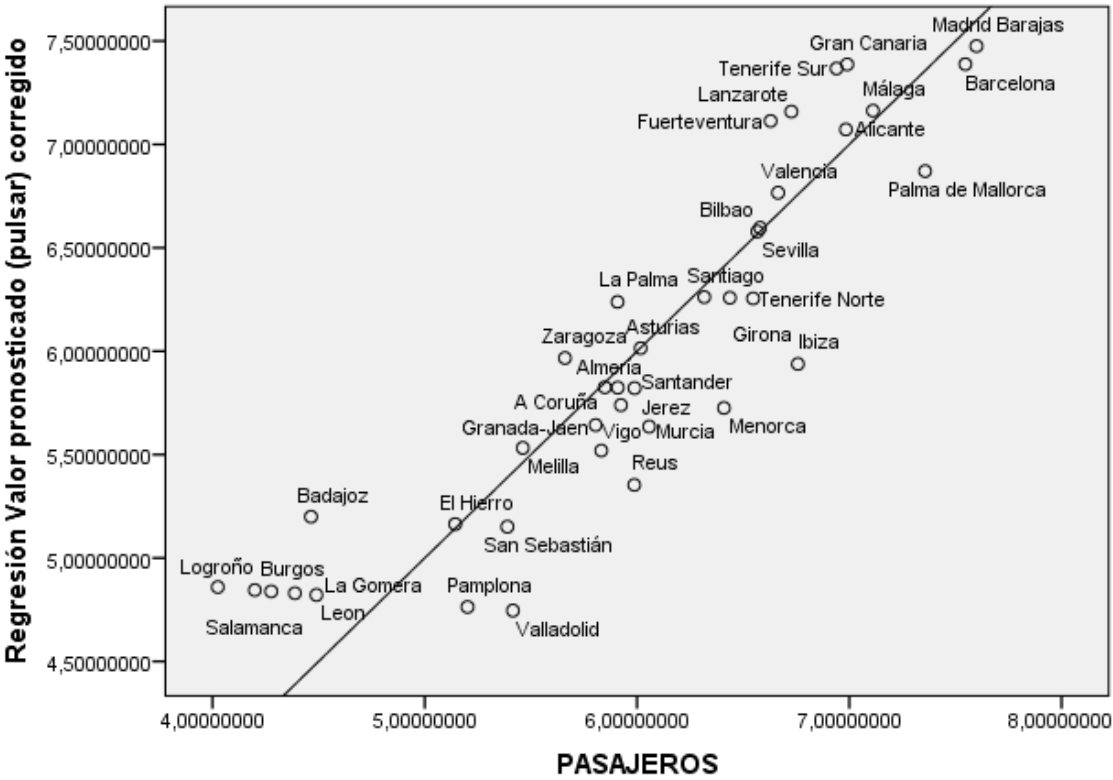
Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
17	,915	,836	,832	,38874707575	1,709

Coefficientes

Modelo		Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
		B	Error típ.	Beta		
17	(Constante)	4,796	,107		44,936	,000
	DESTINOS	1,218	,089	,915	13,750	,000

Gráfico de dispersión

Variable dependiente: PASAJEROS



DESTINOS AL AEROPUERTO MÁS CERCANO

Resumen del modelo

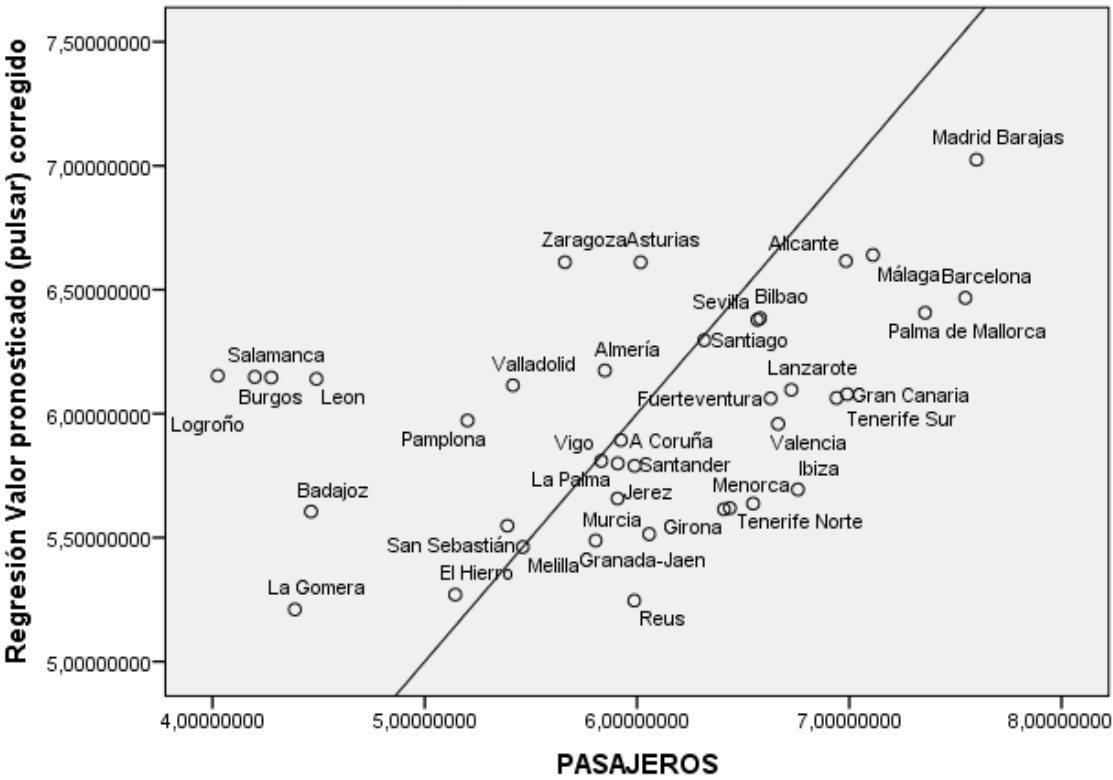
Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
18	,468	,219	,197	,84942238775	1,782

Coeficientes

Modelo		Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
		B	Error típ.	Beta		
18	(Constante)	6,095	,140		43,538	,000
	DESTINOS.CERCA	,473	,147	,468	3,217	,003

Gráfico de dispersión

Variable dependiente: PASAJEROS



DISTANCIA Y MAYORES DESTINOS

Resumen del modelo

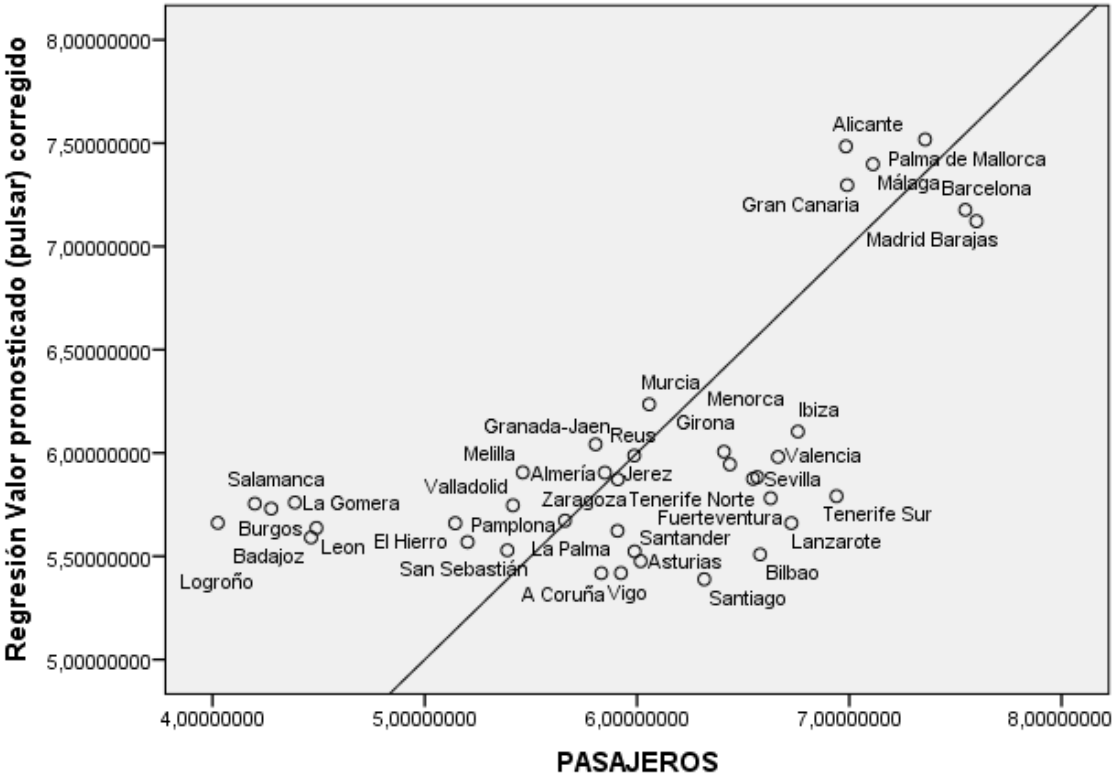
Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
19	,641	,411	,395	,73757164949	1,928

Coefficientes

Modelo		Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
		B	Error típ.	Beta		
19	(Constante)	8,570	,522		16,424	,000
	KMS.DESTINOS	-,628	,124	-,641	-5,080	,000

Gráfico de dispersión

Variable dependiente: PASAJEROS



RENTA PER CÁPITA POR PAÍSES

Resumen del modelo

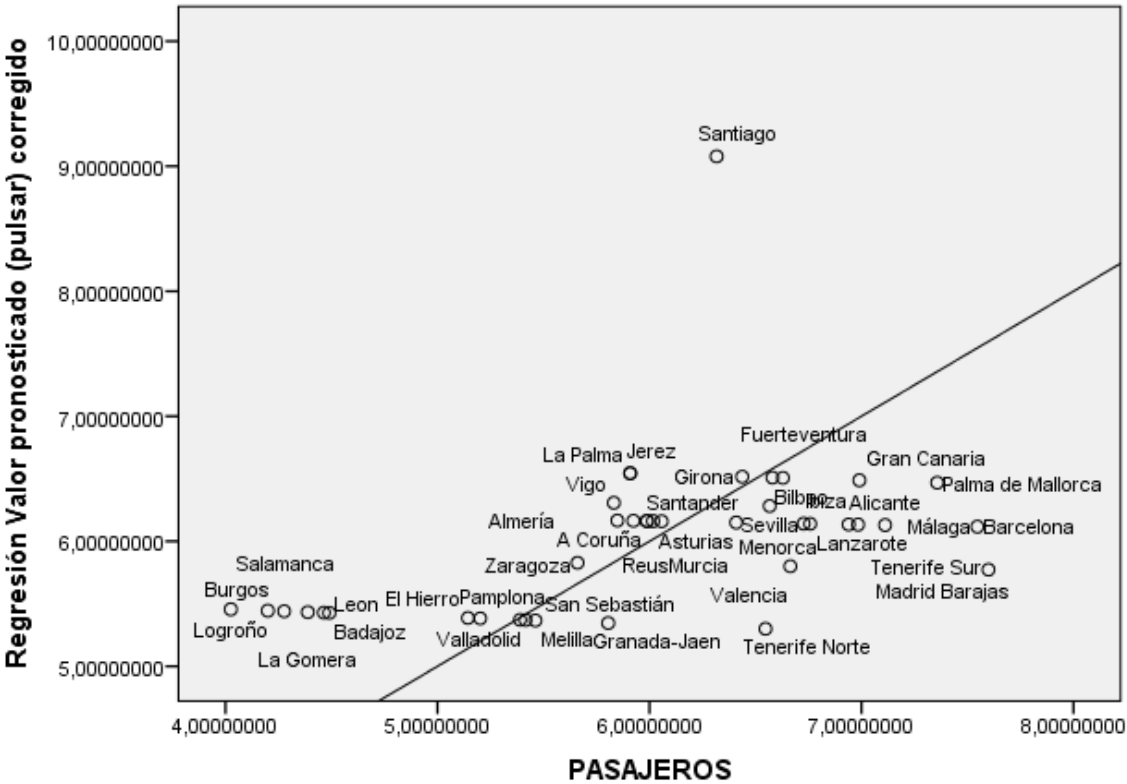
Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
20	,584	,341	,323	,78033124196	1,866

Coefficientes

Modelo		Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
		B	Error típ.	Beta		
20	(Constante)	-21,438	6,275		-3,416	,002
	RENTA.PAÍSES	6,006	1,374	,584	4,371	,000

Gráfico de dispersión

Variable dependiente: PASAJEROS



COMPañÍAS LOW COST CON MS PASAJEROS

Resumen del modelo

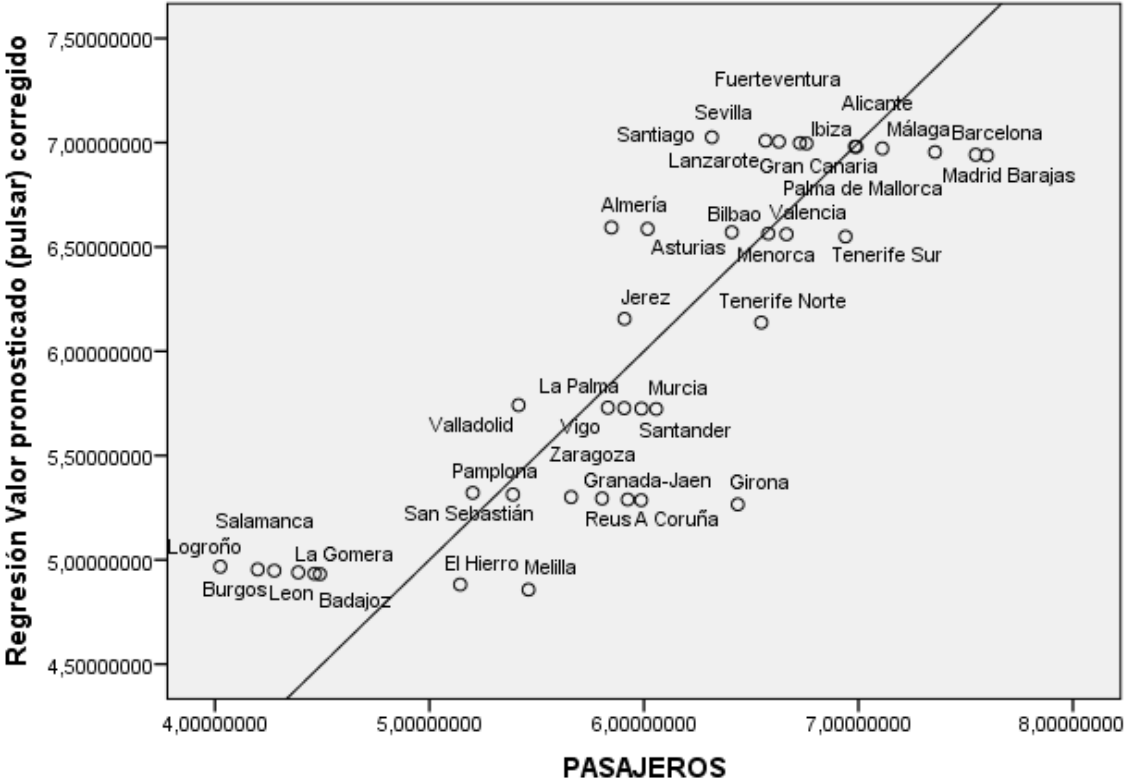
Modelo	R	R cuadrado	R cuadrado corregida	Error tp. de la estimacin	Durbin-Watson
21	,870	,756	,750	,47422463863	2,125

