

## **Análisis estadístico de la realidad socioeconómica en Andalucía. Una aproximación a escala municipal**

Eugenia M.<sup>a</sup> Cruces Pastor \*, Julia de Haro García \*\* y M.<sup>a</sup> Dolores Sarrión Gavilán \*\*\*

**RESUMEN:** En este trabajo aplicamos las técnicas multivariantes de Análisis Factorial y Análisis Cluster a un amplio espectro de indicadores relativos a los entornos físico, demográfico, económico y social a nivel municipal, con la finalidad de extraer un número reducido de variables que reflejen los distintos aspectos latentes de la realidad socioeconómica de las localidades andaluzas. Ello permitirá analizar la existencia de diferencias socioeconómicas significativas entre los municipios andaluces y la posterior clasificación de los mismos en grupos relativamente homogéneos que puedan servir como base para la elaboración de políticas de desarrollo en el ámbito municipal.

**Clasificación JEL:** O18, R11, R12.

**Palabras clave:** análisis factorial, análisis cluster, municipios, socioeconómico, Andalucía.

### **Statistic analysis of the socioeconomic context in Andalusia. An approach at municipal level**

**ABSTRACT:** In this work we apply the multivariate techniques of Factor Analysis and Cluster Analysis to a wide spectrum of indicators of the physical, demographic, economic and social environment at the municipal level, with the aim of extracting a small number of variables that reflect different aspects latent in the socio-economic reality of the towns of Andalusia. This allows an analysis of the existence of significant socio-economic differences between the municipalities of Andalusia and the subsequent classification of them in relatively homogeneous groups which can serve as a basis for policy development at the municipal level.

---

\* Autor para correspondencia. Facultad de Comercio y Gestión. Ampliación del Campus de Teatinos, s/n. Universidad de Málaga. 29071. Málaga. Teléfono: 951952039-951952014. E-mail: [cruces@uma.es](mailto:cruces@uma.es).

\*\* Universidad de Málaga. E-mail: [haro@uma.es](mailto:haro@uma.es).

\*\*\* Universidad de Málaga. E-mail: [dsarrion@uma.es](mailto:dsarrion@uma.es).

*Recibido: 25 de septiembre de 2008 / Aceptado: 8 de junio de 2009.*

**JEL Classification:** O18, R11, R12.

**Keywords:** Factor Analysis, Cluster Analysis, municipalities, socio-economic, Andalusia.

## 1. Introducción

Las diferencias socioeconómicas entre las unidades territoriales que conforman un país, autonomía, provincia, etc., son de gran interés tanto para los economistas como para los políticos. Como se señala en Cano y otros (2003), numerosas instituciones nacionales e internacionales elaboran sistemas de indicadores sociales referidos a diversas unidades territoriales y periodos históricos. Entre los informes más relevantes se encuentran los del PNUD (1999), OCDE (2002), Banco Mundial (1992), Proyectos Urban Audit<sup>1</sup> de la Unión Europea y, en España, el INE (1999 y siguientes). El análisis de estos indicadores se puede utilizar como base para elaborar políticas de desarrollo a distintos niveles de desagregación (regional, provincial, municipal,...) o para cuantificar el bienestar social.

Un análisis en profundidad de la realidad socioeconómica de una región geográfica debe estar basado en un amplio espectro de indicadores de desarrollo socioeconómico y debe tener en cuenta las unidades geográficas más pequeñas para las que se posea información, ya que la proximidad geográfica entre localidades no significa necesariamente proximidad en cuanto a nivel socioeconómico (Soares y otros, 2003). Como se indica en Rúa y otros, 2003, el primer ámbito en el que pueden llevarse a cabo políticas que ayuden a paliar los desequilibrios socioeconómicos es el municipio. En el presente trabajo nos centramos en la Comunidad Autónoma de Andalucía, la región más poblada de España con 8.177.805 de habitantes en 2008 (el 17,82% de la población española) y la segunda en extensión, ocupando el 17,2% de la superficie del territorio nacional. Nuestro objetivo es tratar de establecer, mediante las técnicas estadísticas multivariantes de Análisis Factorial y Análisis *Cluster* o de Conglomerados, si existen diferencias socioeconómicas significativas entre los municipios andaluces y clasificarlos en grupos relativamente homogéneos.

El Análisis *Cluster* o de Conglomerados es el enfoque estadístico estándar para el análisis de las disparidades entre territorios. Sin embargo, cuando se utiliza un número elevado de indicadores socioeconómicos, el análisis de conglomerados se suele combinar con el Análisis Factorial [véase Everitt (1993)]. Mediante esta técnica se resume la información relativa a una amplia gama de variables observadas en unos pocos factores no observables que representan los distintos aspectos de la realidad objeto de estudio. Los factores extraídos se utilizan para, mediante el Análisis de Conglomerados, llevar a cabo la agrupación de municipios, de modo que los grupos obtenidos sean homogéneos en cuanto al comportamiento de los factores que reflejan dicha realidad.

---

<sup>1</sup> En la actualidad se trabaja en el Proyecto Urban Audit III.

En los últimos años estos métodos han sido muy utilizados para caracterizar y tipificar territorios. Como antecedente más inmediato de este trabajo, por estar referido también a Andalucía, cabe mencionar el trabajo realizado por Gutiérrez Jáimez y otros (2000). En el mismo se aprovecha la información que a nivel municipal suministró la base de datos SIMA95 para, mediante el empleo de las técnicas de análisis multivariante antes mencionadas, proponer una tipología de Comarcas Andaluzas. Otros ejemplos de trabajos en los que se emplean estas técnicas estadísticas para caracterizar o agrupar unidades territoriales atendiendo a los valores de un conjunto más o menos numeroso de indicadores socioeconómicos son los realizados por López Ruiz (2005), en el que se analizan las desigualdades económicas y sociales a nivel municipal para la provincia de Almería; Vivo Molina y otros (2004), que realizan un estudio similar para la región de Murcia; el monográfico realizado por Herrero Prieto y otros (2007), relativo a las disparidades territoriales en Castilla y León; el correspondiente a los municipios de la Comunidad de Madrid realizado por Del Castillo Cuervo-Arango y otros (2006) o el que llevan a cabo Pena López y Sánchez Santos (2008), quienes se interesan por la realidad socioeconómica gallega desde una perspectiva municipal. Fuera de nuestras fronteras la combinación del Análisis Factorial y el Análisis *Cluster* se ha aplicado para la identificación de grupos de municipios, comarcas u otras unidades territoriales en Australia (Stimmson y otros, 2001), Estados Unidos (Hill y otros, 1998) o China (Hon y otros, 2005).

Lo que resta de este trabajo está organizado del siguiente modo: en la sección 2 se definen las variables en las que basaremos el estudio y se resumen los resultados del análisis descriptivo previo a nivel unidimensional; en la sección 3 se presentan los métodos estadísticos utilizados en esta investigación: Análisis Factorial y Análisis *Cluster*, más desde el punto de vista del significado de la terminología empleada a lo largo del trabajo que desde el de su fundamento estadístico o matemático, ya que se trata de técnicas clásicas de análisis multivariante cuya descripción formal puede encontrarse en cualquier texto especializado [véanse, por ejemplo, los textos de Cuadras (1996); Hair y otros (2000); Johnson (2000); Morrison (1987); Peña (2002) o Uriel (1995)]. Las secciones 4 y 5 se dedican a la descripción y análisis de los resultados obtenidos en el Análisis Factorial y el Análisis *Cluster*, respectivamente; la sección 6 se dedica a la validación de resultados; las conclusiones del estudio y las referencias bibliográficas se incluyen en las secciones 7 y 8, respectivamente. Por último, se incluye un anexo con información adicional.

## **2. Descripción y análisis univariante de los datos**

En este trabajo hacemos uso de las estadísticas oficiales elaboradas por el Instituto de Estadística de Andalucía (IEA) para el banco de datos del Sistema de Información Multiterritorial de Andalucía (SIMA). En él podemos encontrar datos relativos a los entornos físico, demográfico, económico y social de cualquier ámbito territorial de la Comunidad Autónoma de Andalucía; ya sea regional, provincial, municipal,

entidad de población, capital e incluso distrito censal. Asimismo, y puesto que la unidad de análisis en el presente trabajo es el municipio, utilizamos la información contenida en las fichas municipales elaboradas por el IEA para el SIMA a partir de su banco de datos. Estas fichas ofrecen una síntesis de los valores de los principales indicadores estadísticos relativos a entorno físico, población, sociedad, economía, mercado de trabajo y hacienda, de todos los municipios de la Comunidad Autónoma de Andalucía.

La selección y definición de las variables o indicadores que utilizamos en nuestro análisis queda condicionada, como en casi todos los trabajos de índole práctica, por la información útil disponible en el medio que utilizamos como base de datos, así como por la necesidad de hacer que dicha información resulte comparable a nivel municipal y refleje directa o indirectamente las características socioeconómicas de los municipios que se pretenden clasificar o tipificar. Dichas variables entran en el análisis en términos absolutos, cuando se pretende que sean reflejo del tamaño del municipio y del volumen absoluto de equipamiento, renta, superficie, etc., o como cocientes (*ratios* por habitante, por 100 habitantes, por km<sup>2</sup>, etc.) o porcentajes del total, si lo que se pretende es facilitar la comparación entre municipios con tamaño no similar en cuanto a las características analizadas.

Los indicadores se han clasificado en siete grupos, tratando de respetar la estructura estadística de las fichas municipales del SIMA, denominándose a éstos: Entorno Físico; Población; Educación; Equipamiento y Sociedad; Turismo; Economía y Mercado de trabajo y, por último, Hacienda.

La matriz de datos está constituida por los valores de los 58 indicadores considerados en los 769 municipios andaluces. Los indicadores utilizados, la abreviatura con la que aparecen en el trabajo, las definiciones o fórmulas de cálculo de los mismos y el año al que corresponden los datos utilizados, últimos disponibles, se presentan a continuación:

### Entorno físico

- **Extensión superficial (ES):** La superficie a la que se refiere esta variable es la de cada término municipal completo y está expresada en kilómetros cuadrados. Año 2003.
- **Altitud sobre el nivel del mar (ASM):** Medida en metros, es la correspondiente a un punto de la entidad singular principal. Año 1999.

### Población

- **Población total (PT):** Cifras de población referidas al 1 de enero de 2007, resultantes de la revisión del Padrón municipal, que han sido declaradas oficiales por el Gobierno mediante el Real Decreto 1683/2007, de 14 de diciembre. Año 2007.

- **Densidad de población (DP):** Razón entre la población total del municipio y la extensión superficial del mismo, medida en habitantes/km<sup>2</sup>. Año 2007.
- **Población menor de veinte años (P20):** Porcentaje que la población con edad inferior a 20 años representa del total de población del municipio. Año 2007.
- **Población mayor de sesenta y cuatro años (P64):** Porcentaje que la población con edad superior a 64 años representa del total de población del municipio. Año 2007.
- **Incremento relativo de la población (IRP):** Tasa de variación de la población del municipio en el periodo 1996-2007. Año 2007.
- **Extranjeros<sup>2</sup> (PEX):** Calculados por cada 100 habitantes. Año 2007.
- **Inmigrantes<sup>3</sup> (PIN):** Calculados por cada 100 habitantes. Año 2006.
- **Nacidos vivos por residencia materna (PNV):** Calculados por cada 100 habitantes. Año 2006.
- **Matrimonios por lugar donde fijan su residencia (PMAT):** Calculados por cada 100 habitantes. Año 2006.

## Educación

- **Tasa de analfabetismo en mujeres (TANM):** La tasa de analfabetismo se calcula sobre el conjunto de población residente en viviendas familiares y con dieciséis o más años de edad y recoge la relación porcentual entre la población femenina analfabeta y la población femenina total correspondiente a ese tramo de edad. Año 2001.
- **Tasa de analfabetismo en hombres (TANH):** Se calcula sobre el mismo conjunto que la anterior, en este caso contabilizando sólo a los varones. Año 2001.
- **Mujeres con estudios de segundo grado o superior (MES):** Porcentaje de mujeres con estudios de segundo grado o superior calculado frente a las que tienen dieciséis o más años. Año 2001.
- **Hombres con estudios de segundo grado o superior (HES):** Porcentaje calculado frente a los varones de dieciséis o más años. Año 2001.

## Equipamiento y Sociedad

- **Centros de educación básica (CEB).** Año 2005.
- **Centros de educación secundaria (CES).** Año 2005.
- **Centros de educación de adultos (CEA).** Año 2005.
- **Centros de salud (CS).** Año 2006.

---

<sup>2</sup> Se considera toda la población de nacionalidad extranjera residente en cualquier municipio de la Comunidad Autónoma Andaluza.

<sup>3</sup> Las inmigraciones recogidas son las que tienen como origen cualquiera de España y destino cualquier municipio andaluz. También se recogen las inmigraciones procedentes del extranjero.

- **Centros de Salud y Consultorios (CSCONPH):** Calculados por cada 100 habitantes. Año 2006.
- **Viviendas Familiares Principales (VFP) y Viviendas Familiares Principales por cada 100 habitantes (VFPPH).** Año 2001.
- **Vehículos Turismos (VT).** Año 2006.
- **Oficinas Bancarias (OB).** Año 2006.
- **Líneas Telefónicas (LT) y Líneas Telefónicas por cada 100 habitantes (LTPH).** Año 2006.
- **Líneas ADSL (ADSL) y Líneas ADSL por cada 100 habitantes (ADSLPH).** Año 2006.

## Turismo

- **Número de Restaurantes (R).** Año 2006.
- **Número de Hoteles (H).** Año 2006.
- **Número de Hostales y Pensiones (HP) y Número de Hostales y Pensiones por cada 100 habitantes (HPPH).** Año 2006.
- **Plazas en Hostales y Pensiones por cada 100 habitantes (PHPPH).** Año 2006.
- **Establecimientos Hoteleros por cada 100 habitantes (EHPH).** Año 2006.

## Economía y Mercado de trabajo

- **Número de mujeres registradas como paradas (PRM).** Año 2007.
- **Número de hombres registrados como parados (PRH).** Año 2007.
- **Número de Contratos Registrados correspondientes a Mujeres (CRM).** Año 2007.
- **Número de Contratos Registrados correspondientes a Hombres (CRH).** Año 2007.
- **Número de Contratos Registrados Indefinidos (CRI).** Año 2007.
- **Ocupados en Agricultura (OA):** Porcentaje calculado respecto al total de ocupados [Sección A de la Clasificación Nacional de Actividades Económicas (CNAE-93)]. Año 2001.
- **Superficie cultivada (SC):** Suma de la superficie dedicada a cultivos herbáceos y de la dedicada a cultivos leñosos, expresada en hectáreas. Año 2006.
- **Superficie Cultivada dedicada a Cultivos Herbáceos (SCCH):** Porcentaje del total de superficie cultivada que representa la dedicada a cultivos herbáceos. Año 2006.
- **Superficie Agrícola (SA):** Porcentaje del total de superficie del municipio que representa la superficie agrícola. Año 2005.
- **Trabajadores Eventuales Agrarios Subsidiados Mujeres (TEASM) y Trabajadores Eventuales Agrarios Subsidiados Mujeres por cada 100 habitantes (TEASMPH).** Año 2006.

- **Trabajadores Eventuales Agrarios Subsidiados Hombres (TEASH) y Trabajadores Eventuales Agrarios Subsidiados Hombres por cada 100 habitantes (TEASHPH).** Año 2006.
- **Número de Establecimientos con Actividad Económica (EAE) y Número de Establecimientos con Actividad Económica por cada 100 habitantes (EAEPH).** Año 2006.
- **Número de Establecimientos con Actividad Económica, por cada 100 habitantes, que tienen cinco trabajadores como máximo (EAE5).** Año 2006.
- **Número de Establecimientos con Actividad Económica, por cada 100 habitantes, que tienen veinte o más trabajadores (EAE20).** Año 2006.

## Hacienda

- **Renta Total<sup>4</sup> (RT):** Expresada en euros. Año 2005.
- **Renta Total por habitante (RTPH):** Expresada en euros. Año 2005.
- **Rentas del Trabajo por habitante (RTRPH):** Expresadas en euros. Año 2005.
- **Rentas de Actividades Empresariales por habitante (RAEPH):** Expresadas en euros. Año 2005.
- **Renta Neta Media Declarada (RNMD):** Expresada en euros. Año 2005.
- **Licencias en Actividades Empresariales (LAE).** Año 2005.
- **Licencias en Actividades Empresariales por cada 100 habitantes (LAE-PH).** Año 2005.

La descripción estadística a nivel unidimensional de los indicadores anteriores aparece resumida en el cuadro 1 (Anexo). De la información contenida en el mismo y del análisis de los datos básicos podemos concluir la existencia de diferencias bastante acentuadas entre los municipios, en cuanto a los valores de algunos indicadores. En este sentido cabe señalar que las diferencias más notables, en términos absolutos, se presentan en la renta total (RT), la densidad de población (DP) y la población total (PT). Los valores municipales de estas variables están entre los 29.759,4 € de Benitagla (Almería) y los 5.494.318.659,4 € de Sevilla; entre los 289,19 habitantes por km<sup>2</sup> de Zufre (Huelva) y el 1.062.429,75 habitantes por km<sup>2</sup> de Cádiz; y entre 50 habitantes [Cumbres de En medio (Huelva)] y 699.145 habitantes (Sevilla), respectivamente. Por su parte, los coeficientes de variación (CV) de estas variables son: 499,8%; 361% y 375,4 %, respectivamente; ello indica que tanto las dos variables que son relativas a población como la renta total presentan una fuerte dispersión con respecto al valor medio municipal.

En general, las variables medidas en términos absolutos presentan fuerte variabilidad con respecto al valor medio municipal y, aunque ello conlleva, obviamente,

---

<sup>4</sup> Todos los indicadores de renta que figuran en este apartado son relativos a la sección de IRPF que aparece en las fichas municipales del Banco de Datos SIMA (IEA). La denominada Renta Total (RT) coincide con la renta neta total declarada, es decir, sumas de rentas del trabajo, rentas netas en estimación directa, rentas netas en estimación objetiva y otro tipo de rentas.

te, pérdida de representatividad de dicho valor, es precisamente esa variabilidad la que resulta útil para este tipo de estudio, ya que facilita la clasificación. Además, la inclusión de variables medidas en términos absolutos permite simbolizar la dimensión municipal en cuanto a extensión superficial, población, equipamientos, economía, etc.

En cuanto a las variables que entran en el análisis en términos relativos o en proporción a la población, las que presentan mayor variabilidad son el incremento relativo de población (IRP) y la densidad de población (DP), seguidas de las tres relativas al turismo: hostales y pensiones (HPPH), plazas en hostales y pensiones (PHPPH) y establecimientos hoteleros (EHPH), todas ellas contabilizadas por cada 100 habitantes. En el otro extremo, es decir, variables que presentan poca dispersión, se encuentran las viviendas familiares principales por cada 100 habitantes (VFPPH), con un valor medio municipal de 34,8 viviendas por cada 100 habitantes; el porcentaje de población con edad inferior a veinte años (P20), con un valor medio de un 20,1% y las líneas telefónicas por cada 100 habitantes (LTPH), con una media de 25,6 líneas por cada 100 habitantes. Estas tres variables presentan valores del coeficiente de variación que son inferiores al 25%. Las variables relativas a renta: renta neta media declarada (RNMD), renta total por habitante (RTPH), renta del trabajo por habitante (RTRPH) o renta de actividades empresariales por habitante (RAEPH), presentan niveles medios o bajos de dispersión, con valores del coeficiente de variación inferiores al 70%. Asimismo, presentan niveles de dispersión similares a los anteriores las variables relativas a formación educacional, tanto en sus niveles inferiores, tasas de analfabetismo (TANM y TANH), como las relativas a nivel de estudios de grado medio o superior (MES y HES).

### 3. Metodología

En una primera etapa, y tras analizar la matriz de correlaciones correspondiente a los 58 indicadores seleccionados, aplicamos la técnica de Análisis Factorial con la finalidad de identificar, si es posible, aquellos indicadores que están expresando un mismo aspecto latente (o no medido) de la realidad socioeconómica de los municipios andaluces y, simultáneamente, poder explicar dicha realidad mediante un número reducido de factores que resuman adecuadamente la información contenida en las variables iniciales.

Los factores obtenidos se utilizarán para identificar grupos de municipios con características socioeconómicas similares mediante un Análisis *Cluster*.

#### 3.1. Análisis Factorial

La técnica estadística multivariante de Análisis Factorial está basada en la hipótesis de que la información relativa a un gran número de variables observadas se puede comprimir en un número reducido de variables subyacentes o no observables,



llamadas factores. Su utilización está especialmente indicada en el análisis y explicación de fenómenos complejos, ya que puede permitir la identificación de aquellos indicadores que expresan un mismo aspecto latente (o no medido) de la realidad que es objeto de estudio y, a través de ellos, la identificación de las dimensiones subyacentes del fenómeno analizado.