Coefficientes

Modelo		Coefficients no estandarizados		Coefficientes tipificados	t	Sig.
		B	Error tp.	Beta		
21	(Constante)	4,900	,127		38,650	,000
	LOW.COST.CINCO	,416	,039	,870	10,720	,000

Grfico de dispersin

Variable dependiente: PASAJEROS



COMPañÍAS LOW COST TOTALES

Resumen del modelo

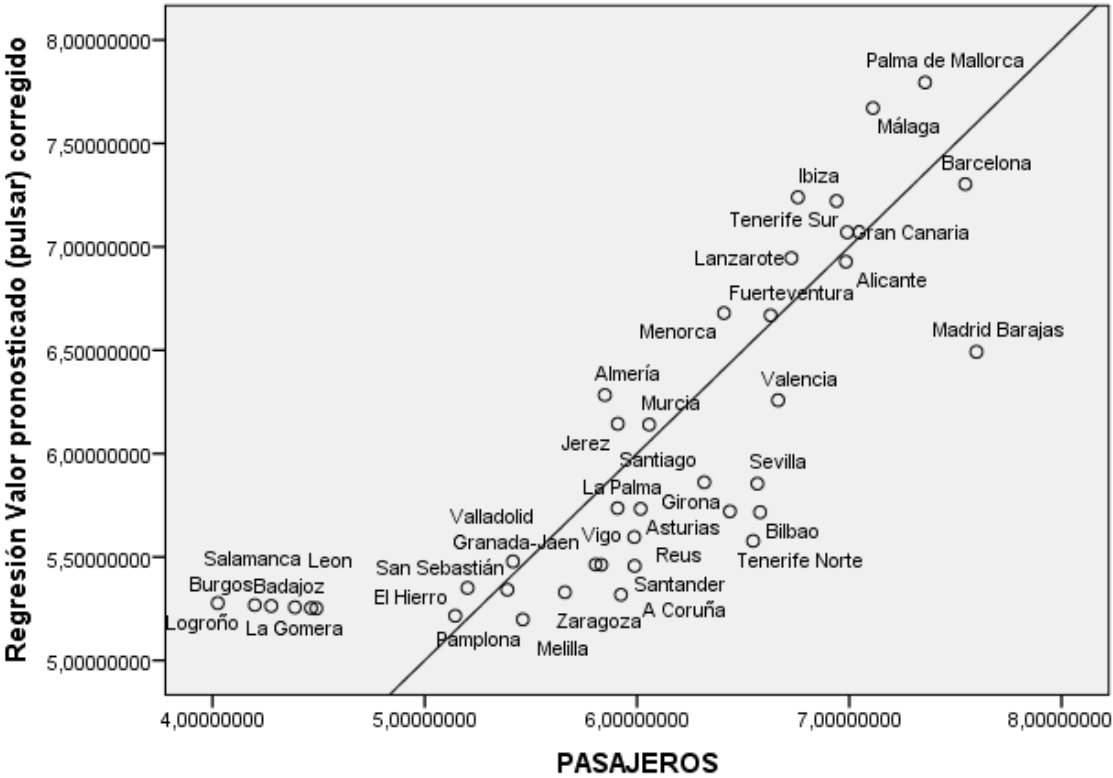
Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
22	,816	,666	,657	,55513589823	2,233

Coefficientes

Modelo		Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
		B	Error típ.	Beta		
22	(Constante)	5,212	,127		41,109	,000
	LOW.COST.TOTAL	,132	,015	,816	8,594	,000

Gráfico de dispersión

Variable dependiente: PASAJEROS



HUBS

Resumen del modelo

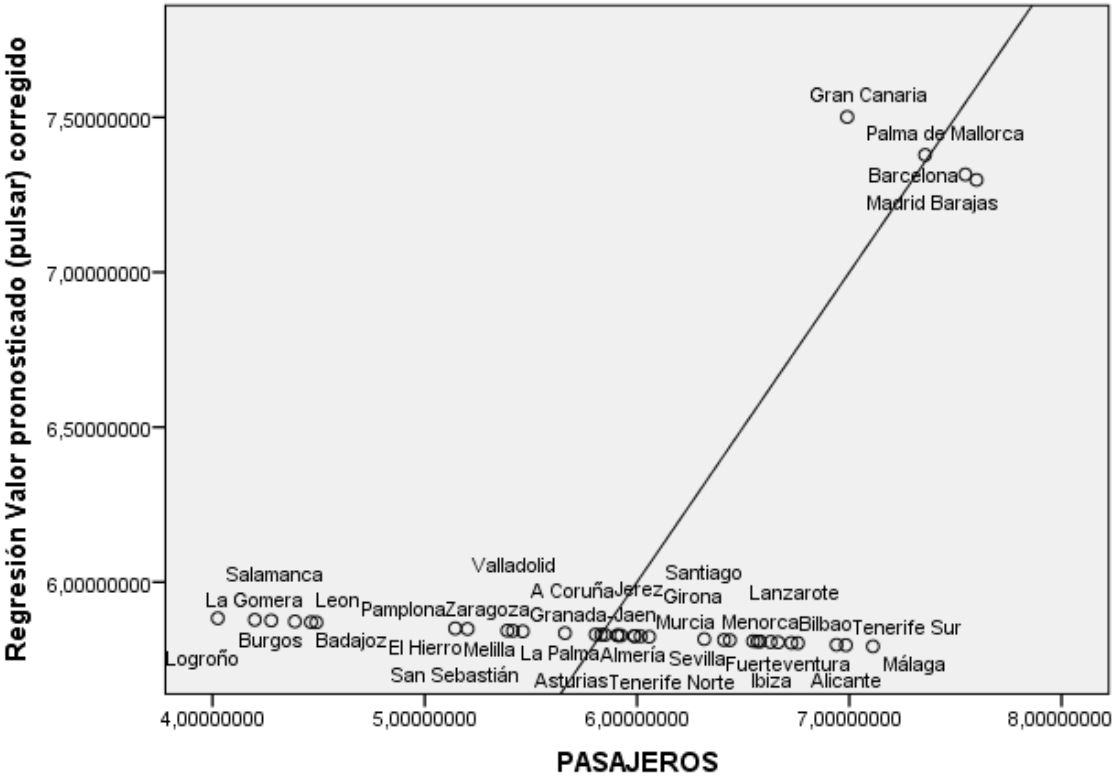
Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
23	,500	,250	,230	,83211655742	1,806

Coefficientes

Modelo		Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
		B	Error típ.	Beta		
23	(Constante)	5,830	,141		41,452	,000
	HUBS	1,543	,439	,500	3,513	,001

Gráfico de dispersión

Variable dependiente: PASAJEROS



POBLACIÓN POR ÁREA DISUASORIA (AÑO 2009)

Resumen del modelo

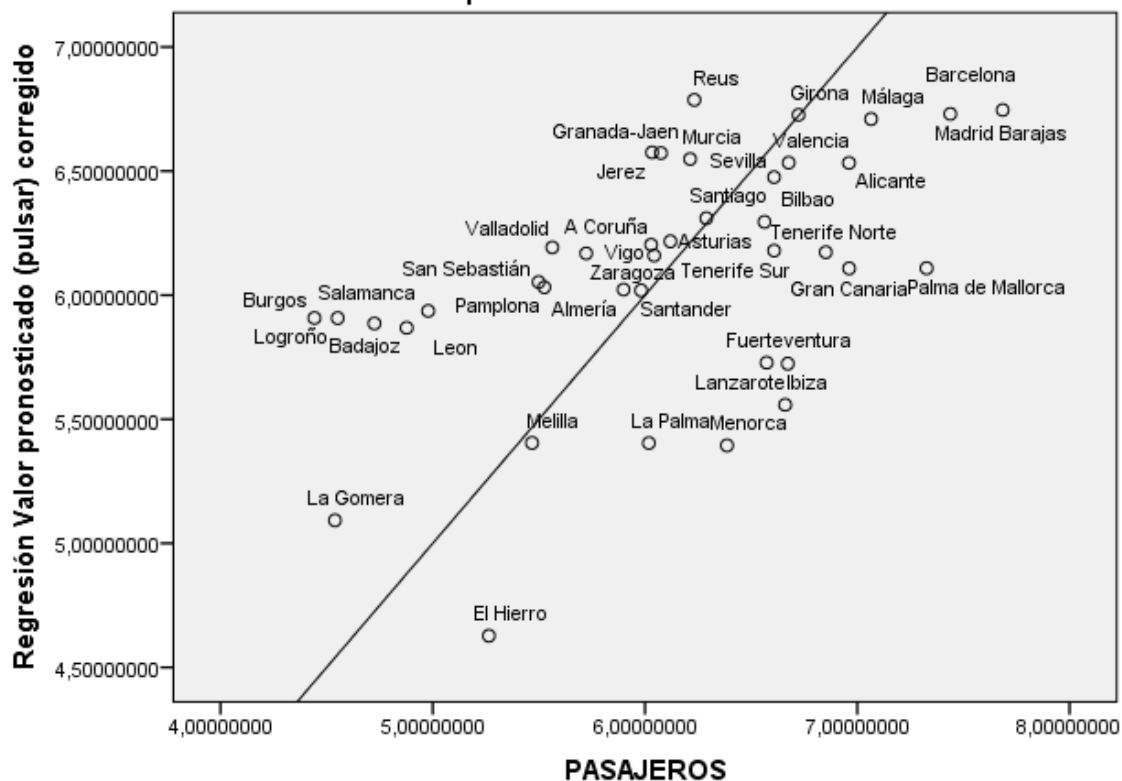
Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación	Durbin-Watson
24	,575	,331	,313	,68255164707	2,198

Coefficientes

Modelo	Coeficientes no estandarizados		Coeficientes tipificados	t	Sig.
	B	Error típ.	Beta		
24	(Constante)	2,037	,955	2,132	,040
	ÁREA.DISUASORIA	,691	,161	4,279	,000

Gráfico de dispersión

Variable dependiente: PASAJEROS



5.8 MÁS ESTADÍSTICOS AÑO 2013

Estadísticos descriptivos

	Media	Desviación típica	N
PASAJEROS	5,9886079728	,94820214288	39
POBLACIÓN	5,749796	,5770142	39
RESIDENTES	4,340882	,5685893	39
RENTA	4,321375	,0836789	39
PARO	1,414430	,1172130	39
Í.COMERCIAL	,10	,307	39
Í.RESTAURACIÓN	3,075776	,5983375	39
Í.TURÍSTICO	3,121616	,5542093	39
Í.ECONÓMICO	3,030533	,6315811	39
Í.INDUSTRIAL	2,983850	,7021521	39
PERNOCTACIONES	7,159333	,5196123	39
IMPORTACIONES	3,240894	,7675959	39
POTENCIAL	2,951012	,6699986	39
KMS.DESTINOS	4,111493	,9678509	39
RENTA.PAÍSES	4,566738	,0921375	39
LOW.COST.CINCO	2,62	1,982	39
LOW.COST.TOTAL	5,87	5,850	39

Correlaciones

		PASAJEROS	POBLACIÓN	RESIDENTES
Sig. (unilateral)	PASAJEROS	.	,002	,000
	POBLACIÓN	,002	.	,000
	RESIDENTES	,000	,000	.
	RENTA	,366	,135	,200
	PARO	,188	,096	,482
	Í.COMERCIAL	,001	,013	,002
	Í.RESTAURACIÓN	,002	,000	,000
	Í.TURÍSTICO	,000	,000	,000
	Í.ECONÓMICO	,006	,000	,000
	Í.INDUSTRIAL	,056	,000	,000
	PERNOCTACIONES	,000	,000	,000
	IMPORTACIONES	,023	,000	,000
	POTENCIAL	,035	,000	,000
	KMS.DESTINOS	,000	,007	,000
	RENTA.PAÍSES	,000	,060	,089
	LOW.COST.CINCO	,000	,007	,000
	LOW.COST.TOTAL	,000	,047	,000

Correlaciones

		RENTA	PARO	Í.COMERCIAL
Sig. (unilateral)	PASAJEROS	,366	,188	,001
	POBLACIÓN	,135	,096	,013
	RESIDENTES	,200	,482	,002
	RENTA	.	,000	,059
	PARO	,000	.	,306
	Í.COMERCIAL	,059	,306	.
	Í.RESTAUACIÓN	,082	,046	,012
	Í.TURÍSTICO	,276	,318	,000
	Í.ECONÓMICO	,020	,013	,014
	Í.INDUSTRIAL	,008	,002	,059
	PERNOCTACIONES	,286	,403	,003
	IMPORTACIONES	,008	,011	,037
	POTENCIAL	,197	,131	,015
	KMS.DESTINOS	,433	,108	,000
	RENTA.PAÍSES	,497	,385	,161
	LOW.COST.CINCO	,403	,172	,005
	LOW.COST.TOTAL	,338	,069	,000