La determinación de estas dimensiones, que puede llevarse a cabo esencialmente de dos maneras, depende del conocimiento que poseamos *a priori* sobre la realidad estudiada. Una posibilidad es la que consiste en la especificación previa de un modelo teórico mediante el que se define la relación de causalidad inobservable existente entre los factores y los indicadores observados. Dicho modelo se pone a prueba (o se confirma) mediante el análisis empírico. Esto se puede hacer mediante un Análisis Factorial Confirmatorio en el que las relaciones que deben someterse a prueba han sido especificadas previamente por el analista. Otra posibilidad es la que consiste en explorar los datos, en un intento de identificar patrones de asociación no aleatoria entre las variables observadas, y dejar que los factores necesarios para la explicación del fenómeno emerjan de dicho análisis. Esto es en esencia lo que permite hacer el Análisis Factorial Exploratorio, técnica que emplearemos en esta investigación. A partir de las variables originales, se obtiene un número reducido de factores correlacionados con las variables observadas e incorrelacionados entre sí, que resumen la información contenida en dichas variables y reflejan la estructura subyacente en el conjunto de datos.

El Análisis Factorial presupone la existencia de un modelo estadístico formal (en nuestro caso el modelo factorial lineal) a través del cual las variables se relacionan con los factores, que es similar al de regresión lineal múltiple. La idea es que las variables observadas, que supondremos tipificadas, se pueden expresar como combinación lineal de factores no observados, del siguiente modo:

$$X_{ij} = \sum_{k=1}^m \lambda_{ik} F_{kj} + U_{ij},$$

donde  $X_{ij}$  es el valor de la variable  $X_i$  en la  $j$ -ésima observación (en nuestro caso, en el municipio  $j$ );  $F_{kj}$  es el valor del  $k$ -ésimo factor en el municipio  $j$  (habitualmente se le denomina puntuación de dicho factor en el municipio correspondiente);  $\lambda_{ik}$  es el coeficiente de regresión estandarizado de la variable  $X_i$  sobre el factor común  $F_k$  (se le denomina carga sobre el factor) y, por último,  $U_{ij}$  representa la parte del valor de la variable  $X_i$  en el  $j$ -ésimo municipio que no viene explicada por los factores comunes a través del modelo especificado.

En el modelo anterior todos los factores se suponen incorrelacionados entre sí. Los factores representados por  $F$  son llamados factores comunes ya que su efecto es compartido conjuntamente por las variables. Por su parte, los factores representados con  $U$  reciben el nombre de factores únicos, o errores en las variables, ya que cada uno actúa exclusivamente sobre una variable. A cada uno de ellos se le atribuye la parte de la variabilidad en la variable correspondiente que no queda explicada por los

factores comunes. La proporción de la varianza de la variable  $X_i$  que queda explicada por los factores comunes  $\left( \sum_{k=1}^m \lambda_{ik}^2 \right)$  se denomina comunalidad de  $X_i$ .

La determinación de los factores está estrechamente relacionada con la obtención de las comunalidades. En este trabajo para generar los factores utilizaremos el Método de Componentes Principales que parte de la no existencia de factores únicos y que, por tanto, no necesita de la estimación previa de las comunalidades iniciales. Todas las variables entran en el análisis inicialmente con comunalidad igual a la unidad.

Los coeficientes de carga,  $\lambda_{ik}$ , describen la proximidad entre la variable  $X_i$  y el factor  $F_k$  e indican, al estar las variables tipificadas, cuánto peso se le asigna a cada factor. Cuando los factores se generan mediante Componentes Principales, dichos coeficientes coinciden con los coeficientes de correlación lineal entre las variables observadas y los factores generados. En este sentido, cuanto más alto sea el valor de la carga sobre un factor más estrecha es la relación entre el mismo y la variable correspondiente.

### 3.2. Análisis Cluster

El otro método estadístico empleado en este trabajo es el Análisis *Cluster* o de Conglomerados. El propósito de esta técnica estadística (o conjunto de técnicas) es agrupar objetos basándose en las características que poseen. La idea básica es que los objetos pertenecientes a un mismo grupo o conglomerado sean muy parecidos entre sí con respecto a algún criterio de selección predeterminado y que dichos grupos sean, en cuanto a ese criterio, tan diferentes entre ellos como sea posible (Hair y otros, 2000).

Las técnicas empleadas en este tipo de análisis pueden ser jerárquicas o no jerárquicas. Las primeras parten de considerar inicialmente tantos grupos como objetos: cada objeto o individuo forma un grupo. En el primer paso se agrupan los dos casos con menor distancia<sup>5</sup> (o mayor similitud). En cada una de las sucesivas etapas, siguiendo el criterio de mínima distancia, se forma un nuevo grupo por unión de dos individuos aislados, de dos grupos, o por la adhesión de un individuo a un grupo ya formado en una etapa anterior. El proceso culmina con la obtención de un único grupo, que engloba a todos los individuos u objetos, formado por las aglutinaciones sucesivas, con niveles de similitud o distancia distintos<sup>6</sup>. La representación gráfica

<sup>5</sup> La creación de grupos basados en la similitud de los elementos en ellos englobados exige una definición de este concepto, o de su complementario de «distancia» entre individuos. Existen diferentes formas de medir diferencias multidimensionales que, a su vez, proporcionan diversas posibilidades de análisis. La utilizada en este trabajo es el cuadrado de la denominada distancia Euclídea, que mide la distancia entre dos elementos mediante la raíz cuadrada de la suma de los cuadrados de las diferencias entre los valores que cada variable toma en dichos elementos. Hair y otros (2000) indican que la elección de una u otra medida no influye significativamente en el resultado final del análisis.

<sup>6</sup> Este procedimiento jerárquico de formación de grupos es de tipo aglomerativo o ascendente. También existe el disociativo o descendente (véase González López-Valcarcel, 1991).

de estas etapas de formación de grupos, a modo de árbol invertido, se denomina dendograma.

Los métodos no jerárquicos, también conocidos como partitivos o de optimización, tienen el objetivo de formar clases excluyentes, en número fijado de antemano por el investigador, de modo que sean homogéneas en su interior y con máxima divergencia entre ellas. Las clases que se forman constituyen una única partición y no están organizadas jerárquicamente ni relacionadas entre sí. La asignación de individuos a los grupos se hace mediante algún proceso que optimice el criterio de selección.

Autores como Hair y otros (2000) destacan la utilidad de los métodos jerárquicos para decidir el número de *cluster* a formar, pero reconocen el mayor uso y las mayores ventajas que ofrecen los algoritmos no jerárquicos. Entre las ventajas se incluyen el poder trabajar con un número muy elevado de objetos y de variables y, sobre todo, la reasignación de elementos que ellos permiten al ser iterativos.

Fijado el número de conglomerados a obtener, el primer paso es seleccionar una semilla, como centro de conglomerado inicial, y todos los elementos dentro de una distancia umbral que ha sido previamente especificada se incluyen dentro del conglomerado resultante. Entonces se selecciona otra semilla de conglomerado y la asignación continúa hasta que todos los elementos están distribuidos (Hair y otros, 2000). Los procedimientos de aglomeración no jerárquicos se denominan frecuentemente como aglomeración de K-medias y difieren, esencialmente, en los criterios empleados para la selección de la semilla y la asignación de los elementos.

## 4. Resultados del Análisis Factorial

Como se ha señalado antes el estudio de la realidad socioeconómica municipal en Andalucía lo basamos en el comportamiento conjunto de los 58 indicadores ya descritos y, más concretamente, en el análisis de la matriz de correlaciones asociada a los mismos, paso previo para la aplicación del Análisis Factorial.

El valor del determinante de la matriz de correlaciones ( $1,46 \times 10^{-56}$ ), muy próximo a cero, indica la existencia de dependencia lineal en los datos y la no existencia de indicadores redundantes, es decir, ninguno de los indicadores incluidos es combinación lineal perfecta de otros que entren también en el análisis. A partir de dicho determinante se ha calculado el valor del coeficiente de dependencia efectiva<sup>7</sup>,  $D(R) = 0,895$ , lo que indica que el grado de dependencia lineal entre las variables es apreciable.

---

<sup>7</sup> Este coeficiente se define como  $D(R) = 1 - |R|^{1/(p-1)}$ , siendo  $|R|$  el determinante de la matriz de correlaciones. Sabemos que si una de las variables es combinación lineal perfecta de otras, también incluidas en el análisis, la matriz de correlaciones es singular,  $|R| = 0$  y, por tanto,  $D(R) = 1$ . Por otra parte, si entre los distintos pares de variables la correlación lineal es nula, la matriz de correlaciones coincide con la identidad, su determinante es 1 y, en consecuencia,  $D(R) = 0$ . La comparación del coeficiente de dependencia con estos casos extremos puede darnos una buena idea del grado de dependencia lineal que existe entre los indicadores utilizados (D. Peña, 2002).

Como complemento del análisis del grado de dependencia lineal se ha calculado el índice KMO (Kaiser-Meyer-Olkin)<sup>8</sup>. El valor de dicho índice,  $KMO = 0,915$ , indica que la suma de los cuadrados de los coeficientes de correlación parcial es pequeña comparada con la correspondiente a los cuadrados de los coeficientes de correlación y avala, por tanto, la pertinencia de la realización de un Análisis Factorial exploratorio.

Dado que los resultados del Análisis Factorial se van a utilizar posteriormente para obtener agrupaciones de municipios que sean homogéneos en cuanto a sus características socioeconómicas y que para ello son necesarias las puntuaciones factoriales, el método de extracción empleado ha sido el de Componentes Principales (F. D. Morrison, 1987), reteniendo aquellos factores con autovalor mayor que la unidad (H. F. Kaiser, 1960). Por último, para facilitar la interpretación de los factores retenidos se ha aplicado el procedimiento de rotación ortogonal *varimax*. Los autovalores asociados a los factores retenidos, el porcentaje del total de varianza explicada por cada uno de ellos después de la rotación *varimax* y el porcentaje acumulado se presentan en la siguiente tabla (tabla 1).