Correlaciones

		Í.RESTAUACIÓN	Í.TURÍSTICO	Í.ECONÓMICO
Sig. (unilateral)	PASAJEROS	,002	,000	,006
	POBLACIÓN	,000	,000	,000
	RESIDENTES	,000	,000	,000
	RENTA	,082	,276	,020
	PARO	,046	,318	,013
	Í.COMERCIAL	,012	,000	,014
	Í.RESTAUACIÓN	.	,000	,000
	Í.TURÍSTICO	,000	.	,000
	Í.ECONÓMICO	,000	,000	.
	Í.INDUSTRIAL	,000	,000	,000
	PERNOCTACIONES	,000	,000	,000
	IMPORTACIONES	,000	,000	,000
	POTENCIAL	,000	,000	,000
	KMS.DESTINOS	,007	,000	,014
	RENTA.PAÍSES	,061	,004	,088
	LOW.COST.CINCO	,006	,000	,015
	LOW.COST.TOTAL	,044	,000	,081

Correlaciones

		Í.INDUSTRIAL	PERNOCTACIONES	IMPORTACIONES
Sig. (unilateral)	PASAJEROS	,056	,000	,023
	POBLACIÓN	,000	,000	,000
	RESIDENTES	,000	,000	,000
	RENTA	,008	,286	,008
	PARO	,002	,403	,011
	Í.COMERCIAL	,059	,003	,037
	Í.RESTAUACIÓN	,000	,000	,000
	Í.TURÍSTICO	,000	,000	,000
	Í.ECONÓMICO	,000	,000	,000
	Í.INDUSTRIAL	.	,000	,000
	PERNOCTACIONES	,000	.	,000
	IMPORTACIONES	,000	,000	.
	POTENCIAL	,000	,000	,000
	KMS.DESTINOS	,074	,000	,114
	RENTA.PAÍSES	,147	,013	,092
	LOW.COST.CINCO	,092	,000	,086
	LOW.COST.TOTAL	,279	,000	,320

Correlaciones

		POTENCIAL	KMS.DESTINOS	RENTA.PAÍSES
Sig. (unilateral)	PASAJEROS	,035	,000	,000
	POBLACIÓN	,000	,007	,060
	RESIDENTES	,000	,000	,089
	RENTA	,197	,433	,497
	PARO	,131	,108	,385
	Í.COMERCIAL	,015	,000	,161
	Í.RESTAUACIÓN	,000	,007	,061
	Í.TURÍSTICO	,000	,000	,004
	Í.ECONÓMICO	,000	,014	,088
	Í.INDUSTRIAL	,000	,074	,147
	PERNOCTACIONES	,000	,000	,013
	IMPORTACIONES	,000	,114	,092
	POTENCIAL	.	,024	,321
	KMS.DESTINOS	,024	.	,161
	RENTA.PAÍSES	,321	,161	.
	LOW.COST.CINCO	,098	,000	,000
	LOW.COST.TOTAL	,251	,000	,001

Correlaciones		LOW.COST.CINCO	LOW.COST.TOTAL
Sig. (unilateral)	PASAJEROS	,000	,000
	POBLACIÓN	,007	,047
	RESIDENTES	,000	,000
	RENTA	,403	,338
	PARO	,172	,069
	Í.COMERCIAL	,005	,000
	Í.RESTAURACIÓN	,006	,044
	Í.TURÍSTICO	,000	,000
	Í.ECONÓMICO	,015	,081
	Í.INDUSTRIAL	,092	,279
	PERNOCTACIONES	,000	,000
	IMPORTACIONES	,086	,320
	POTENCIAL	,098	,251
	KMS.DESTINOS	,000	,000
	RENTA.PAÍSES	,000	,001
	LOW.COST.CINCO	.	,000
	LOW.COST.TOTAL	,000	.

ANOVA ^a					
Modelo	Suma de cuadrados	gl	Media cuadrática	F	Sig.
1 Regresión	33,241	21	1,583	29,122	,000 ^b
Residual	,924	17	,054		
Total	34,165	38			

a. Variable dependiente: PASAJEROS

b. Variables predictoras: (Constante), HUBS, PARO, RENTA.PAÍSES, POTENCIAL, LOW.COST.TOTAL, DESTINOS.CERCA, PERNOCTACIONES, DESTINOS, KMS.DESTINOS, RENTA, LOW.COST.CINCO, IMPORTACIONES, RESIDENTES, COMPAÑÍAS, Í.TURÍSTICO, EXPORTACIONES, Í.RESTAURACIÓN, Í.INDUSTRIAL, Í.COMERCIAL, POBLACIÓN, Í.ECONÓMICO

Coefficients^a

Modelo	Sig.	Correlaciones			Estadísticos de colinealidad
		Orden cero	Parcial	Semiparcial	Tolerancia
(Constante)	,233				
POBLACIÓN	,293	,459	,254	,043	,002
RESIDENTES	,411	,627	-,200	-,034	,050
RENTA	,102	,057	,387	,069	,051
PARO	,543	,146	-,149	-,025	,062
Í.COMERCIAL	,006	,500	-,610	-,127	,210
Í.RESTAURACIÓN	,121	,445	,368	,065	,003
Í.TURÍSTICO	,799	,785	,063	,010	,027
1 Í.ECONÓMICO	,468	,400	,177	,030	,001
Í.INDUSTRIAL	,039	,259	-,477	-,089	,003
PERNOCTACIONES	,735	,672	-,083	-,014	,029
IMPORTACIONES	,428	,321	,193	,032	,036
POTENCIAL	,191	,294	,313	,054	,031
KMS.DESTINOS	,010	-,641	-,574	-,115	,132
RENTA.PAÍSES	,692	,584	,097	,016	,262
LOW.COST.CINCO	,028	,870	,505	,096	,073
LOW.COST.TOTAL	,924	,816	,023	,004	,067

Coeficientes ^a		
Modelo	Estadísticos de colinealidad	
	FIV	
1	(Constante)	
	POBLACIÓN	606,866
	RESIDENTES	19,976
	RENTA	19,418
	PARO	16,067
	Í.COMERCIAL	4,751
	Í.RESTAURACIÓN	286,515
	Í.TURÍSTICO	37,631
	Í.ECONÓMICO	804,153
	Í.INDUSTRIAL	342,062
	PERNOCTACIONES	34,327
	IMPORTACIONES	27,823
	POTENCIAL	32,176
	KMS.DESTINOS	7,552
	RENTA.PAÍSES	3,815
	LOW.COST.CINCO	13,763
	LOW.COST.TOTAL	14,951

a. Variable dependiente: PASAJEROS

Diagnósticos de colinealidad^a

Modelo	Dimensión	Autovalores	Índice de condición	Proporciones de la varianza		
				(Constante)	POBLACIÓN	RESIDENTES
1	1	18,768	1,000	,00	,00	,00
	2	1,889	3,152	,00	,00	,00
	3	,577	5,703	,00	,00	,00
	4	,515	6,036	,00	,00	,00
	5	,079	15,436	,00	,00	,00
	6	,074	15,907	,00	,00	,00
	7	,050	19,460	,00	,00	,00
	8	,017	33,554	,00	,00	,00
	9	,012	38,869	,00	,00	,01
	10	,007	53,134	,00	,00	,00
	11	,005	62,966	,00	,00	,00
	12	,003	76,586	,00	,00	,01
	13	,002	99,960	,00	,00	,00
	14	,001	136,505	,00	,00	,15
	15	,001	153,727	,00	,00	,52
	16	,000	214,006	,00	,00	,00
	17	,000	262,701	,00	,00	,01
	18	,000	382,800	,00	,00	,15
	19	7,335E-005	505,830	,01	,00	,07
	20	3,516E-005	730,574	,02	,03	,00
	21	1,292E-005	1205,442	,01	,93	,06
	22	6,911E-006	1647,967	,97	,04	,00

Diagnósticos de colinealidad^a

Modelo	Dimensión	Proporciones de la varianza				
		RENTA	PARO	Í.COMERCIAL	Í.RESTAURACIÓN	Í.TURÍSTICO
1	1	,00	,00	,00	,00	,00
	2	,00	,00	,02	,00	,00
	3	,00	,00	,03	,00	,00
	4	,00	,00	,23	,00	,00
	5	,00	,00	,04	,00	,00
	6	,00	,00	,00	,00	,00
	7	,00	,00	,06	,00	,00
	8	,00	,00	,00	,00	,00
	9	,00	,00	,02	,00	,01
	10	,00	,01	,00	,00	,03
	11	,00	,00	,07	,00	,00
	12	,00	,00	,01	,00	,09
	13	,00	,04	,03	,00	,00
	14	,00	,04	,00	,03	,01
	15	,00	,01	,23	,00	,17
	16	,00	,12	,16	,02	,01
	17	,00	,03	,03	,08	,14
	18	,00	,00	,00	,08	,06
	19	,04	,09	,00	,28	,01
	20	,00	,00	,05	,07	,47
	21	,05	,17	,02	,43	,00
	22	,90	,48	,01	,01	,01

Diagnósticos de colinealidad^a

Modelo	Dimensión	Proporciones de la varianza			
		Í.ECONÓMICO	Í.INDUSTRIAL	PERNOCTACIONES	IMPORTACIONES
1	1	,00	,00	,00	,00
	2	,00	,00	,00	,00
	3	,00	,00	,00	,00
	4	,00	,00	,00	,00
	5	,00	,00	,00	,00
	6	,00	,00	,00	,00
	7	,00	,00	,00	,00
	8	,00	,00	,00	,00
	9	,00	,00	,00	,00
	10	,00	,00	,00	,03
	11	,00	,00	,00	,10
	12	,00	,00	,00	,11
	13	,00	,00	,00	,14
	14	,00	,00	,00	,08
	15	,00	,00	,01	,00
	16	,00	,17	,01	,05
	17	,05	,00	,18	,01
	18	,00	,06	,42	,08
	19	,00	,00	,00	,00
	20	,67	,65	,29	,07
	21	,07	,10	,08	,20
	22	,20	,00	,00	,11

Diagnósticos de colinealidad^a

Modelo	Dimensión	Proporciones de la varianza			
		POTENCIAL	KMS.DESTINOS	RENTA.PAÍSES	LOW.COST.CINCO
1	1	,00	,00	,00	,00
	2	,00	,00	,00	,00
	3	,00	,00	,00	,00
	4	,00	,00	,00	,01
	5	,00	,02	,00	,00
	6	,00	,00	,00	,01
	7	,00	,01	,00	,27
	8	,00	,04	,00	,29
	9	,00	,16	,00	,02
	10	,06	,01	,00	,00
	11	,06	,22	,00	,00
	12	,00	,14	,00	,00
	13	,01	,00	,00	,01
	14	,38	,00	,00	,26
	15	,14	,30	,01	,00
	16	,04	,08	,00	,02
	17	,01	,00	,01	,00
	18	,16	,00	,15	,00
	19	,01	,00	,63	,05
	20	,06	,00	,13	,00
	21	,05	,00	,04	,02
	22	,01	,03	,03	,02

Diagnósticos de colinealidad^a

Modelo	Dimensión	Proporciones de la varianza
		LOW.COST.TOTAL
1	1	,00
	2	,00
	3	,01
	4	,01
	5	,02
	6	,17
	7	,03
	8	,00
	9	,37
	10	,02
	11	,03
	12	,04
	13	,00
	14	,02
	15	,04
	16	,01
	17	,02
	18	,00
	19	,02
	20	,00
	21	,02
	22	,17

a. Variable dependiente: PASAJEROS

Estadísticos sobre los residuos^a

	Mínimo	Máximo	Media	Desviación típica
Valor pronosticado	4,2863440514	7,7030768394	5,9886079728	,93529190135
Valor pronosticado tip.	-1,820	1,833	,000	1,000
Error típico de valor pronosticado	,113	,213	,174	,024
Valor pronosticado corregido	3,8157160282	7,9345679283	5,9316994270	1,05927356358
Residual	-,49510571361	,27047574520	0E-11	,15593704834
Residuo típ.	-2,124	1,160	,000	,669
Residuo estud.	-2,583	2,862	,069	1,067
Residuo eliminado	-,73232996464	1,64600908756	,05690854583	,45830907648
Residuo eliminado estud.	-3,214	3,857	,090	1,210
Dist. de Mahalanobis	7,971	30,890	20,462	5,593
Distancia de Cook	,000	1,893	,123	,337
Valor de influencia centrado	,210	,813	,538	,147

Estadísticos sobre los residuos^a

	N
Valor pronosticado	39
Valor pronosticado tip.	39
Error típico de valor pronosticado	39
Valor pronosticado corregido	39
Residual	39
Residuo típ.	39
Residuo estud.	39
Residuo eliminado	39
Residuo eliminado estud.	39
Dist. de Mahalanobis	39
Distancia de Cook	39
Valor de influencia centrado	39

a. Variable dependiente: PASAJEROS

Gráficos

Histograma

Variable dependiente: PASAJEROS

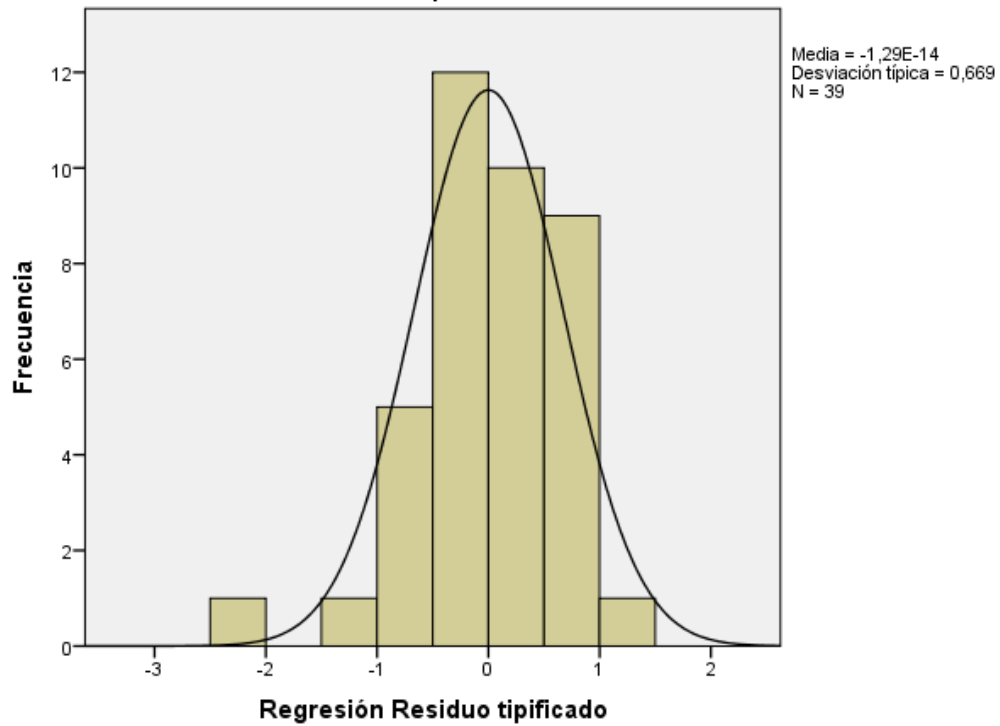
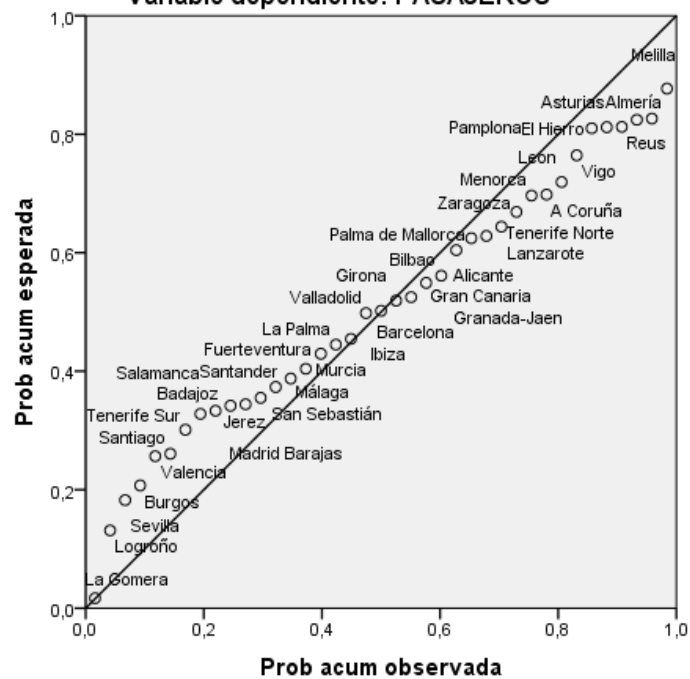
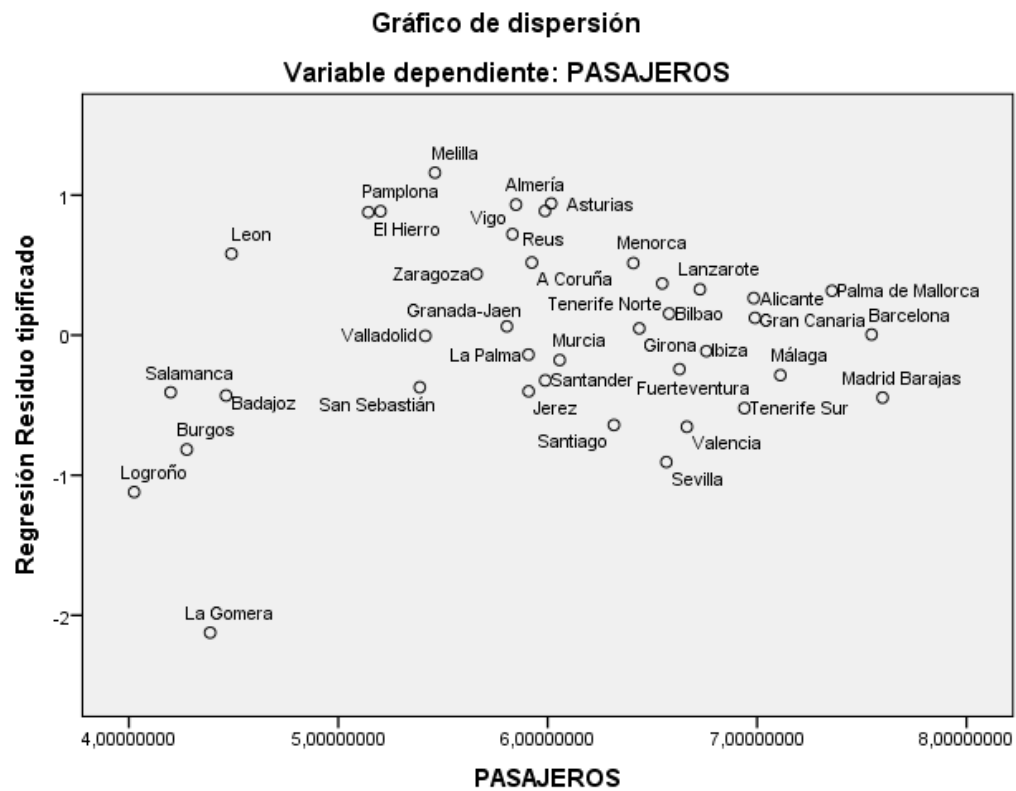


Gráfico P-P normal de regresión Residuo tipificado

Variable dependiente: PASAJEROS





5.9 MODELOS NO LINEALES AÑO 2013

COMPARACIÓN DE R CUADRADO ENTRE MODELOS NO LINEALES

	POBLACIÓN	RESIDENTES U.E.	RENTA	PARO
Ecuación	R cuadrado	R cuadrado	R cuadrado	R cuadrado
Lineal	0,211	0,393	0,003	0,021
Logarítmica	0,192	0,367	0,003	0,023
Inversa	0,174	0,334	0,003	0,024
Cuadrático	0,281	0,43	0,003	0,031
Cúbico	0,292	0,43	0,013	0,031
Compuesto	0,197	0,36	0,002	0,024
Potencia	0,181	0,337	0,002	0,025
S	0,165	0,308	0,002	0,026
Crecimiento	0,197	0,36	0,002	0,024
Exponencial	0,197	0,36	0,002	0,024
Logística	0,197	0,36	0,002	0,024

	Í.COMERCIAL	Í.RESTAURACIÓN	Í.TURÍSTICO	Í.ECONÓMICO
Ecuación	R cuadrado	R cuadrado	R cuadrado	R cuadrado
Lineal	0,214	0,198	0,617	0,16
Logarítmica	0,178	0,164	0,554	0,134
Inversa	0,142	0,133	0,462	0,111
Cuadrático	0,28	0,265	0,658	0,209
Cúbico	0,322	0,318	0,652	0,277
Compuesto	0,199	0,182	0,583	0,146
Potencia	0,166	0,153	0,527	0,123
S	0,134	0,124	0,442	0,103
Crecimiento	0,199	0,182	0,583	0,146
Exponencial	0,199	0,182	0,583	0,146
Logística	0,199	0,182	0,583	0,146

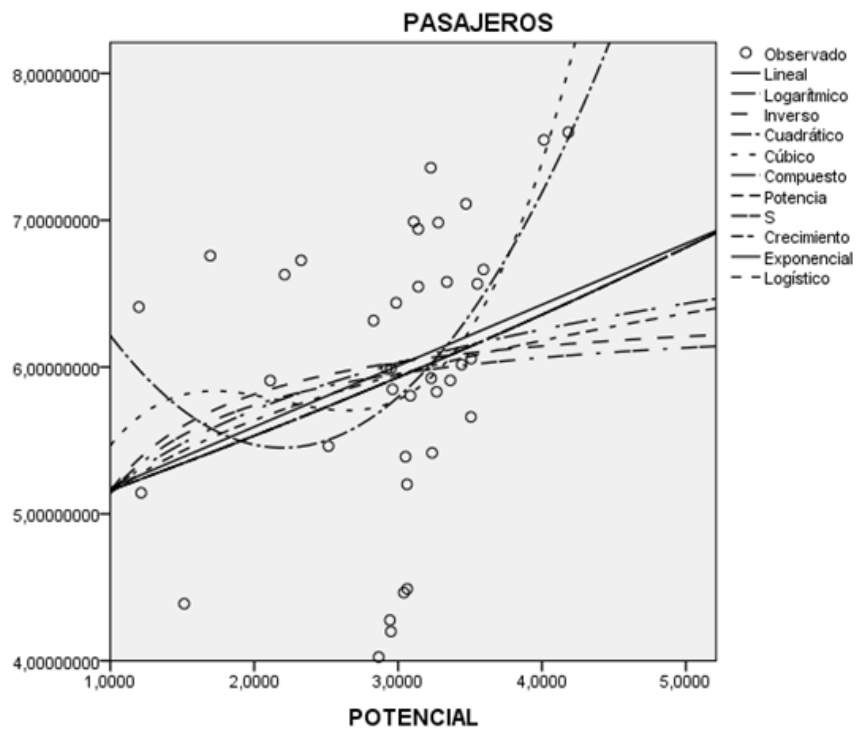
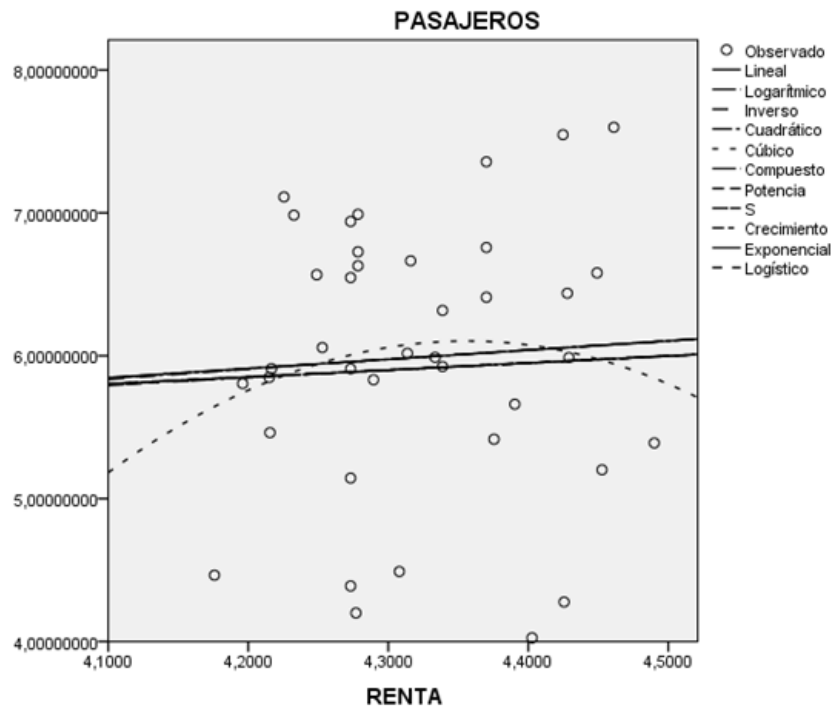
	Í.INDUSTRIAL	PERNOCTACIONES	EXPORTACIONES	IMPORTACIONES
Ecuación	R cuadrado	R cuadrado	R cuadrado	R cuadrado
Lineal	0,067	0,452	0,055	0,103
Logarítmica	0,064	0,434	0,056	0,094
Inversa	0,063	0,413	0,06	0,085
Cuadrático	0,08	0,505	0,062	0,118
Cúbico	0,258	0,504	0,257	0,179
Compuesto	0,06	0,42	0,053	0,102
Potencia	0,058	0,404	0,053	0,093
S	0,058	0,384	0,057	0,084
Crecimiento	0,06	0,42	0,053	0,102
Exponencial	0,06	0,42	0,053	0,102
Logística	0,06	0,42	0,053	0,102

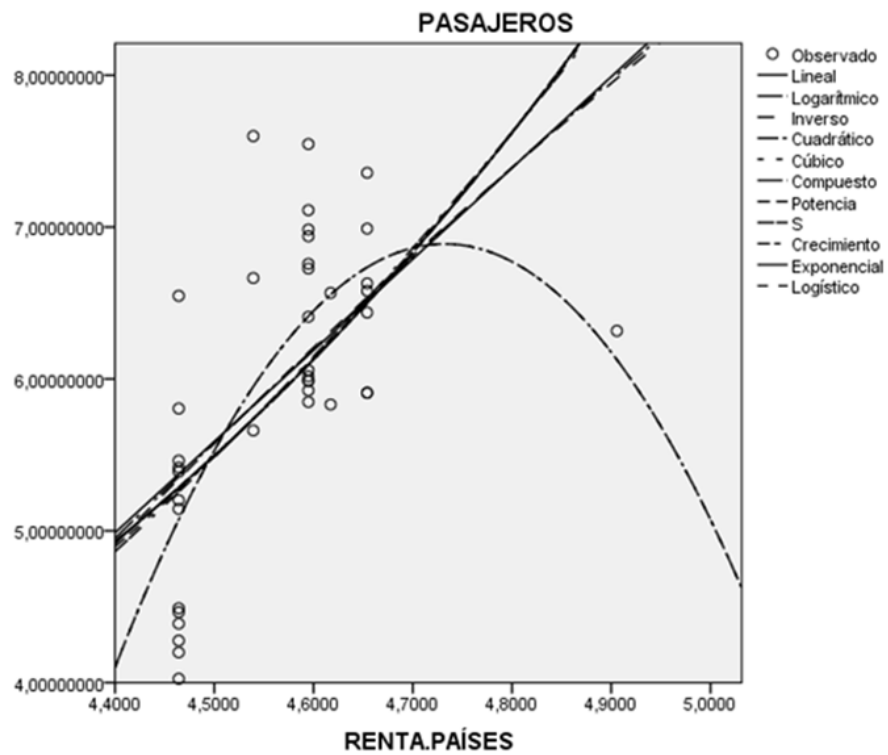
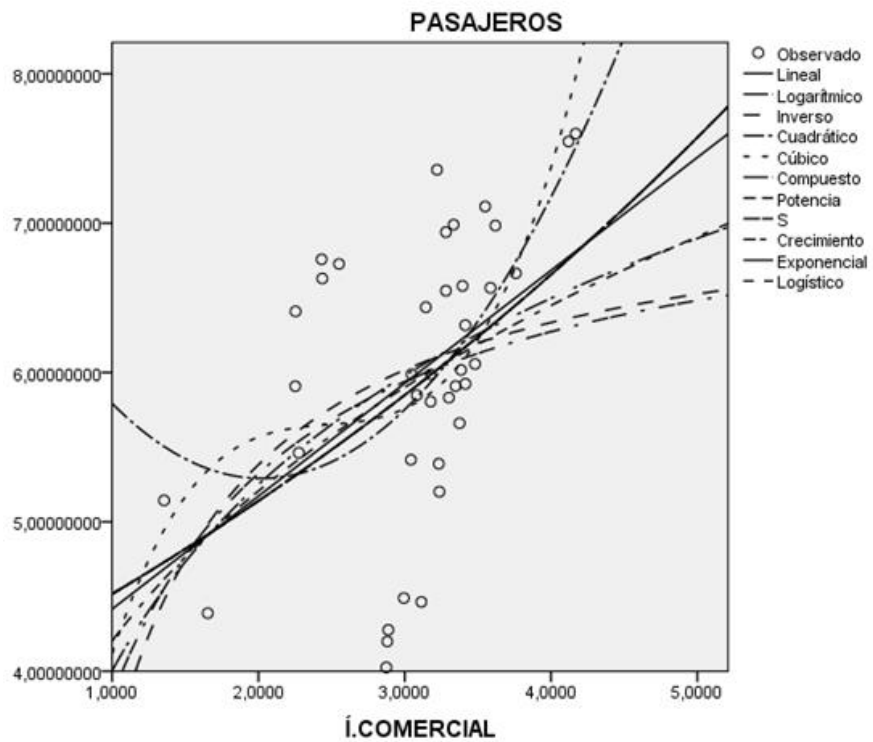
	MERCADO POTENCIAL	CUOTA DE MERCADO	COMPAÑÍAS	DESTINOS
Ecuación	R cuadrado	R cuadrado	R cuadrado	R cuadrado
Lineal	0,086	0,206	0,826	0,836
Logarítmica	0,057	0,172	X	X
Inversa	0,036	0,139	X	X
Cuadrático	0,203	0,279	0,828	0,862
Cúbico	0,233	0,329	0,828	0,869
Compuesto	0,077	0,192	0,793	0,797
Potencia	0,05	0,161	X	X
S	0,032	0,131	X	X
Crecimiento	0,077	0,192	0,793	0,797
Exponencial	0,077	0,192	0,793	0,797
Logística	0,077	0,192	0,793	0,797