**Tabla 1.** Varianza explicada

	<i>Autovalor</i>	<i>% varianza</i>	<i>% acumulado</i>
Factor 1: <i>Dimensión municipal</i>	19,7393	34,0332	34,0332
Factor 2: <i>Dinamismo económico-demográfico</i>	7,6792	13,2400	47,2733
Factor 3: <i>Actividad agrícola</i>	3,2459	5,5964	52,8697
Factor 4: <i>Actividad económica intermedia</i>	3,0659	5,2860	58,1557
Factor 5: <i>Oferta turística relativa</i>	2,8599	4,9308	63,0865
Factor 6: <i>Educativo</i>	2,7626	4,7632	67,8497
Factor 7: <i>Desarrollo agrícola</i>	2,7252	4,6987	72,5483
Factor 8: <i>Movilidad poblacional</i>	2,1320	3,6758	76,2241
Factor 9: <i>Analfabetismo</i>	1,8772	3,2366	79,4607
Factor 10	1,6935	2,9199	82,3806

Los diez factores retenidos, de acuerdo con el criterio basado en seleccionar los correspondientes a autovalores mayores que la unidad, explican el 82% de la variabilidad total, porcentaje más que aceptable si se tiene en cuenta que en estudios relativos a las Ciencias Sociales el límite inferior de admisibilidad es el 60% (Hair y otros; 2000).

<sup>8</sup> Este índice se define como: 
$$KMO = \frac{\sum_{j \neq k} r_{jk}^2}{\sum_{j \neq k} r_{jk}^2 + \sum_{j \neq k} a_{jk}^2},$$

donde  $r_{jk}$  es el coeficiente de correlación lineal simple entre las variables observadas  $j$  y  $k$  y  $a_{jk}$  es el coeficiente de correlación parcial entre dichas variables.

La tabla 2 contiene la matriz factorial rotada o matriz de componentes rotados, que está formada por los coeficientes de correlación lineal entre los factores y los indicadores a partir de los cuales éstos han sido estimados. Para facilitar la interpretación de dichos factores, en términos de las características socioeconómicas de los municipios que los mismos explican, se han puesto en negrita los coeficientes que en valor absoluto son superiores a 0,4.

**Tabla 2.** Matriz factorial rotada

	<i>Factor 1</i>	<i>Factor 2</i>	<i>Factor 3</i>	<i>Factor 4</i>	<i>Factor 5</i>	<i>Factor 6</i>	<i>Factor 7</i>	<i>Factor 8</i>	<i>Factor 9</i>	<i>Factor 10</i>
ES	0,205	0,018	-0,171	0,084	-0,007	-0,071	<b>0,820</b>	-0,116	0,078	-0,031
ASM	-0,164	<b>-0,530</b>	-0,026	-0,102	0,221	-0,023	-0,058	-0,154	-0,121	0,435
PT	<b>0,986</b>	0,104	-0,031	0,008	-0,023	0,030	0,067	-0,015	-0,015	-0,025
DP	0,101	0,118	0,050	0,011	-0,070	<b>0,710</b>	-0,101	-0,020	0,059	-0,125
P20	0,044	<b>0,806</b>	0,235	0,102	-0,067	-0,002	0,131	-0,141	0,076	-0,157
P64	-0,111	<b>-0,877</b>	-0,051	-0,097	-0,004	-0,103	-0,065	-0,087	0,014	0,209
IRP	0,014	<b>0,649</b>	-0,202	-0,051	-0,025	0,063	-0,090	<b>0,552</b>	-0,082	0,075
PEX	0,039	-0,010	-0,116	0,215	0,022	0,048	-0,073	<b>0,844</b>	0,019	0,008
PIN	-0,044	-0,014	-0,365	-0,181	-0,005	0,010	-0,163	<b>0,724</b>	-0,170	0,093
PNV	0,077	<b>0,726</b>	0,079	0,017	0,018	0,039	0,034	0,034	-0,100	-0,086
PMAT	0,046	<b>0,719</b>	0,008	0,024	-0,009	0,009	0,054	-0,096	-0,102	-0,116
TANM	-0,052	-0,059	0,187	-0,065	0,069	-0,332	0,097	-0,128	<b>0,809</b>	0,038
TANH	-0,032	-0,025	0,067	-0,082	0,048	-0,235	0,020	-0,005	<b>0,883</b>	-0,041
MES	0,084	0,280	-0,121	0,095	0,056	<b>0,811</b>	0,049	0,039	-0,242	0,078
HES	0,085	0,222	-0,154	0,043	0,006	<b>0,836</b>	0,057	0,047	-0,246	0,106
CEB	<b>0,976</b>	0,095	-0,034	-0,001	-0,017	0,037	0,111	-0,035	-0,011	-0,025
CES	<b>0,979</b>	0,078	-0,025	0,000	-0,021	0,054	0,064	-0,052	-0,010	-0,020
CEA	<b>0,877</b>	0,095	0,018	0,033	-0,032	-0,024	0,103	-0,089	-0,010	-0,082
CS	<b>0,964</b>	0,091	-0,023	0,025	-0,022	0,051	0,087	-0,041	-0,008	-0,025
CSCONSPH	-0,082	<b>-0,617</b>	-0,229	-0,211	0,134	-0,045	-0,116	0,111	-0,209	0,147
VFP	<b>0,989</b>	0,079	-0,016	0,004	-0,021	0,029	0,058	-0,030	-0,014	-0,021
VFPPH	-0,090	<b>-0,823</b>	0,115	-0,109	0,058	-0,074	-0,098	-0,344	0,029	0,168
VT	<b>0,990</b>	0,094	-0,029	0,017	-0,022	0,023	0,040	-0,008	-0,021	-0,015
OB	<b>0,991</b>	0,063	-0,006	0,032	-0,018	0,020	0,013	0,000	-0,015	0,000
LT	<b>0,988</b>	0,094	-0,027	0,043	-0,020	0,018	0,031	0,014	-0,012	0,023
LTPH	0,022	-0,136	-0,276	0,337	0,052	0,063	-0,104	-0,005	-0,063	<b>0,678</b>
ADSL	<b>0,973</b>	0,124	-0,046	0,055	-0,023	0,021	-0,002	0,033	-0,011	0,039
ADSLPH	0,149	<b>0,765</b>	-0,192	0,259	-0,083	0,174	0,020	0,096	0,002	0,250
R	<b>0,820</b>	0,109	-0,061	0,196	0,015	0,063	-0,006	0,155	0,034	0,073
H	<b>0,809</b>	0,110	-0,083	0,185	0,061	0,047	0,073	0,134	0,024	0,069
HP	<b>0,809</b>	0,079	-0,031	0,143	0,151	0,072	0,022	0,086	0,039	-0,008
HPPH	-0,019	-0,106	-0,032	0,088	<b>0,944</b>	0,006	-0,036	-0,003	0,030	0,000
PHPPH	-0,012	-0,065	-0,058	0,131	<b>0,926</b>	0,021	-0,025	-0,032	0,042	-0,019
EHPH	-0,022	-0,132	-0,090	0,164	<b>0,914</b>	0,011	-0,042	0,048	0,020	0,049
PRM	<b>0,966</b>	0,087	-0,068	-0,019	-0,021	0,041	0,084	-0,057	-0,018	-0,043
PRH	<b>0,978</b>	0,085	-0,043	-0,005	-0,019	0,050	0,068	-0,033	-0,017	-0,050
CRM	<b>0,979</b>	0,065	0,028	0,016	-0,014	0,008	0,079	-0,001	-0,017	-0,021

**Tabla 2.** (Continuación)

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8	Factor 9	Factor 10
CRH	<b>0,978</b>	0,097	0,005	0,017	-0,021	0,007	0,107	-0,011	-0,022	-0,013
CRI	<b>0,986</b>	0,045	-0,009	0,017	-0,014	-0,013	-0,030	-0,001	-0,032	0,000
OA	0,011	-0,028	0,265	-0,030	-0,087	<b>-0,611</b>	0,049	-0,028	0,270	-0,060
SC	0,183	0,117	0,060	0,026	-0,044	-0,048	<b>0,884</b>	-0,095	-0,014	0,041
SCCH	0,085	0,218	-0,254	0,102	0,067	-0,017	0,160	-0,252	-0,059	<b>-0,586</b>
SA	0,029	0,275	<b>0,555</b>	-0,068	-0,074	-0,288	0,159	-0,159	-0,144	0,119
TEASM	0,167	0,219	<b>0,451</b>	0,018	-0,053	0,033	<b>0,712</b>	-0,010	0,026	-0,192
TEASMPH	-0,138	-0,192	<b>0,795</b>	-0,156	-0,080	-0,091	0,055	-0,143	0,102	-0,156
TEASH	0,096	0,180	<b>0,503</b>	-0,004	-0,055	0,029	<b>0,691</b>	-0,010	0,058	-0,213
TEASHPH	-0,137	-0,362	<b>0,712</b>	-0,169	-0,054	-0,063	0,010	-0,146	0,144	-0,119
EAE	<b>0,990</b>	0,086	-0,022	0,052	-0,021	0,030	0,024	0,007	-0,015	0,005
EAEPH	0,158	0,337	-0,074	<b>0,871</b>	0,116	0,093	0,048	0,024	-0,031	0,045
EAE5	0,133	0,272	-0,040	<b>0,876</b>	0,147	0,091	0,056	-0,017	-0,006	0,039
EAE20	0,161	<b>0,454</b>	-0,159	0,363	-0,106	0,032	0,041	0,061	-0,101	-0,028
RT	<b>0,989</b>	0,071	-0,028	-0,006	-0,017	0,025	0,009	-0,035	-0,016	0,002
RTPH	0,296	<b>0,752</b>	-0,194	0,143	-0,053	0,171	0,027	-0,246	0,023	0,331
RTRPH	0,292	<b>0,751</b>	-0,285	0,044	-0,057	0,203	0,002	-0,232	-0,008	0,257
RAEPH	-0,081	-0,025	<b>0,569</b>	0,106	-0,036	-0,179	0,059	-0,131	0,160	0,284
RNMD	0,278	<b>0,753</b>	-0,382	0,132	0,009	0,216	0,023	-0,005	-0,001	0,127
LAE	<b>0,985</b>	0,101	-0,025	0,065	-0,022	0,025	0,049	0,020	-0,019	0,003
LAEPH	0,052	0,167	-0,041	<b>0,864</b>	0,227	-0,018	0,024	0,058	-0,086	0,050

El factor 1 está fuertemente asociado con 21 de los 58 indicadores empleados y explica el 34% de la variabilidad total. Este factor, al que denominamos *Dimensión municipal*, representa a la población total del municipio y a todas las variables que reflejan algún aspecto de la realidad socioeconómica del mismo y están medidas en términos absolutos, salvo a la extensión superficial y las variables relacionadas con la agricultura, ya que los coeficientes de carga de dichas variables sobre el primer factor son, en valor absoluto, superiores a 0,8. Ello, teniendo en cuenta además que las correlaciones son positivas, significa que valores altos (bajos) del factor están asociados a municipios con muchos (pocos) habitantes, con independencia de su tamaño en términos de superficie y que, paralelamente, presentan un elevado (bajo) número de parados o de contratos registrados, alta (baja) renta y un alto (bajo) nivel de equipamientos: centros de enseñanza, centros de salud, viviendas familiares, vehículos, oficinas bancarias, líneas telefónicas y ADSL, establecimientos hoteleros o dedicados a alguna actividad económica, restaurantes, etcétera.