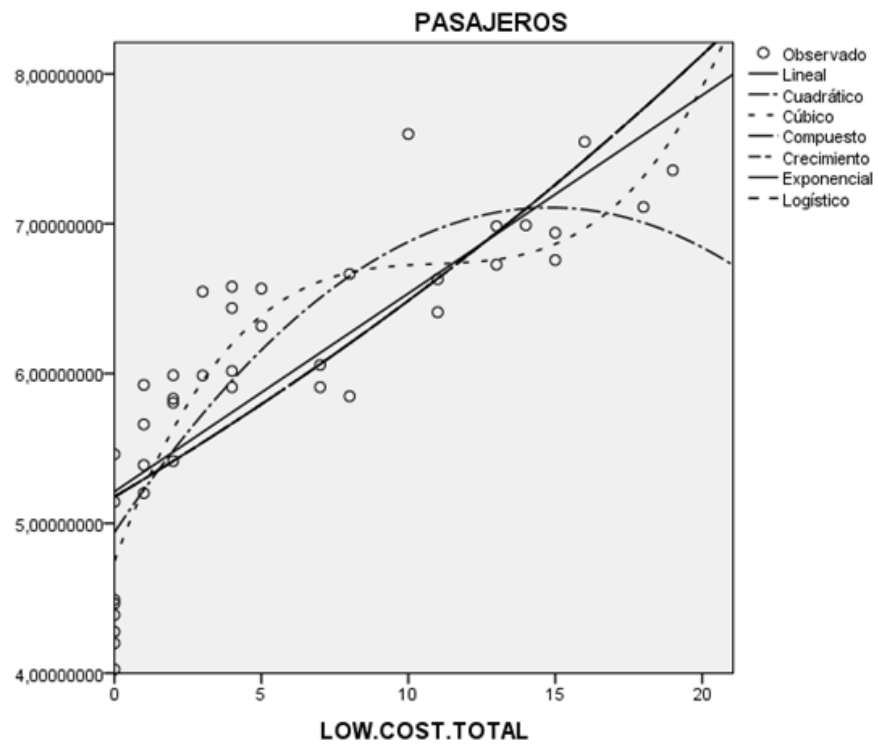
	DESTINOS.CERCA	KMS.DESTINOS	RENTA.PAÍSES	LOW.COST.CINCO
Ecuación	R cuadrado	R cuadrado	R cuadrado	R cuadrado
Lineal	0,219	0,411	0,341	0,756
Logarítmica	X	0,392	0,35	X
Inversa	X	0,368	0,358	X
Cuadrático	0,23	0,419	0,52	0,779
Cúbico	0,252	0,419	0,52	0,822
Compuesto	0,182	0,352	0,357	0,732
Potencia	X	0,334	0,366	X
S	X	0,311	0,375	X
Crecimiento	0,182	0,352	0,357	0,732
Exponencial	0,182	0,352	0,357	0,732
Logística	0,182	0,352	0,357	0,732

	LOW.COST.TOTAL	HUBS	POBLACIÓN DISUASORIA
Ecuación	R cuadrado	R cuadrado	R cuadrado
Lineal	0,666	0,25	0,331
Logarítmica	X	X	0,313
Inversa	X	X	0,29
Cuadrático	0,753	0,25	0,359
Cúbico	0,809	0,25	0,361
Compuesto	0,615	0,202	0,326
Potencia	X	X	0,309
S	X	X	0,289
Crecimiento	0,615	0,202	0,326
Exponencial	0,615	0,202	0,326
Logística	0,615	0,202	0,326

GRÁFICOS DE EJEMPLO







6. CONCLUSIONES

De ahora en adelante se tratará de analizar, en la manera de lo posible, los resultados obtenidos con el software SPSS; hay que resaltar que las conclusiones venideras estarán argumentadas en su mayoría, pero en ciertas ocasiones, algunas pueden ser de interpretación subjetiva.

6.1 CORRELACIONES

Trabajando con el programa estadístico a lo largo del proyecto, los primeros resultados en los que nos fijamos son las correlaciones, pues nos dan una primera idea de cuál es el grado de relación entre variables. Esto supone una gran ayuda a la hora de introducir o descartar variables en el modelo, pues el coeficiente de determinación nos señala el grado de relación de una variable independiente respecto a la dependiente (pasajeros), pero no la relación que guardan entre ellas.

Es de certeza absoluta que las siguientes variables que se exponen tienen un grado de correlación muy alto, pues todas ellas superan entre sí un 0,9, y además esto se cumple sistemáticamente para todos los modelos desde el año 2009 al 2013. La siguiente tabla muestra las correlaciones mayores de 0,9 con una “X”.

	PASAJEROS	POBLACIÓN	POBLACIÓN UE RESIDENTE	RENTA	PARO	Í.COMERCIAL	Í.RESTAURA CIÓN	Í.TURÍSTICO	Í.ECONÓMIC O	Í.INDUSTRIAL
PASAJEROS										
POBLACIÓN										
POBLACIÓN UE RESIDENTE										
RENTA										
PARO										
Í.COMERCIAL										
Í.RESTAURA CIÓN										
Í.TURÍSTICO			X							
Í.ECONÓMIC O		X				X	X			
Í.INDUSTRIAL		X				X	X		X	
PERNOCTACI ONES			X					X		
IMPORTACIO NES									X	X
MERCADO POTENCIAL		X				X	X		X	X

Tabla 58. Correlaciones mayores de 0,9

Como se puede observar la variable población y los índices comercial, de restauración y bares y económico están fuertemente correlacionados entre ellos y a su vez con el mercado potencial.

Por otra parte, también tienen una gran correlación las pernoctaciones, el índice turístico y la población residente de la UE. Por último, y siendo lógico, el índice industrial explica la variable importaciones y viceversa.

Todo esto será importante más adelante para ver qué variables tener en cuenta o no en los resúmenes de los diversos modelos. Como ya se ha comentado esto se cumple casi de manera pulcra para los 5 años de estudio, lo que nos da una primera idea de que los modelos no variarán demasiado dependiendo del año; lo que es una gran noticia a la hora de sacar conclusiones y de tomar dichos modelos como fiables.

Como se puede observar en la tabla de arriba, no aparece ninguna de las llamadas variables aeroportuarias, esto es debido a que ninguna de ellas tiene una fuerte correlación con el resto (tomando más de 0,9 como referencia) para el único año de estudio en su caso (2013).

6.2 COEFICIENTES Y DEMÁS ESTADÍSTICOS

Otras de las tablas de gran importancia incluidas en los resultados de cada modelo son los coeficientes no estandarizados y tipificados, ya que con estos podríamos construir la ecuación de regresión y ver cómo influye cada término en esta. Además puede ser de interés para el lector observar qué elasticidades son positivas o negativas. En dichas tablas también se muestra el valor de las pruebas “t” (que debe ser mayor de 2,5 para que la variable sea significativa); y en el bloque de resultados de más estadísticos para el año 2013 también encontramos los niveles críticos “Sig” (que deberán tener valores menores de 0,05 para que la variable sea significativa; “sig” es más conocido como p-value).

En cuanto al análisis de los residuos para cada uno de los modelos mostrados, el estadístico en el que nos fijamos es el de Durbin-Watson, y se puede concluir que en todos los modelos existe independencia entre los residuos ya que DW toma valores entre 1,5 y 2,5.

Además, en el bloque de “más estadísticos” encontramos el gráfico de pronósticos tipificados frente a residuos tipificados, donde se observa que la nube de puntos no sigue ninguna pauta de asociación clara, ni lineal ni de otro tipo, por lo que podemos afirmar que sus varianzas son homogéneas; y lo mismo ocurre para el resto de años aunque no se hayan incluido las gráficas en dicho bloque.

En ese bloque también se muestra el histograma del modelo de 2013, donde la distribución de los residuos parece seguir el modelo de probabilidad normal, ya que aproximadamente se acumulan los mismos casos de los que existen en una curva normal; además, la distribución es más o menos simétrica. Para el resto de modelos de diferentes años ocurre algo similar, por lo que los resultados son aceptados.

El otro gráfico que se muestra en el bloque es el gráfico de probabilidad normal, donde en el eje de abscisas está representada la probabilidad acumulada que corresponde a cada residuo tipificado, y el de ordenadas representa la probabilidad acumulada teórica que corresponde a cada puntuación típica en una curva normal con media 0 y desviación típica 1. De forma aproximada, la nube de puntos se encuentra alineada sobre la diagonal del gráfico, por lo que se puede afirmar que los residuos se distribuyen normalmente.

Cabe destacar que estos gráficos podrían estar más ajustados dependiendo de la tolerancia que deseemos darle a los residuos, y por tanto se debería entrar a valorar los problemas asociados que estos tienen y que ya se han visto en la parte de explicación estadística. Por último, añadir que se adjunta en ese bloque los diagnósticos de colinealidad y las correlaciones parcial y semiparcial del año 2013 a modo informativo, pero no se entrará de nuevo en un análisis de los resultados, pues ya se hizo en la explicación estadística general.

6.3 MODELOS TOTALES

En cuanto al coeficiente de determinación (R^2 corregido), la siguiente tabla nos muestra cada uno de ellos para los distintos años (modelos 1 al 5) en función del método de selección de variables. Los valores se muestran en tanto por cien.

También se adjunta un gráfico en el que se puede observar esta evolución.

MÉTODO	MODELO 1 2013	MODELO 2 2012	MODELO 3 2011	MODELO 4 2010	MODELO 5 2009
INTRODUCIR	86	82,6	83,3	81,8	80,8
ATRÁS	86,8	84,7	85,2	82,5	83,2
ADELANTE	71,5	65,9	68,5	67,9	81,7

Tabla 59. R2 corregido por modelos y métodos de selección

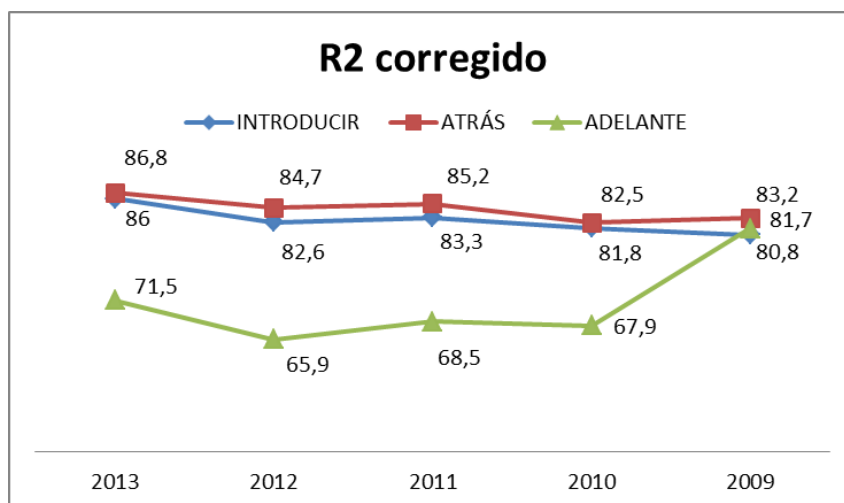


Gráfico 25. Evolución del R2 corregido

Esto nos muestra que nuestros modelos son muy parecidos a lo largo del tiempo en cuanto a explicación de la variable dependiente se refiere; ya que sus valores de R^2 corregido a lo largo de los 5 años son muy próximos para cada método de selección de variable. La única desviación que se produce es en el año 2009 para el método “adelante”, donde se observa un aumento considerable del R^2 corregido llegando a las cotas de otros métodos.

Los R^2 corregidos para nuestros modelos son muy positivos, teniendo una media de 83,7 en los métodos “introducir” y “atrás”, y de 71,1 en el método “adelante”; lo cual dice que nuestros modelos se explican con una alta fiabilidad.

La otra cara de la moneda al hablar del R^2 corregido es la elección de uno u otro método para saber cuáles serán las variables con las que trabajaremos. Esto es posiblemente lo más subjetivo del presente proyecto fin de carrera, pues no hay unas reglas o estándares que definan cuál debe ser la elección, sino que dependerá de la facilidad que se tenga para obtener los datos de las variables. En el siguiente cuadro se muestra un resumen de las variables escogidas por el método “adelante”, que en nuestra modesta opinión es el más adecuado, ya que aunque su coeficiente de determinación es algo inferior al resto

de métodos, tiene en cuenta un número mucho menor de variables, lo que lo hace idóneo para nuestro análisis. En cualquier caso el lector puede encontrar las variables incluidas en cada método de cada modelo en los extensos resultados expuestos en el bloque anterior, y hacer su propia elección.