Las puntuaciones más altas del primer factor se alcanzan en capitales de provincia. Concretamente, la puntuación más alta es la correspondiente a la ciudad de Sevilla con 18,34 puntos, seguida de Málaga con 13,67 puntos. A continuación se encuentran Granada y Córdoba con una puntuación de 7,4; Almería con 4,4, similar a la que presentan Jerez de la Frontera (Cádiz) o Marbella (Málaga) (4,2 y 4,1, respectivamente) y, por último, Huelva (3,1), Cádiz (2,8) y Jaén (2,5).

El factor 2 explica el 13,24% de la varianza total y está positivamente correlacionado con 13 variables: el porcentaje de población con edad inferior a veinte años (P20); el incremento relativo de población en el periodo 1996-2007 (IRP); el porcentaje de nacidos (PNV); el porcentaje de matrimonios (PMAT); las líneas ADSL por cada 100 habitantes (ADSLPH); la renta por habitante (RTPH y RTRPH); la renta neta media declarada (RNMD) y los grandes establecimientos dedicados a alguna actividad económica o empresarial (EAE20). Simultáneamente, este factor presenta correlación negativa con: la altitud del municipio sobre el nivel del mar (ASM); el porcentaje de población con edad superior a sesenta y cuatro años (P64) y, por último, los centros de salud y las viviendas familiares (CSPH y VFPPH) cuando se contabilizan por cada 100 habitantes. De ello se deduce que valores altos de este factor están asociados a municipios con población joven, altas tasas de natalidad, elevado número de matrimonios, con buen nivel de renta por habitante y que generan un volumen importante de actividad económica. Según este perfil, hemos denominado a este factor de *Dinamismo económico-demográfico*. Las primeras posiciones en cuanto a las puntuaciones de este factor las ocupan municipios limítrofes con las capitales de provincia, o muy próximos a éstas, que son utilizados como ciudad dormitorio. Entre ellos destacamos: Castilleja de Guzmán, Espartinas, Palomares del Río, Bormujos, Tomares, Gelves o Gines (Sevilla); Vegas del Genil (Granada); Huércal de Almería (Almería); La Guardia de Jaén (Jaén) o Rincón de la Victoria (Málaga).

En el tercer factor adquieren importancia aspectos relacionados con la agricultura, como son el número de trabajadores eventuales agrarios subsidiados, tanto masculinos como femeninos (TEASH, TEASM, TEASHPH y TEASMPH), que presentan cargas superiores a 0,7 si están expresados en términos relativos o que rondan el 0,5 cuando se consideran en términos absolutos (TEASH y TEASM). La superficie dedicada a la agricultura (SA) y las rentas de actividades empresariales por habitante (RAEPH) intervienen con una carga cercana a 0,6. Según esta información, podemos decir que este factor, cuyas puntuaciones varían entre los 3,2 puntos de Jódar en la Comarca de Sierra Mágina (Jaén) y los -2,77 de Aldeaquemada, pequeño municipio de la Comarca Norte también en la provincia de Jaén, refleja la *Actividad agrícola* del municipio.

El cuarto factor recibe el nombre de *Actividad económica intermedia*. En él resultan relevantes, por las altas correlaciones positivas alcanzadas, el número de establecimientos con actividad económica por cada 100 habitantes (EAEPH) y el número de licencias destinadas a este fin (LAEPH), también en términos relativos. Por otro lado, la variable «número de establecimientos con actividad económica con cinco trabajadores como máximo» (EAE5), que entra en el análisis medida por cada 100 habitantes y carga sobre este factor con un coeficiente de casi 0,9, nos hace suponer que pueden ser negocios de tipo familiar, autónomos o pequeñas y medianas empresas. Las puntuaciones que este factor alcanza en los municipios andaluces se encuentran entre 5,5 y -3,72 puntos y, teniendo en cuenta que entre los municipios con puntuación más alta se encuentran Mojácar (Almería), Bubión, Pampaneira, Trevélez (Granada), Marbella, Fuengirola, Nerja (Málaga), etc., no parece muy arriesgado pensar que en buena medida los valores altos de este factor pueden ir asociados a municipios

en los que el número de establecimientos con actividad económica dedicados al turismo o las licencias dedicadas a este fin son, en términos relativos, elevados.

La interpretación global que creemos adecuada para el quinto factor es la de *Oferta turística relativa*, ya que presenta correlaciones superiores a 0,9 con HPPH (Número de Hostales y Pensiones por cada 100 habitantes), PHPPH (Plazas en Hostales y Pensiones por cada 100 habitantes) y EHPH (Establecimientos Hoteleros por cada 100 habitantes); es decir, el factor aparece fuertemente correlacionado, tanto en número como en plazas, con los distintos tipos de establecimientos hoteleros por cada cien habitantes. Ahora bien, teniendo en cuenta la distribución municipal de las puntuaciones de este factor, creemos conveniente observar que los valores más altos del mismo se alcanzan en pequeños municipios turísticos situados en el interior. Concretamente, los primeros lugares, en orden descendente, están ocupados por las localidades de Cortes y Graena (al norte de Sierra Nevada, Granada); Bayárcal (el municipio más alto de la provincia de Almería); Pampaneira, Capileira y Juviles (en la Alpujarra Granadina); Parauta (en el paraje natural Sierra de las Nieves, en Málaga) o Atajate, en la Serranía de Ronda, también en Málaga.

El sexto es el factor *Educativo*, ya que en él adquieren una significación relevante variables como el porcentaje de mujeres y hombres con estudios de segundo grado o superior (MES y HES). Este factor presenta también correlación positiva fuerte con la densidad de población (DP) y negativa con el grado de ocupación en agricultura (OA).

El factor séptimo complementa al tercero, reflejando también aspectos relativos a la agricultura, como son la extensión de la superficie cultivada (SC) y el número de trabajadores con subsidios (TEASM y TEASH), ahora medidos en términos absolutos. También está correlacionado positivamente con la extensión superficial (ES), por lo que este factor podría denominarse de *Desarrollo agrícola*. Los valores más altos de este factor se presentan en municipios de gran extensión superficial en los que el sector agrícola es un importante motor de su economía. Entre ellos destacamos los de Écija, Carmona y Osuna, en la provincia de Sevilla; Jérez de la Frontera y Arcos de la Frontera, en Cádiz; Antequera, en la provincia de Málaga; Córdoba, Baena o Palma del Río, en Córdoba, o Almonte, Andújar y Loja, en las provincias de Huelva, Jaén y Granada, respectivamente.

El factor octavo, al que denominamos *Movilidad poblacional*, presenta cargas positivas con los extranjeros y los inmigrantes medidos en términos relativos (PEX y PIN) y una correlación moderada con el incremento relativo de la población (IRP). *Analfabetismo* es el nombre que recibe el noveno factor, que viene explicado por las tasas de analfabetismo tanto masculinas como femeninas (TANH y TANM). Finalmente, sobre el factor décimo cargan las variables líneas telefónicas por cada 100 habitantes (LTPH) y superficie cultivada dedicada a cultivos herbáceos (SCCH) presentando esta última signo negativo<sup>9</sup>.

---

<sup>9</sup> En este trabajo el factor núm. 10, a pesar de tener un autovalor mayor que 1, no se identifica claramente con ningún grupo de variables, por lo que no se le ha asignado una denominación como a los restantes. Sin embargo, creemos conveniente su inclusión por permitir la identificación clara del factor núm. 9 (*Analfabetismo*).

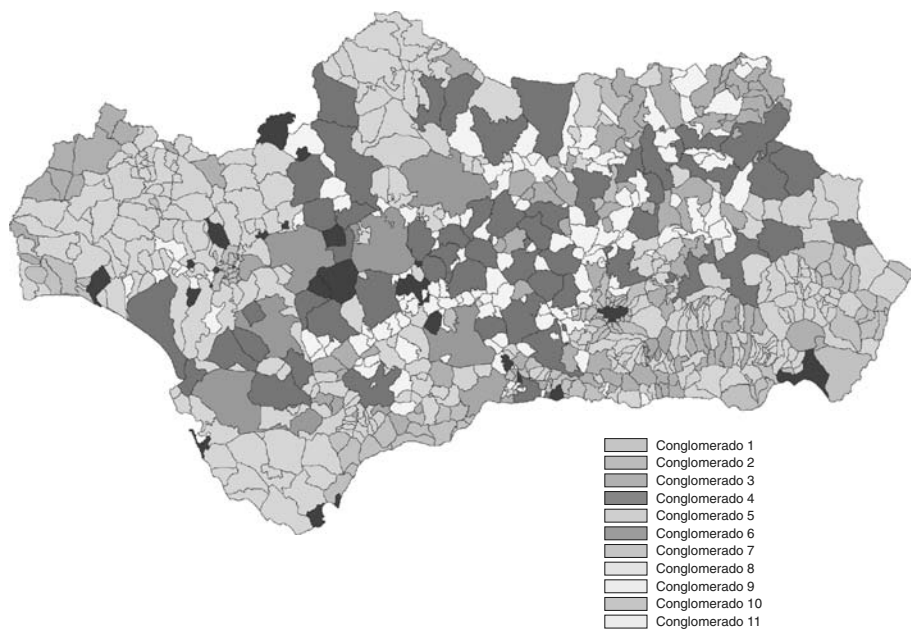


## 5. Resultados del Análisis *Cluster*

Como se ha señalado en la introducción, el siguiente paso en esta investigación es la aplicación de un método de aglomeración no jerárquico de Análisis *Cluster*, k-medias, a las puntuaciones municipales de los diez factores retenidos, para formar grupos de municipios que sean homogéneos en cuanto a las características socioeconómicas que resumen dichos factores y de modo que estas características difieran tanto como sea posible para municipios pertenecientes a distintos grupos. En cuanto a las razones que pueden avalar la pertinencia de utilizar un método no jerárquico frente a uno jerárquico, podemos argumentar el hecho de que las técnicas no jerárquicas están especialmente indicadas cuando, como en el caso que nos ocupa, el conjunto de datos es muy numeroso.

Tras un análisis previo en el que se han realizado múltiples pruebas con análisis no jerárquico, para analizar las variaciones experimentadas por la varianza residual al ir aumentando el número de grupos, y teniendo en cuenta los «saltos» en la distancia de agrupamiento en un análisis jerárquico <sup>10</sup>, hemos segmentado los municipios en once clases o conglomerados. En el Anexo aparecen los componentes de cada uno de los once grupos obtenidos. Igualmente, en el mapa 1 se puede observar la distribu-

Mapa 1



<sup>10</sup> El dendograma que resulta en la clasificación jerárquica de los municipios andaluces en base a los factores retenidos, utilizando la distancia euclídea al cuadrado y el método de Ward, se presenta en el Anexo (figura 1).

ción geográfica de los municipios englobados en los distintos *clusters* y en la tabla 3 las puntuaciones medias que los factores alcanzan en los mismos. La heterogeneidad y las diferencias intrarregionales existentes en la extensa Comunidad andaluza quedan patentes en este mapa.

**Tabla 3.** Centros de los Conglomerados finales

	1	2	3	4	5	6	7	8	9	10	11
Factor 1: <i>Dimensión municipal</i>	16,01	2,11	-0,15	0,72	-0,21	0,10	0,19	-0,14	-0,12	-0,09	-0,10
Factor 2: <i>Dinamismo económico-demográfico</i>	-0,84	0,22	-0,55	0,21	1,95	0,31	0,28	-1,10	0,03	-0,43	0,32
Factor 3: <i>Actividad agrícola</i>	1,06	-1,20	-0,20	0,39	-0,76	0,39	-0,19	-0,59	1,33	0,05	-0,46
Factor 4: <i>Actividad económica intermedia</i>	-0,90	-0,50	0,68	0,02	-0,50	0,22	1,02	-0,47	-0,12	-0,24	0,35
Factor 5: <i>Oferta turística relativa</i>	0,01	-0,04	5,92	-0,30	-0,04	0,01	-0,08	-0,18	-0,12	-0,02	-0,10
Factor 6: <i>Educativo</i>	-2,01	0,23	0,24	3,50	0,54	-0,13	-0,09	-0,02	-0,27	-0,18	-0,26
Factor 7: <i>Desarrollo agrícola</i>	-2,72	6,91	-0,31	-0,23	-0,59	2,03	-0,21	-0,17	-0,20	-0,14	-0,15
Factor 8: <i>Movilidad poblacional</i>	-0,39	-0,22	-0,39	-0,20	-0,20	-0,15	2,61	0,20	-0,25	-0,08	-0,44
Factor 9: <i>Analfabetismo</i>	-1,08	-0,52	-0,63	0,10	-0,46	-0,03	0,07	-0,56	-0,21	1,63	-0,03
Factor 10	-0,38	0,29	-0,04	-0,89	1,26	0,24	-0,03	0,20	0,23	0,37	-0,80

En los Gráficos 1, 2 y 3, se pueden visualizar las diferencias entre conglomerados, en cuanto a las puntuaciones medias que en ellos alcanzan los factores.

A partir de la matriz de centros de conglomerados finales (tabla 3) y de los gráficos anteriores podemos extraer los siguientes resultados.

Los conglomerados que engloban a un menor número de municipios (representados en el gráfico 1) son el 1 (dos municipios), el 2 (seis municipios) y el 3 (trece municipios). En contrapartida los *clusters* más numerosos (representados en el gráfico 3) son el 8 (142 municipios), el 9 (145 municipios), el 10 (91 municipios) y el 11 (189 municipios).

En las dos capitales de provincia que integran el conglomerado 1 (Málaga y Sevilla) se alcanzan las máximas puntuaciones de Andalucía relativas al factor que representa la *Dimensión municipal*, de hecho podemos afirmar que este conglomerado se caracteriza por el predominio de los valores de dicho factor frente a los alcanzados en estas capitales por los restantes. Por otra parte, es necesario destacar que en este conglomerado los valores medios de ocho de los restantes factores son inferiores a la media municipal.

Gráfico 1

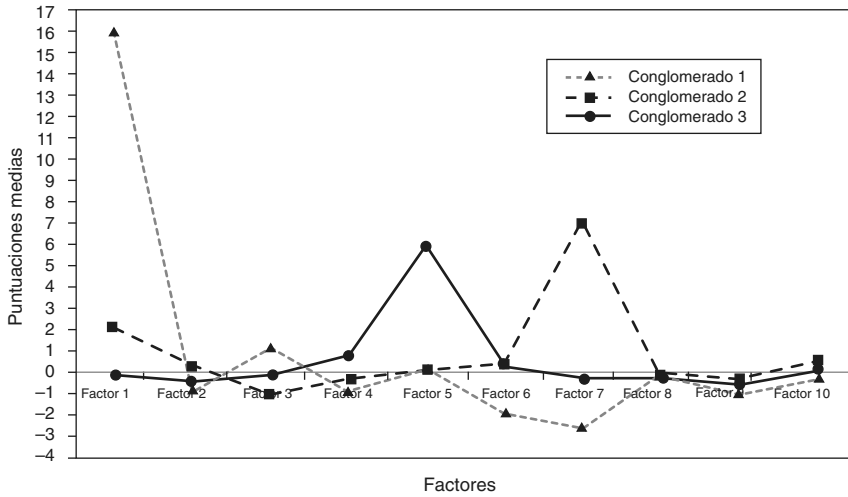
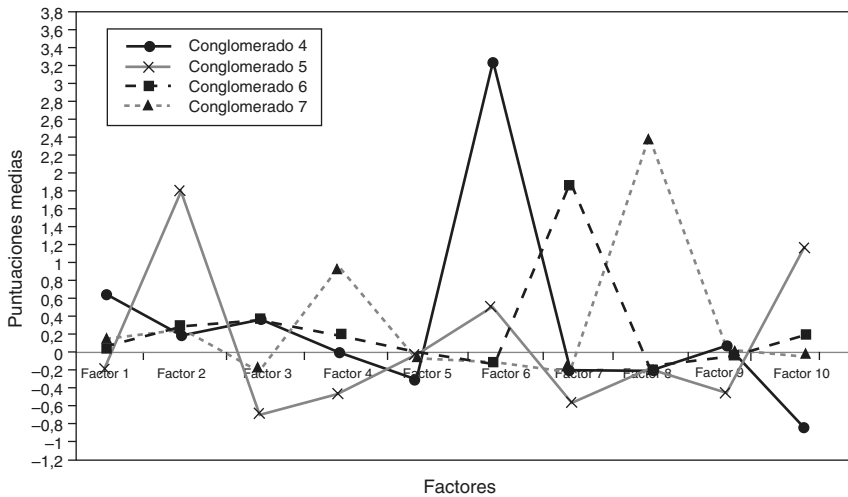
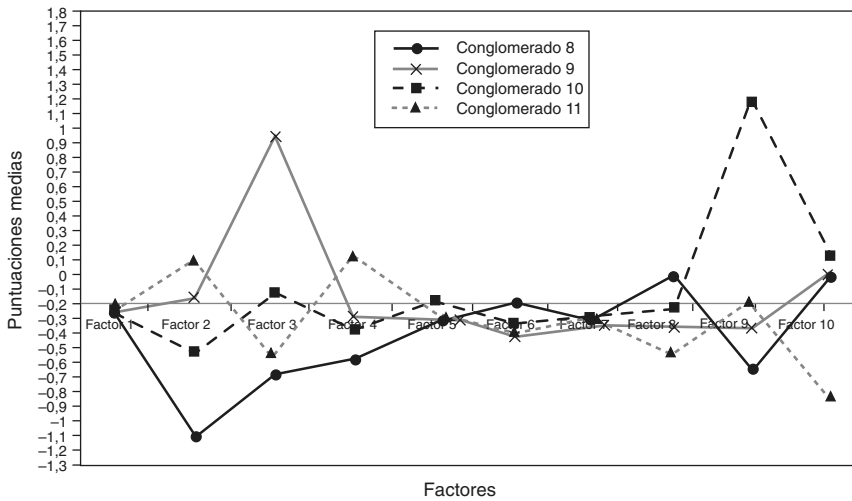


Gráfico 2



Los seis municipios que conforman el conglomerado 2, Jerez de la Frontera (Cádiz); Córdoba; Antequera (Málaga); y Carmona, Écija y Utrera (Sevilla), se caracterizan por tener una amplia extensión superficial. En cuanto a sus características socioeconómicas cabe destacar los valores medios de las puntuaciones que en ellos alcanzan los factores denominados como *Desarrollo agrícola* (factor 7, que presenta en este conglomerado el valor medio más alto, 6,91) y *Dimensión municipal* (factor 1, con una puntuación media de 2,11, el segundo en importancia a nivel de toda Andalucía). También se observan en este conglomerado puntuaciones medias aceptables

Gráfico 3



para los factores que representan al *dinamismo económico-demográfico* (factor 2) o al *nivel de educación* (factor 6) (0,22 y 0,23, respectivamente; en ambos casos, por encima de sus valores medios municipales para el conjunto de Andalucía). Con todo ello, podemos decir que se trata de grandes municipios en cuya economía tiene bastante importancia el sector de la Agricultura, ello no significa que sus habitantes vivan preferentemente de esta actividad ya que dichos niveles coexisten con unos altos niveles de equipamiento tanto de enseñanza como de sanidad; establecimientos hoteleros y otros tipos de actividades económicas no exclusivas del sector primario; buen nivel de renta por habitante así como de tasas de natalidad acompañadas de una población predominantemente joven pero instruida. La inclusión de Córdoba capital en este *cluster* queda justificada por el peso que supone su extensión superficial<sup>11</sup>. El *Dinamismo económico-demográfico* de estos municipios así como el nivel de formación de sus habitantes son superiores a la media andaluza.