MÉTODO ADELANTE	MODELO 1 2013	MODELO 2 2012		MODELO 3 2011
	Í. TURÍSTICO Í. ECONÓMICO Í. INDUSTRIAL	POBLACIÓN UE RESIDENTE Í. TURÍSTICO		Í. TURÍSTICO PERNOCTACIONES
	MODELO 4 2010		MODELO 5 2009	MODELO 6 2013 (23V)
	POBLACIÓN UE RESIDENTE Í. TURÍSTICO		RENTA Í. COMERCIAL Í. TURÍSTICO Í. ECONÓMICO Í. INDUSTRIAL	DESTINOS COMPAÑÍAS

Tabla 60. Variables escogidas para cada modelo con el método adelante

De forma general las conclusiones de este apartado pueden ser que las variables que tiene en cuenta cada modelo son las relacionadas con el turismo en todos los casos (bien sea solo índice turístico o con pernoctaciones y residentes de la UE); y a su vez en alguno de los modelos introduce otros índices de carácter económico e industrial.

La única excepción a esto se observa para el modelo 6 del año 2013, en el que se introdujeron las variables aeroportuarias, y se escogen como predictoras destinos y compañías.

Aunque ya se comentó en la parte explicativa de las variables, es muy importante destacar que este modelo número 6 no debe ser tomado con la misma confianza que los que solo incluyen variables socioeconómicas, pues las propias variables aeroportuarias están muy relacionadas con la variable dependiente “pasajeros” en la propia obtención de datos por aeropuerto.

Esto se ve representado en los altos coeficientes de determinación obtenidos y que se exponen a continuación, en comparación con el modelo 1 donde solo se incluyen las variables socioeconómicas. Debajo se adjunta un gráfico aclaratorio.

MÉTODO	MODELO 1 2013 con 15 variables	MODELO 6* 2013 con 23 variables
INTRODUCIR	86	94
ATRÁS	86,8	94,7
ADELANTE	71,5	87,4

Tabla 61. R2 corregido de los modelos 1 y 6 del año 2013

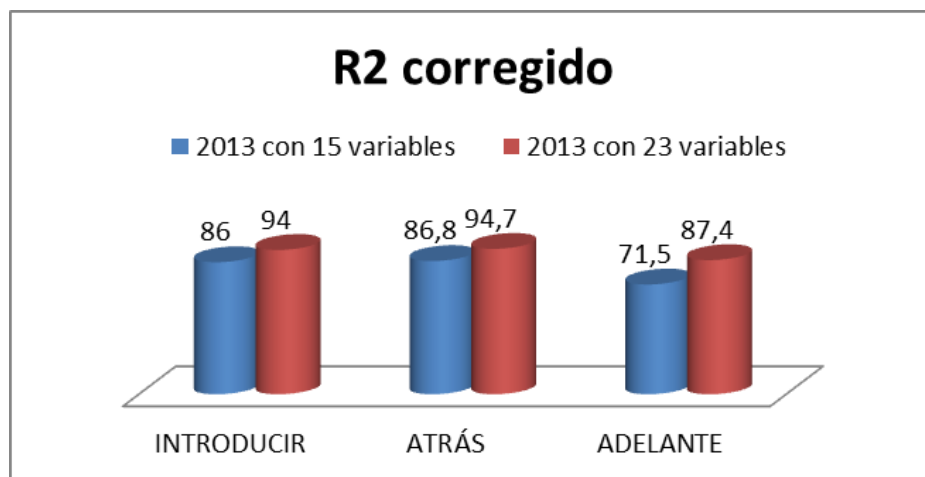


Gráfico 26. Diferencia de los R2 corregidos de los modelos 1 y 6 el año 2013

Otra de las variables a analizar con cautela es “la población en los entornos aeroportuarios por distancias disuasorias”, donde se hace una comparativa para el año 2009 (del que se tienen datos), introduciendo o no dicha variable. Los resultados se muestran a continuación.

MÉTODO	MODELO 5 2009 con 15 variables	MODELO 5* 2009 con 16 variables (Área Disuasoria)
INTRODUCIR	80,8	82,5
ATRÁS	83,2	84,5
ADELANTE	81,7	84,1

Tabla 62. R2 corregido del modelo 5 con y sin la variable área disuasoria (año 2009)

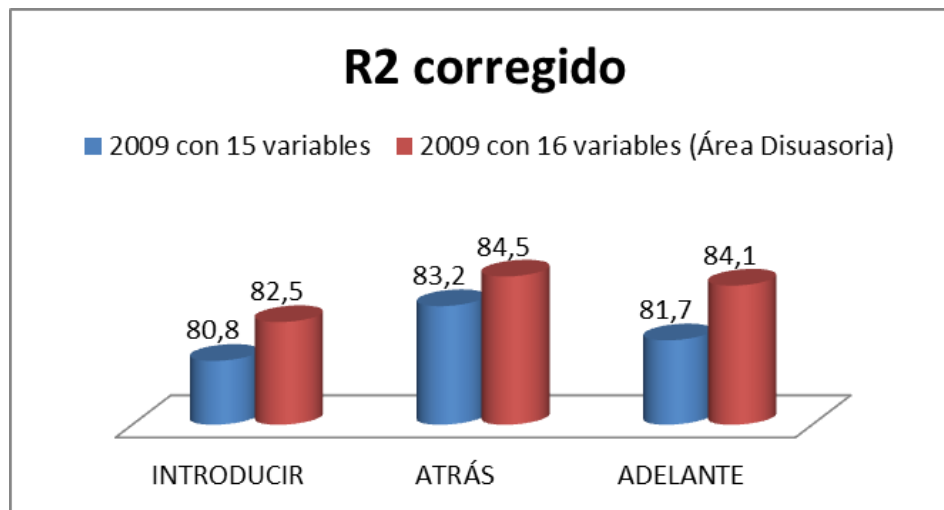


Gráfico 27. Diferencia de los R2 corregidos del modelo 5 con y sin la variable área disuasoria (año 2009)

Se puede apreciar que los coeficientes de determinación apenas varían, y aunque si son un poco mejores introduciendo esta variable en el modelo, esta no se debe tener en cuenta debido a que la forma de obtención de los datos es laboriosa y la elección de los radios que delimitan el área de cada aeropuerto han sido elegidos de una forma muy subjetiva. Es decir, al igual que como se ha procedido, podrían elegirse otras distancias disuasorias, pero eso implica tener en cuenta de antemano cual es el tráfico de pasajeros de cada aeropuerto para aplicar una relación con su radio de acción.

Por todo esto y debido a que la variable “población” es prácticamente igual de predictora que la variable “población con área disuasoria”, podemos afirmar que aplicar los datos por provincias a cada aeropuerto (o islas en cada caso) es lo más adecuado, aprovechando los datos que se tienen para las propias fronteras estatales.

6.4 MODELO DE ÚNICA VARIABLE

Una vez analizados los resultados de todos los modelos completos, es el momento de profundizar en lo que ocurre en cada variable por separado respecto a la variable dependiente “pasajeros”. Los resultados detallados, al igual que con el resto de modelos y estadísticos, están en el anterior bloque.

El siguiente gráfico nos da una buena idea de qué variables son más o menos predictoras (según su R^2) en relación con la variable dependiente y de manera única.

R2 para cada modelo de única variable

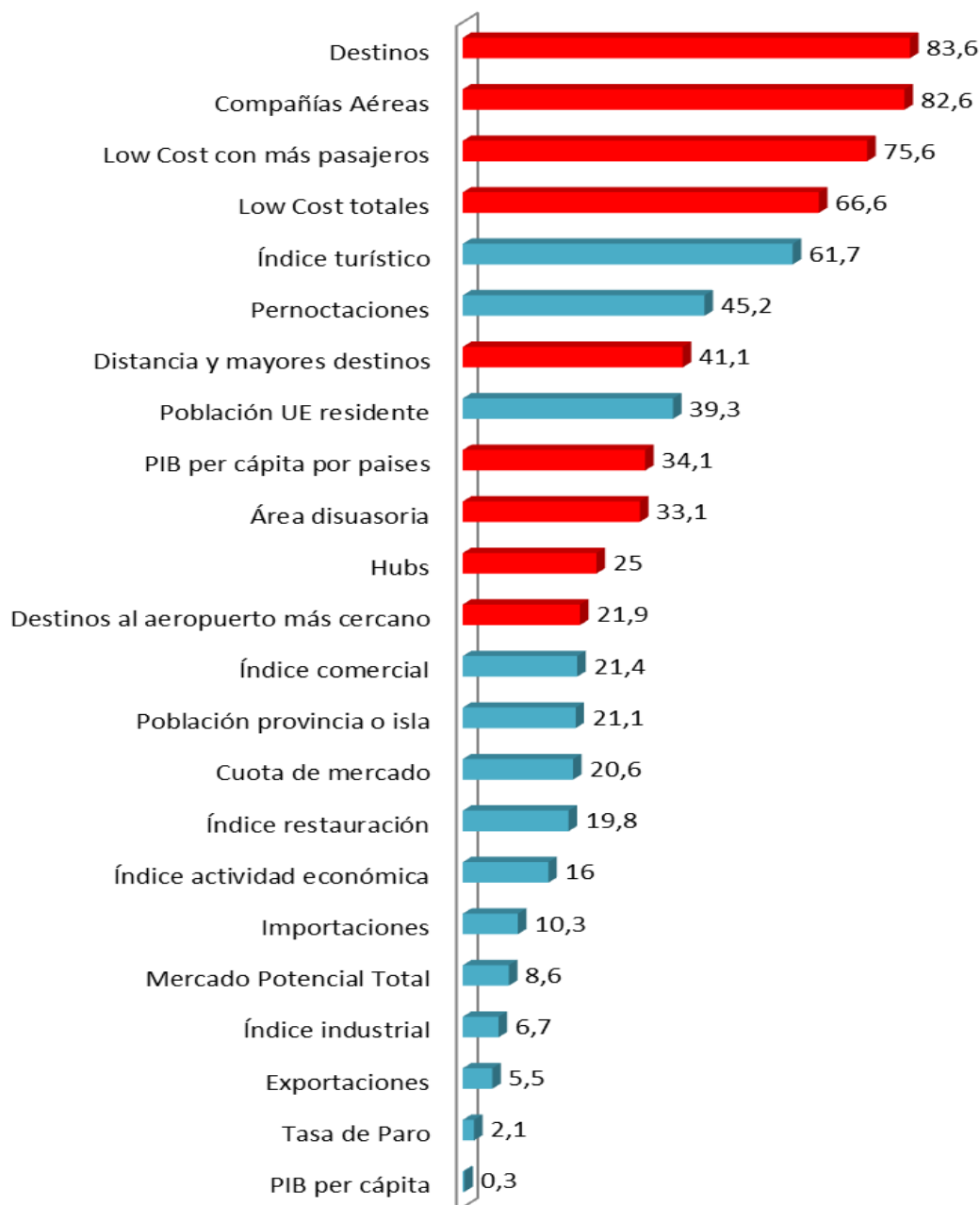


Gráfico 28. Orden de cada variable según el valor de su R2

De forma aclaratoria, se muestran las barras en color azul o rojo, según sea la variable socioeconómica o aeroportuaria respectivamente. Este gráfico es muy intuitivo y deja claro lo que antes se mencionó sobre la dudosa fiabilidad de usar las variables

aeroportuarias en un modelo, pues como se aprecia en rojo, estas tienen altos coeficientes de determinación al analizarlas por separado (todas están en la parte alta).

Lo cual quiere decir que guardan una gran relación con la variable dependiente “pasajeros”; esto a priori no debería ser algo negativo en los resultados, pero si lo es si tenemos en cuenta la propia esencia de la variable. Para explicarlo con un ejemplo, tomemos la variable “destinos”, de entrada no es una variable que directamente involucre el número de pasajeros dentro de ella, pero si lo hace de manera indirecta ya que es obvio que a mayor tráfico de pasajeros en un aeropuerto, mayores deben ser sus destinos y viceversa. Del mismo modo no podemos conocer si aumentando el número de destinos de un aeropuerto aumentará el número de pasajeros tanto como se espera, ya que eso depende de otros muchos factores involucrados que se escapan del objetivo de este proyecto.

Sin ánimo de ser exhaustivo, se van a mostrar y analizar solo algunos de los gráficos más representativos de los modelos de única variable. Ya que en estos esquemas no se muestran los ejes y sus unidades (si se muestran en el bloque anterior de resultados), hay que aclarar que en el eje de ordenadas se muestra el valor pronosticado del logaritmo de pasajeros, y en el eje de abscisas se muestra el logaritmo de pasajeros observados. La línea recta que aparece tiene por ecuación $x = y$, por lo que todo punto que se encuentre encima de esta, estaría en el equilibrio entre pasajeros pronosticados y observados.



Gráfico 29. Logaritmo de pasajeros pronosticados VS logaritmo de pasajeros observados (Residentes UE)

La conclusión a la que se llega con este esbozo es que la población residente de la UE en las provincias donde están los aeropuertos como Burgos, Badajoz, Logroño, Salamanca o León entre otros, no usan estos para desplazarse, ya que preferirán usar otros medios de transporte. Sin embargo los aeropuertos de las provincias o islas como Melilla, El Hierro, Fuerteventura, Lanzarote, Ibiza o Barcelona entre otros, si son usados por la población residente de la UE para sus desplazamientos. Esto tiene mucha lógica viendo que muchos de ellos son insulares y otros son aeropuertos de importancia relevante, por lo que el avión es su principal medio de transporte. De hecho al estar por debajo de la diagonal están bastante infraestimados.