El tercer conglomerado se caracteriza principalmente por una oferta turística importante, en términos relativos, dependiente plenamente del turismo rural y categorías hoteleras inferiores como hostales y pensiones. El *cluster* se conforma de la siguiente forma: ocho poblaciones se ubican en la zona de la Alpujarra granadina (Bubión, Calahorra (La), Capileira, Cortes y Graena, Juviles, Pampaneira, Pórtugos, Trevélez); la provincia de Cádiz aporta los municipios de Villaluenga del Rosario y Benaocaz y la provincia de Málaga los municipios de Atajate y Parauta, en conjunto forman parte de un mismo entorno orográfico constituido por la Sierra de Cádiz y la Serranía de Ronda, respectivamente. Finalmente, Bayárcal en Almería entra a formar parte del radio de acción de las Alpujarras granadinas. Son municipios con un nivel intermedio de actividad económica, con una elevada densidad de población y con

<sup>11</sup> Aunque la agricultura no sea su principal fuente de ingresos, posee extensiones significativas de cereales y olivos.

un nivel de estudios de segundo grado o superior importante dentro del contexto andaluz. Ello queda patente en el hecho de que los valores medios alcanzados por los factores 4 (*Actividad económica-intermedia*) y 6 (*Educativo*) en este conglomerado son superiores a los correspondientes al resto de los conglomerados en los que hemos dividido Andalucía. Por otra parte es de destacar que a pesar de ser pueblos de pequeño tamaño del interior de esta Comunidad ubicados en serranías<sup>12</sup>, no resulta relevante en este conglomerado el *factor Agrícola*.

A los conglomerados anteriores les siguen en importancia, por el número de municipios que aglutinan, los conglomerados 4, 5, 6 y 7 formados por 26, 52, 54 y 49 municipios, respectivamente.

En el gráfico 2 se pueden observar las puntuaciones medias de los factores en este grupo de conglomerados. La simple visión de dicho gráfico permite señalar que en el conglomerado 4 predomina el factor 6, por su alto valor medio en comparación con los correspondientes a los restantes factores. Este factor, al que hemos denominado *Educativo*, presenta en los municipios de este conglomerado sus valores más altos. Ello indica que el porcentaje de personas con dieciséis o más años que tienen estudios de grado medio o superior es en estas localidades muy superior a la media andaluza a nivel municipal. Por otra parte, en este grupo también presentan valores superiores a los medios municipales los factores que representan la *Dimensión municipal* (0,72), el *Dinamismo económico-demográfico* (0,21) o el *grado de Actividad agrícola* (0,39) de sus municipios.

De todo ello se puede concluir que los municipios que forman este conglomerado se caracterizan por una *Dimensión municipal* medio-alta, con una elevada densidad de población y un *Dinamismo económico-demográfico* también superior al medio, siendo importante en su economía la actividad agrícola.

Entre los municipios de este conglomerado se encuentran algunos de las provincias de Málaga y Sevilla; las cuatro capitales de provincia que presentan una *dimensión* medio-alta (Almería, Granada, Cádiz y Huelva) y algunos de sus municipios. En el Anexo se presentan los municipios que forman este conglomerado y en el mapa 2 se puede observar su situación geográfica.

En el conglomerado 5, al que pertenecen 52 municipios (ver Anexo\*), se alcanzan los valores más elevados del factor que representa el *Dinamismo económico-demográfico* (factor 2). Dichos municipios tienen, también, una elevada densidad de población y un porcentaje de mayores de dieciséis años con estudios de segundo grado o superior que es muy superior al medio municipal, lo que se puede concluir del valor medio que alcanza el factor 6 (*Educativo*) en este conglomerado. Asimismo, el porcentaje de analfabetos o sin estudios está muy por debajo del valor medio municipal andaluz (-0,46 de media para el factor 9, *Analfabetismo*). En contrapartida, en estas localida-

---

<sup>12</sup> En Cruces Pastor *et al.* (2006) se analiza a nivel municipal la oferta turística en Andalucía para los años 1993 y 2003, quedando patente el gran aumento experimentado por el turismo rural en esa década. Además en dicho artículo se pueden apreciar detalladamente los enclaves que destacan en este tipo de oferta turística.

\* El lector interesado en conocer la relación completa de los municipios que conforman cada uno de los conglomerados puede solicitarla directamente a los autores.

**Mapa 2.** Conglomerado 4



des el valor medio del factor denominado *Dimensión municipal* está por debajo de la media municipal andaluza ( $-0,21$ ) y, en conjunto, se encuentran en el penúltimo lugar de Andalucía con respecto a lo que se ha denominado *Desarrollo agrícola* ( $-0,59$ ). Es de destacar que la *Actividad agrícola* (medida por las variables que entran en el factor 3) y la *Actividad económica intermedia* (factor 4) presentan en las localidades de este conglomerado valores medios inferiores a la media municipal de Andalucía.

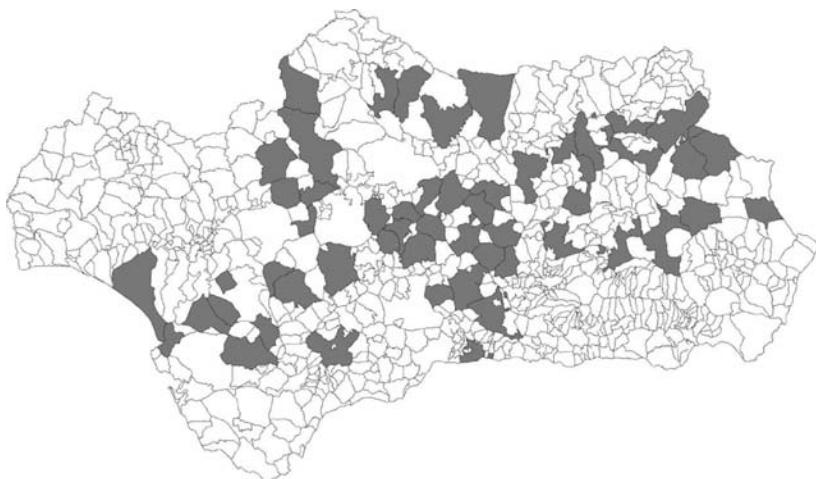
En el mapa que mostramos a continuación, se observa que se trata de municipios que en su mayoría pertenecen a las provincias de Sevilla y Granada y que están muy próximos a la capital.

**Mapa 3.** Conglomerado 5



Por otra parte, los municipios pertenecientes al conglomerado 6, cuya situación geográfica se puede visualizar con claridad en el mapa 4, se caracterizan por poseer valores muy superiores a la media en el factor 7 denominado *Desarrollo agrícola* que recoge, en términos absolutos, la extensión superficial del municipio y la dedicada a cultivo así como los trabajadores empleados agrarios subsidiados. Se trata de municipios con un elevado *Dinamismo económico-demográfico* (0,31), en los que el sector agrícola adquiere bastante importancia, tanto en términos absolutos (2,03) como en términos relativos (0,39) y con un nivel medio de la denominada *Actividad económica intermedia* que es, también, superior al medio municipal (0,22), presentando en conjunto valores para la denominada *Dimensión municipal* en torno al valor medio municipal andaluz (0,1).

**Mapa 4.** Conglomerado 6



En cuanto al último conglomerado de este bloque (conglomerado 7), está constituido por 49 municipios que se reparten mayoritariamente a lo largo de toda la costa andaluza, como se puede observar en el mapa 5. Hemos de destacar el elevado dinamismo poblacional de los municipios que lo integran, ya que es en ellos donde el factor 8 presenta las puntuaciones más altas<sup>13</sup>. Quizás por ello las provincias de Almería y Málaga engloban más de la mitad de sus municipios. Forman parte de este *cluster* municipios de gran atracción de extranjeros como El Ejido en Almería y Fuengirola, Torremolinos, Benalmádena, Nerja y Estepona en la costa malagueña<sup>14</sup>.

<sup>13</sup> Recordemos que este factor presenta correlaciones bastante altas con los porcentajes de extranjeros e inmigrantes así como con el incremento relativo de población en el periodo 1996-2007.

<sup>14</sup> De hecho estas provincias son las que mayor porcentaje de población extranjera poseen en Andalucía, en concreto, en 2006 la población extranjera representa en la provincia de Málaga el 13,7% de su población, casi nueve puntos porcentuales superior al que se registraba en 1998, y más del doble del peso que la población extranjera tiene por término medio en Andalucía (6,1%). Más elevado es aún el crecimiento experimentado por la población extranjera en la provincia de Almería representando esta población el 17,4% de la población total de la provincia. (Informe Económico de Andalucía 2006).

Mapa 5. Conglomerado 7



Se trata del grupo de municipios en los que el factor al que hemos denominado *Actividad económica intermedia* (factor 4) alcanza el valor medio más elevado (1,02). Aunque, como indican las puntuaciones medias obtenidas por los factores que reflejan el carácter agrícola del municipio ( $-0,19$  para el factor 3 y  $-0,21$  para el factor 7), son municipios en los que la actividad agrícola no es importante, presentan un elevado *Dinamismo económico-demográfico* (factor 2 con un valor medio de 0,28) y su nivel medio para la *Dimensión municipal* es superior al nivel medio municipal (0,19 puntos de media en el factor 1). En cuanto a las variables de oferta turística que quedan recogidas en el factor 5, los municipios de este conglomerado están ligeramente por debajo del valor medio municipal andaluz ( $-0,1$ ); sin embargo, como ya se ha comentado a lo largo de este trabajo, la importancia económica de la actividad turística en los municipios de este conglomerado queda reflejada en los elevados valores que en ellos presenta el factor 4, denominado *Actividad económica intermedia*, de ahí que destaquen en términos relativos el número de establecimientos con actividad económica entre los que se incluyen restaurantes, establecimientos hoteleros, etc. Son por tanto municipios dedicados al sector servicios, con un destacado peso turístico, entre los que cabe mencionar Roquetas de Mar (Almería), Almuñécar (Granada), Fuengirola o Torremolinos (Málaga).

El último grupo de conglomerados lo forman los que están representados, a través de las puntuaciones medias que en ellos alcanzan los factores, en el gráfico 3. Son los conglomerados 8, 9, 10 y 11, formados por 142, 145, 91 y 189; es decir, aquellos que aglutinan un mayor número de municipios. Como se puede observar en dicho gráfico o en la tabla 3, estos conglomerados se caracterizan por presentar puntuaciones medias inferiores a la media en casi todos los factores que, en este trabajo, resumen las características socioeconómicas municipales. En este sentido cabe señalar que de estos cuatro conglomerados es el conglomerado 8 el que presenta una menor *Dimensión municipal* y, también, un menor *Dinamismo económico-demográfico*. De



hecho, es en él donde se alcanza el menor valor medio para el factor así denominado (factor 2). Asimismo, los factores que representan a la *Actividad o desarrollo agrícola* presentan aquí también valores bastante inferiores al medio municipal ( $-0,59$  y  $-0,17$ , respectivamente) y lo mismo ocurre con la *Actividad económica intermedia* ( $-0,47$ ) o la *Oferta turística* ( $-0,18$ ). La mayoría de sus componentes se corresponden con municipios pertenecientes a las provincias de Almería y entornos cercanos de Jaén y Granada, fundamentalmente (mapa 6).