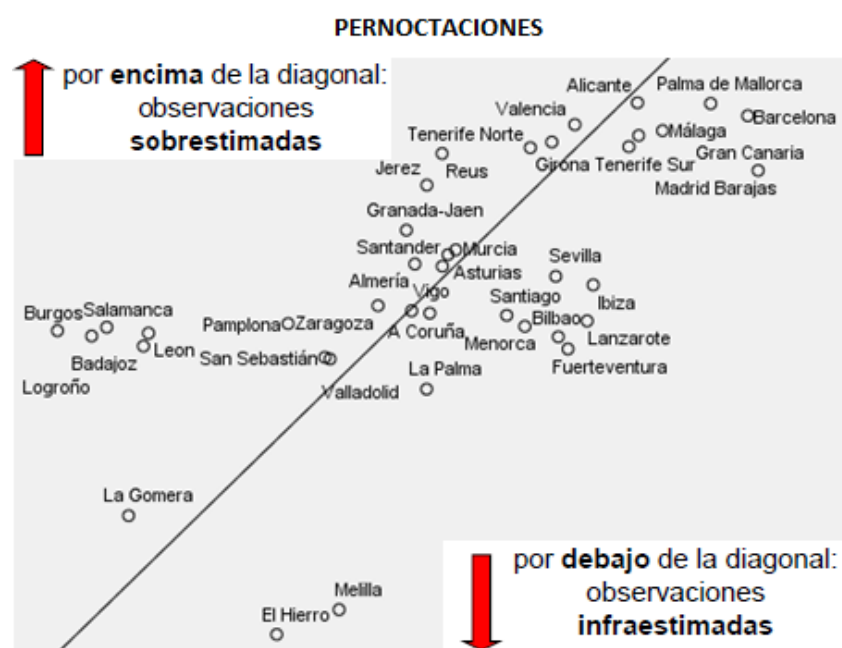


Gráfico 30. Logaritmo de pasajeros pronosticados VS logaritmo de pasajeros observados
(Pernoctaciones)

Al igual que en el ejemplo anterior se puede observar que la gente que pernocta (plazas hoteleras) en los aeropuertos de las provincias que se encuentran por encima de la diagonal no usan como medio de transporte el avión, sino otros; y del mismo modo pero a la inversa ocurre para los aeropuertos por debajo de la diagonal.

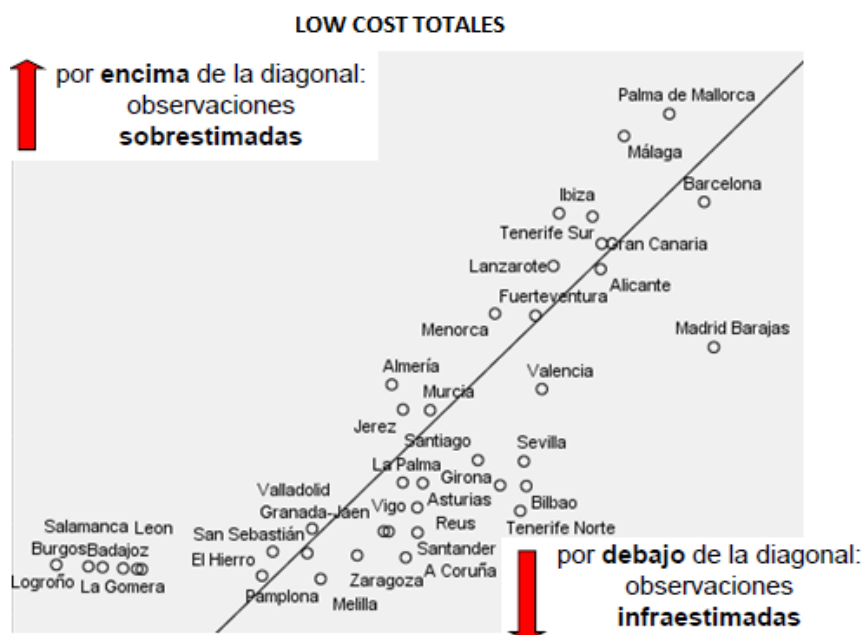


Gráfico 31. Logaritmo de pasajeros pronosticados VS logaritmo de pasajeros observados (Low Cost)

En el caso de las aerolíneas low cost que operan en cada aeropuerto ocurre algo parecido, pues es coherente pensar que los aeropuertos sobrestimados cubren más que de sobra la demanda de pasajeros con las aerolíneas low cost existentes; mientras que en los aeropuertos infraestimados el proceso se da a la inversa, pensando que podrían atraer a más compañías de bajo coste en función de la demanda.

Un caso curioso aquí es el del aeropuerto de Madrid Barajas, siendo este el más infraestimado de todos los del modelo; esto se debe a que su relación entre low cost y tráfico de pasajeros no es tan elevada como en otros aeropuertos. Aunque sería muy complejo valorar si debería haber más aerolíneas de este tipo, pues entra en juego a su vez que Madrid es un aeropuerto “hub”, por lo que algunas aerolíneas tradicionales absorben gran parte de la demanda.

6.5 MODELOS NO LINEALES

En la última parte del bloque de resultados se hace una comparación del coeficiente de determinación (R^2) para los distintos modelos no lineales que tiene implantado el software SPSS. En este fragmento solo se trata de conocer si nuestros modelos de única variable lineales (realmente son exponenciales pero se han linealizado aplicando logaritmos) podrían ser aproximados de manera más eficaz por modelos no lineales, y de ser así, por cuáles.

El procedimiento de Estimación Curvilínea de SPSS ha generado estadísticos de estimación curvilínea por regresión y gráficos relacionados para 11 modelos diferentes de estimación curvilínea por regresión. Se produce un modelo diferente para cada variable independiente respecto a la variable dependiente “pasajeros”.

En los resultados no se incluyeron los parámetros β_0 , β_1 , β_2 y β_3 que definen la ecuación cúbica para cada variable, pese a que el programa los adjunta con el resto de resultados, debido a que solo nos interesa conocer la diferencia explicativa entre el modelo lineal y el cúbico.

El siguiente gráfico resume los resultados anteriores, haciendo una comparación por cada variable entre los coeficientes de determinación para el modelo lineal y el modelo cúbico, ya que la ecuación cúbica es la que mayores R^2 obtiene por variable (excepto para “índice turístico” y “pernoctaciones” que se aproxima mejor con una ecuación cuadrática). Un modelo cúbico está definido por la ecuación $Y = \beta_0 + \beta_1X + \beta_2X^2 + \beta_3X^3$.

En el gráfico se observa que el aumento del coeficiente de determinación de los modelos cúbicos respecto a los lineales es mínimo en la mayoría de los casos, por lo que podemos afirmar que es suficiente y acertada la regresión lineal múltiple para el análisis de nuestros modelos de única variable, y por tanto también lo es para los modelos completos.

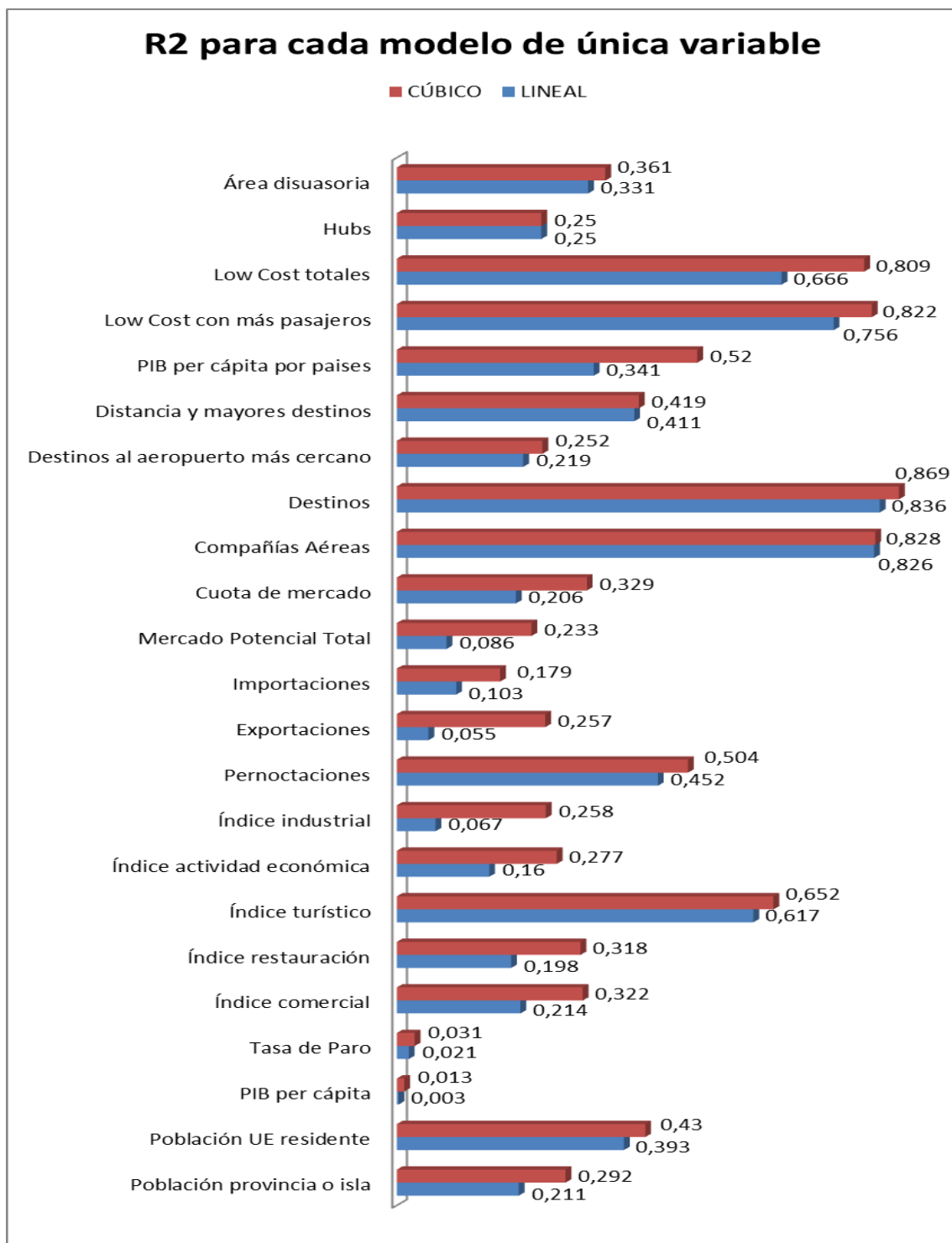


Gráfico 32. Comparación de R2 lineales y R2 cúbicos

6.6 AEROPUERTOS

Por último haremos un análisis por aeropuerto, así como por sectores territoriales; hay que resaltar que tratar de predecir lo que ocurre en cada aeropuerto es una tarea complicada, pues en este proyecto cada aeropuerto es un caso o grado de libertad, y de lo que se obtienen resultados son de las variables, tanto separadas como en conjunto formando modelos. Por lo tanto esta información debemos obtenerla de las gráficas, y precisamente de las de valor pronosticado frente al valor observado de los modelos completos. Las maneras de interpretar las gráficas pueden ser diversas, y una de las pensadas es la siguiente, dividimos cada gráfico de manera aproximada como se indica en el dibujo de abajo, y asignamos un color a cada aeropuerto en función de su posición (siempre de la forma más aproximada posible). El color amarillo significa que hay una fuerte sobrestimación o infraestimación, y aunque en el dibujo se ha colocado en las esquinas a modo de una mejor visualización, no quiere decir que un aeropuerto deba estar tan alejado de la diagonal para estar dentro de este grupo.

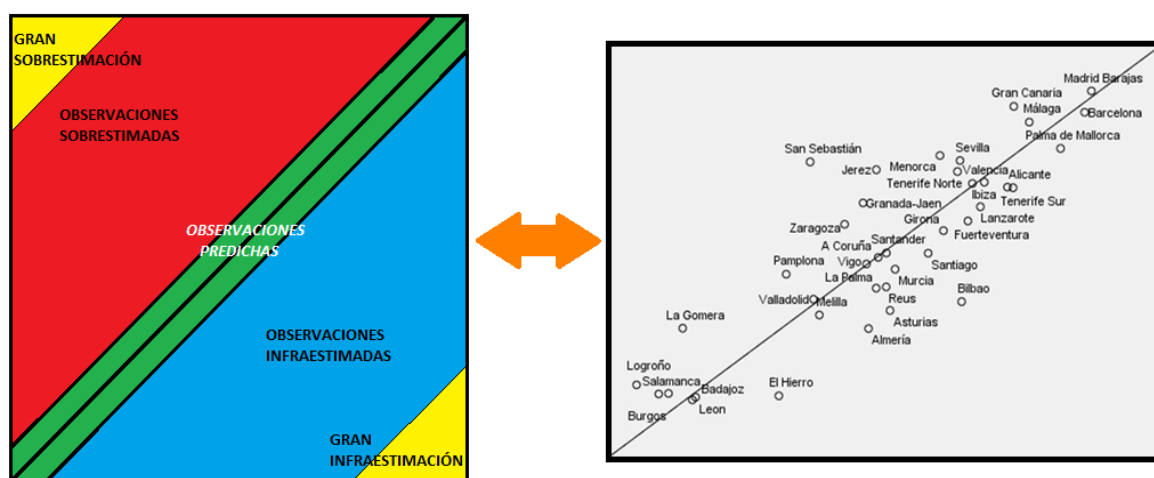


Gráfico 33. División por colores de un gráfico de pasajeros observados y pronosticados

Una vez se ha asignado uno de estos cuatro colores a cada aeropuerto de cada modelo, podemos hacernos una idea de la evolución de cada aeropuerto a lo largo de los años de estudio, analizando si entra o no dentro de unos márgenes estables de estimación y, por tanto, si los modelos pueden ser más o menos fiables para dicho aeropuerto. De este modo obtenemos el cuadro adjunto.

AEROPUERTO	MODELO 1	MODELO 2	MODELO 3	MODELO 4	MODELO 5	MODELO 6
	2013	2012	2011	2010	2009	2013*
A Coruña						
Alicante						
Almería						
Asturias						
Badajoz						
Barcelona						
Bilbao						
Burgos						
El Hierro						
Fuerteventura						
Girona						
Gran Canaria						
Granada-Jaén						
Ibiza						
Jerez						
La Gomera						
La Palma						
Lanzarote						
León						
Logroño						
Madrid Barajas						
Málaga						
Melilla						
Menorca						
Murcia						
Palma de Mallorca						
Pamplona						
Reus						
Salamanca						
San Sebastián						
Santander						
Santiago						
Sevilla						
Tenerife Norte						
Tenerife Sur						
Valencia						
Valladolid						
Vigo						
Zaragoza						

Tabla 63. Cuadro de colores por aeropuertos y modelos

Como se analizó en algún ejemplo con anterioridad, predecir lo que ocurre en los gráficos de única variable se puede hacer con relativa facilidad, pues al solo influir una variable se pueden sacar conclusiones respecto a un aeropuerto o región; otra cosa bien distinta es tratar de hacer lo mismo con los modelos completos, pues al haber tantas variables en acción es fácil caer en errores y establecer relaciones falaces. Además no es objeto de este proyecto sacar ese tipo de conclusiones socioeconómicas, ni establecer relaciones de demanda más allá de los propios resultados estadísticos. Dicho esto, lo más sensato es exponer qué aeropuertos cumplen mejor o peor los modelos, tratando de visualizarlo todo como un único modelo de 5 años con las variables socioeconómicas y otro modelo para el año 2013 con las variables socioeconómicas y aeroportuarias.

Los aeropuertos que tienen todos los modelos con muy buenos valores pronosticados son Barcelona, Girona, La Palma, León, Málaga, Palma de Mallorca, Santander, Tenerife Norte, Tenerife Sur, Valencia, Valladolid y Vigo.

Los aeropuertos en que prácticamente todos sus modelos están sobrestimados son Gran Canaria, Granada-Jaén, Logroño, Madrid, San Sebastián y Zaragoza.

Los aeropuertos con casi todos sus modelos infraestimados son Almería, Asturias, Fuerteventura, Ibiza y Lanzarote.

Los aeropuertos donde la mayoría de sus modelos tienen un gran error de estimación son Alicante, Bilbao, El Hierro, Melilla, Jerez y La Gomera; en el caso de los cuatro primeros hay una fuerte infraestimación por lo que estos aeropuertos absorben parte del tráfico de otros de su alrededor. El caso más claro es el de Bilbao, que recibe la mayoría de sus pasajeros de las provincias colindantes donde sus aeropuertos tienen menos destinos; mientras que por ejemplo para el aeropuerto de Melilla se dan otras circunstancias, como el lugar geográfico que ocupa alejado de la península. En contraposición tendríamos el aeropuerto de Jerez donde se da una gran sobrestimación, ya que aeropuertos de gran importancia como Málaga y Sevilla están muy próximos a este y por tanto sus pasajeros tienen opciones mejores a corta distancia.

Esto mismo ocurre en muchos de los aeropuertos incluidos en los modelos, aunque de una manera menos intuitiva, solo basta con fijarse en cuáles están sobrestimados o infraestimados respecto a sus más próximos (dentro o fuera de su sector); como ejemplo tenemos el aeropuerto de Santiago de Compostela que atrae parte del tráfico de La

Coruña (tiene casi el triple de pasajeros) pese a ser una ciudad de menor población e importancia; la causa principal de que esto ocurra, entre otros factores, es que Santiago disfruta de un mayor número de destinos y compañías.

En este aspecto, en las islas suele producirse un efecto dispar en función de la que analicemos, por lo que habría que ver qué variables afectan a cada una de ellas y tener en cuenta que las opciones de transporte son más limitadas.

El resto de aeropuertos, dependiendo del modelo, se comportan de una manera diferente sin seguir una pauta determinada a lo largo de los años. Hay que señalar que algunos aeropuertos se han comportado de manera sistemática en el límite de una gran sobrestimación o infraestimación, aunque no se han incluido en esos grupos; son los aeropuertos de Madrid, Gran Canaria, Almería y Asturias.

Algo curioso es lo que ocurre con los aeropuertos de Alicante, Bilbao, Granada-Jaén, Jerez, Lanzarote y Zaragoza donde en el modelo del año 2013 con 15 variables están sobrestimados, infraestimados o en el margen de error, y en el modelo de 2013 de 23 variables están dentro de las observaciones predichas. Esto se debe a que las variables aeroportuarias para estos aeropuertos equilibran y ajustan la predicción de la demanda, por lo que es mejor tener en cuenta los dos tipos de variables para obtener un buen coeficiente de determinación en ellos.

Al ser muy difícil y arriesgado sacar conclusiones aeropuerto por aeropuerto, vamos a agruparlos por sectores geográficos y tratar de poner puntos en común.

6.7 SECTORES AEROPORTUARIOS

La manera de proceder es hacer una serie de subdivisiones territoriales englobando a cada aeropuerto en una de las 9 siguientes (en el mapa de España que sigue se puede visualizar de forma más clara):

- 1- Sector Noroeste: A Coruña, Santiago y Vigo.
- 2- Sector Norte: Asturias, Bilbao, Pamplona, San Sebastián y Santander.
- 3- Sector Noreste: Barcelona, Girona, Reus y Zaragoza.
- 4- Sector Centro-Norte: Logroño, Burgos, León, Salamanca y Valladolid.
- 5- Sector Centro: Madrid.
- 6- Sector Este: Alicante, Murcia y Valencia.

- 7- Sector Sur: Almería, Badajoz, Granada-Jaén, Jerez, Málaga, Melilla y Sevilla.
- 8- Sector Baleares: Ibiza, Menorca y Palma de Mallorca.
- 9- Sector Canarias: El Hierro, Fuerteventura, Gran Canaria, La Gomera, La Palma, Lanzarote, Tenerife Norte y Tenerife Sur.



Figura 23. División por sectores aeroportuarios

La forma de proceder de este método es asignar a cada aeropuerto un número para ordenarlos dentro del grupo al que pertenecen, después contar para cada grupo el número de casos que tiene (casos predichos, sobrestimados, infraestimados y erróneos) y con esos datos obtener el siguiente gráfico que hace más cómoda la extracción de conclusiones finales.

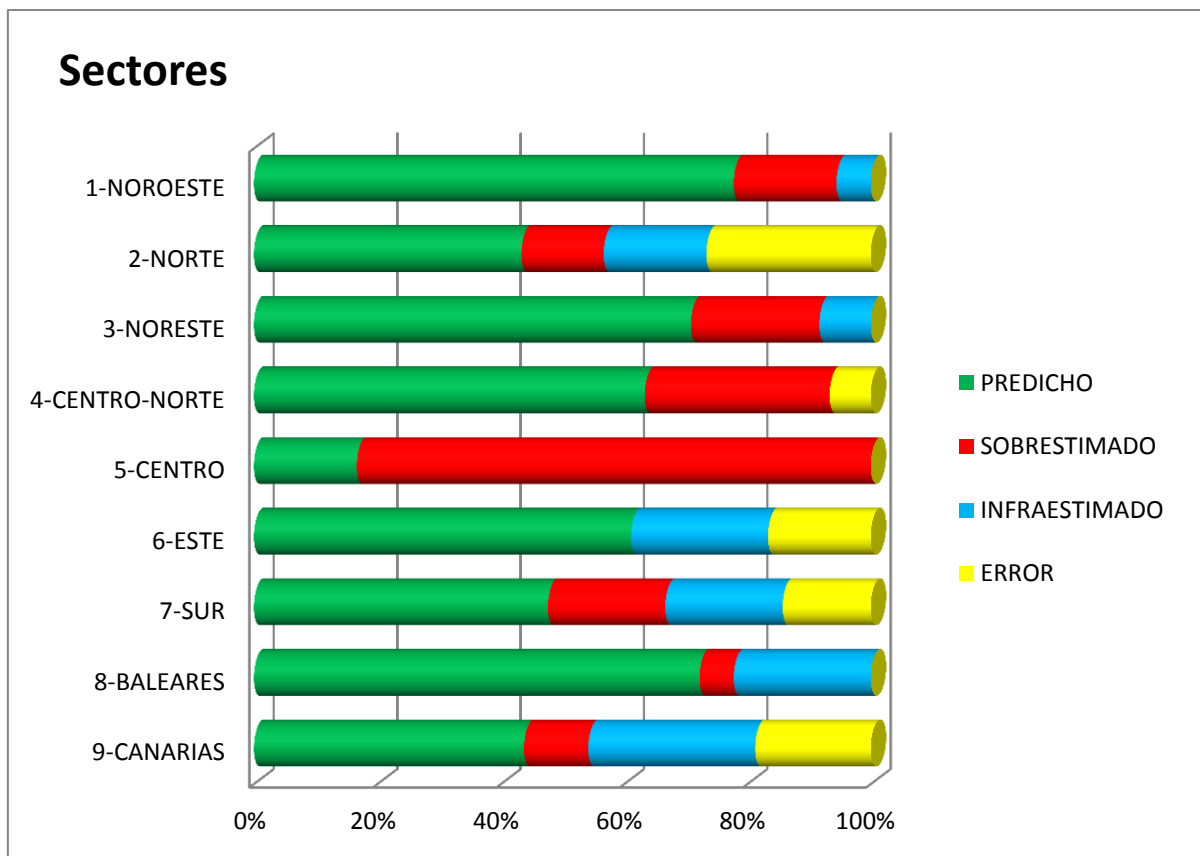


Gráfico 34. Resumen por sectores

Se debe recordar que este gráfico es la suma de los resultados de los 6 modelos estudiados, a modo de analizar el comportamiento de estos de forma conjunta, tanto en las variables como en el tiempo.

Los sectores donde más se ajustan los modelos son el Noroeste, Noreste y Balear donde se ronda o supera el 70% de los casos predichos, y les siguen los sectores Centro-Norte y Este con un 60% de casos predichos. El sector Centro es el que peores resultados obtiene, pero al solo incluir el aeropuerto de Madrid Barajas puede no ser representativo. Esto tiende a parecer algo contradictorio, ya que Madrid tiene el mayor tráfico de pasajeros del país y además es un “hub”, por lo que absorbe parte de los pasajeros de otros aeropuertos, y sus resultados no deberían estar sobrestimados; pero la solución puede ser que algunas de las variables estén de por si sobrestimadas por ser Madrid la capital del país y centro de diversas actividades.

Además, cambiando los casos de error (color amarillo), por su correspondiente sobrestimación o infraestimación, tenemos que el sector Centro-Norte es el más sobrestimado, por lo que estaría trabajando por debajo de su demanda real; esto se debe

a que la población elige los aeropuertos de otros sectores para desplazarse. Mientras que el sector más infraestimado es el Este, lo cual concluye que está trabajando por encima de su nivel real de demanda. Los motivos de que esté atrayendo a pasajeros de otros sectores pueden ser diversos, aunque muy posiblemente el turismo es el más importante.

Por otro lado, la falta de destinos y frecuencias en los aeropuertos pequeños, así como sus elevados precios y falta de alternativas low cost están haciendo que los clientes recorran distancias cada vez más largas para coger un vuelo en uno de los principales aeropuertos, especialmente en aquellos que se ven favorecidos por su situación geográfica y buenas conexiones.

A la vista de todos los datos expuestos hasta ahora, se propone la progresión hacia un sistema futuro basado en los principios del modelo hub & spoke, lo que permitiría un tratamiento ágil de los flujos de tráfico. Los aeropuertos de Madrid, Barcelona y Málaga podrían establecerse como los tres grandes hubs españoles y el resto crearían una red de aeropuertos de menor entidad encargados de aportar el pasaje.

7. BIBLIOGRAFÍA

Recursos Web:

- <http://www.aena.es>
- <http://www.ine.es>
- <http://www.anuarieco.lacaixa.comunicacions.com>
- <http://www.mineco.gob.es>
- <http://www.ibm.com/software/es/analytics/spss>
- <http://www.iet.tourspain.es>
- <http://www.expansion.com>
- <http://www.iata.org>
- <http://www.bravofly.es>
- <http://www.atag.org>
- <http://www.fomento.gob.es>
- <http://www.minetur.gob.es>
- <http://www.oag.com/Insight>
- <http://www.ryanair.com>
- <http://www.iberia.com>
- <http://www.aci.aero>
- <http://www.fedea.net>
- <http://ec.europa.eu/eurostat>
- <http://datos.bancomundial.org>

Libros, publicaciones y tesis doctorales consultadas:

- Apuntes de Explotación del Transporte Aéreo de 5º curso de Ingeniería Aeronáutica de la ESI de la Universidad de Sevilla. Profesor: Jose María del Castillo Granados. (2014).
- NAVIDI, WILLIAM. (2006). Estadística para ingenieros y científicos. McGraw-Hill.
- Memorias corporativas de AENA para los años 2009, 2010, 2011, 2012 y 2013.
- Manual IBM SPSS Statistics 20.
- BEL, G., FAGEDA, X. (2006). Aeropuertos y globalización. Opciones de gestión aeroportuaria e implicaciones sobre el territorio.

- BENÍTEZ ROCHEL, J. J. (2000). Factores determinantes de la demanda de transporte aéreo y modelos de previsión.
- ADIEGO, P. (2004). El proceso de urbanización y los sistemas urbanos españoles.
- APPILD, S. J., KASARDA, J. D. (2006). The appropriate scale of US airport retail activities. *Journal of Air Transport Management*, Vol. 12.
- BURGHOUWT, G., VELDHUIS, J. (2006). The Competitive Position of Hub Airports in the Transatlantic Market. *Journal of Air Transportation*, Vol. 11.
- CAMAGNI, R. (1992). Economía urbana.
- GÁMIR, A., RAMOS, D. (2002) Transporte aéreo y territorio. Barcelona, Ariel.
- KRAUS, A., KOCH, B. (2006). The Opportunities and Threats of Turning Airports into Hubs.
- LÖSH, A. (1954). The Economics of Location. Yale University Press. New Haven.
- PITA, L. A. (2003). Airport connections of high-speed lines. Ediciones UPC.
- PRED, A. (1977). City-systems in advanced economies. London. Hutchinson.
- PUEBLA, J. G. (1998) Redes, espacio y tiempo. *Anales de Geografía de la Universidad Complutense*.
- VILLA, A. R. (2008). Los aeropuertos en el sistema de transporte. Fundación Aena.
- TAPIADOR, F. J., MATEOS, A. (2008). The geographical efficiency of Spain's regional airports: A quantitative analysis. *Journal of Air Transport Management*, Vol. 14.
- BOIX DOMÉNECH, RAFAEL. (2003). Redes de ciudades y externalidades. Universitat Autònoma de Barcelona.
- HORMIGO VENTURA, JUAN P. (2006). La evolución de los factores de localización de actividades. Universitat Politècnica de Catalunya.
- SERRANO MARUGÁN, JOSE M. (2007). Análisis del impacto en los mercados de transporte de la aparición de un operador aéreo de bajo coste y razones de elección de los mercados. Universidad Pontificia de Comillas (Madrid).
- VELASCO, CARLOS. (2008). Análisis de regresión múltiple. *Econometría*. Universidad Carlos III de Madrid.