**Mapa 6.** Conglomerado 8



A diferencia de los municipios que conforman el conglomerado 8, en los pertenecientes al conglomerado 9 el *Dinamismo económico-demográfico* está al nivel medio municipal andaluz y la *Actividad agrícola* adquiere en ellos una especial relevancia, siendo este conglomerado el que presenta la puntuación media más alta del factor que la representa (1,33). La *Dimensión municipal* es, como en el conglomerado 8, inferior a la media municipal así como el grado de *Actividad económica intermedia* presentando, además, sus municipios un porcentaje de personas mayores de dieciséis años con estudios de segundo grado o superior que, a nivel medio, es de los más bajos de Andalucía. Los 145 municipios que forman este conglomerado (mapa 7) son en su gran mayoría de la zona del interior oriental de Andalucía, descartando la provincia de Almería.

Como se puede concluir de los comentarios anteriores son municipios típicamente rurales cuya principal fuente de ingresos es la agricultura y que presentan niveles bastante bajos del factor denominado *Dimensión municipal*.

Por último, los municipios de los conglomerados 10 y 11 presentan valores medios del factor *Dimensión municipal* muy similares y ligeramente inferiores a la media municipal. Eso mismo ocurre con el factor que representa el *Desarrollo agrícola* (factor 7). Sin embargo, mientras que los municipios englobados en el conglomerado

**Mapa 7.** Conglomerado 9



10 presentan una considerable tasa de analfabetismo<sup>15</sup> (la más alta de Andalucía con una puntuación media de 1,63), un *Dinamismo económico-demográfico* inferior al medio (-0,43) y un grado de *Actividad económica intermedia* inferior a la media (-0,24), con una importancia de la *Actividad agrícola* que es similar a la del municipio medio andaluz (0,05), los que conforman el conglomerado 11 presentan tasas de analfabetismo similares a la media municipal andaluza, niveles de *Dinamismo económico-poblacional* superiores al nivel medio municipal (0,32) y están también por encima de la media municipal andaluza en cuanto al grado de *Actividad económica intermedia* (0,35) por lo que no se encuentran en tan mala situación como el resto de los *cluster* representados en el gráfico 3. En el mapa 8 se muestran de modo conjunto ambos conglomerados<sup>16</sup>.

**Mapa 8.** Conglomerados 10 y 11



<sup>15</sup> El 18% de la población andaluza es analfabeta o no tiene estudios, una cifra que supera en 5,63 puntos la media nacional (Datos INE).

<sup>16</sup> En gris oscuro el conglomerado 10 y en gris claro el conglomerado 11.

## 6. Validación de los resultados

Para evaluar la bondad de la agrupación y analizar la capacidad de discriminación de los factores considerados se ha realizado un Análisis Discriminante, utilizando como valor de la variable dependiente para cada municipio el grupo al que el mismo pertenece de acuerdo con los resultados del Análisis *Cluster* realizado anteriormente y como variables explicativas los diez factores retenidos y utilizados en la clasificación.

De la información estadística básica relativa a dicho análisis (tabla 4) podemos concluir que los resultados son bastante satisfactorios desde el punto de vista estadístico, ya que los valores de la *lambda de Wilks*, tanto para el modelo final (0,00036) como los correspondientes a la participación de cada variable en el modelo (segunda columna de la tabla), son todos muy próximos a cero<sup>17</sup>. En cuanto a la contribución única de cada variable (factor) a la discriminación entre grupos (*lambda parcial*), las variables que más contribuyen son: el factor *Dimensión municipal* (0,25), seguido del *Desarrollo agrícola* (0,26) y de la *Oferta turística relativa* (0,38). Por su parte, los niveles de significación y los valores de tolerancia abundan en la significación estadística de los factores y en su no redundancia<sup>18</sup>.

**Tabla 4.** Análisis discriminante. Indicadores de validación

Núm. de variables en el modelo: 10 (11 grupos) Lambda de Wilks: 0,00036 approx. $F(100,5374) = 108,69$ $p < 0,0000$				
$N = 769$	Lambda de Wilks	Lambda Parcial	Nivel de significación	Tolerancia
Factor 1: <i>Dimensión municipal</i>	0,001453	0,250500	0,00	0,948981
Factor 2: <i>Dinamismo económico-demográfico</i>	0,000865	0,420598	0,00	0,928617
Factor 3: <i>Actividad agrícola</i>	0,000803	0,453192	0,00	0,951231
Factor 4: <i>Actividad económica intermedia</i>	0,000491	0,740996	0,00	0,903823
Factor 5: <i>Oferta turística relativa</i>	0,000963	0,377698	0,00	0,963382
Factor 6: <i>Educativo</i>	0,000737	0,493699	0,00	0,956340
Factor 7: <i>Desarrollo agrícola</i>	0,001396	0,260706	0,00	0,973887
Factor 8: <i>Movilidad poblacional</i>	0,000804	0,452633	0,00	0,925352
Factor 9: <i>Analfabetismo</i>	0,000636	0,571822	0,00	0,969625
Factor 10	0,000581	0,626799	0,00	0,935718

<sup>17</sup> La *lambda de Wilks* es un estadístico que expresa la significación estadística del poder de discriminación del modelo. Su valor está comprendido entre 0 y 1. Un valor de lambda igual a cero significa perfecto poder de discriminación.

<sup>18</sup> Para una variable  $x_i$  la tolerancia se define como  $1 - R_i^2$ , donde  $R_i^2$  es el coeficiente de determinación de la variable independiente  $x_i$  y el resto de las variables independientes que están seleccionadas en ese momento. Valores bajos en ella significan que la variable  $x_i$  puede ser considerada como una combinación lineal de las otras variables independientes y no debe entrar en el análisis.

Del análisis de la matriz de clasificación se concluye que, haciendo uso de las funciones clasificadoras resultantes del análisis, aproximadamente el 93% de los municipios se han clasificado correctamente, siendo los conglomerados 7 y 10 los que tienen un mayor porcentaje de municipios mal clasificados.

Por último, hemos de señalar que las cinco primeras funciones clasificadoras explican el 79% de la discriminación entre grupos y que en ellas pesan fuertemente la *Dimensión municipal* (-0,84), el *Desarrollo agrícola* (-0,76), la *Oferta turística relativa* (0,93), el *Dinamismo económico-demográfico* (-0,79) y la *Movilidad poblacional* (0,77), respectivamente.

## 7. Conclusiones

Los 769 municipios que conforman la Comunidad andaluza se han agrupado en once conglomerados que, atendiendo a las puntuaciones medias que en ellos alcanzan los factores, hemos subdividido en tres niveles: Conglomerados que tienen una mejor situación socioeconómica (*clusters* 1, 2 y 3), que han resultado ser los que aglutinan un menor número de municipios; Conglomerados que poseen una situación socio-económico media-alta (*clusters* 4, 5, 6 y 7); y por último Conglomerados con una mala situación socioeconómica (*clusters* 8, 9, 10 y 11) entre los que se encuentran aproximadamente el 74% de los municipios de Andalucía. Sin embargo, es de destacar que los 189 municipios del *cluster*, once no presentan una situación socioeconómica tan desfavorable como el resto de los componentes de este grupo.

En Andalucía oriental las disparidades socioeconómicas son más significativas que en la parte occidental dando lugar a una mayor aparición de conglomerados.

La zona costera de Andalucía junto con las capitales de provincia de Málaga y Sevilla, presenta buenos niveles en los indicadores socioeconómicos analizados. Si bien la costa oriental presenta mejor situación que la occidental.

Los municipios pequeños del interior de Andalucía muestran grandes deficiencias en los indicadores incluidos en este estudio, muy especialmente en los relativos a equipamientos e infraestructuras. La mayoría de estas localidades se encuentran ubicadas en la zona oriental de Andalucía formando parte de los *clusters* que resultan con una peor situación socioeconómica. A ello habría que añadir que en 91 de los 378 municipios que se encuentran en este grupo (los que constituyen el *cluster* 10) el factor que representa el *Analfabetismo* (factor 9) alcanza la mayor puntuación media (1,63) de toda Andalucía.

El sector primario sigue teniendo una gran importancia como fuente de ingresos en Andalucía, pero hay que destacar que no existe un prototipo único de actividad agrícola. Hemos detectado una agricultura más actualizada que genera otras actividades económicas que presentan los municipios recogidos en los *clusters* 2 y 6. Y por otro lado, una agricultura más tradicional, casi de subsistencia, de pueblos del interior incluidos en los *clusters* 9 y 10.

Por último, los resultados obtenidos manifiestan no sólo la heterogeneidad socioeconómica existente en la Comunidad andaluza sino también el hecho de que proximidad geográfica entre localidades no significa necesariamente proximidad en cuanto a nivel socioeconómico. Por todo ello queremos resaltar la importancia que el municipio adquiere en la elaboración de políticas de desarrollo regional, donde a menudo se apuesta por medidas de carácter «comarcal» que no recogen adecuadamente la especificidad de cada localidad.

## 8. Referencias bibliográficas

- Banco Mundial (1992): *Social indicators of development 1991-1992*, Johns Hopkins University Press, Baltimore and London.
- Borrás Palá, F. (2002): «Análisis y caracterización socioeconómica de las regiones de la Unión Europea mediante técnicas multidimensionales», *ICADE: Revista de las Facultades de Derecho y Ciencias Económicas y Empresariales*, 56, 307-322.
- Cano Torres, T.; Hernández Asensio, I., y Chasco Irigoyen, C. (2003): «Medición del bienestar social provincial a través de indicadores objetivos», *Anales de Economía Aplicada*, ISBN: 84-607-7655-7.
- Cruces Pastor, E. M.; Fernández Morales, A., y De Haro García, J. (2006): «Categorización territorial de la oferta turística de alojamientos reglados en Andalucía», *Papers de Turisme*, 39, 61-77.
- Cuadras, C. M. (1996): *Métodos de Análisis multivariantes*. EUB, S. L., Barcelona.
- Del Castillo Cuervo-Arango, F.; Fernández Huete, T.; Pedrazuela Frías, C., y Sacristán Moreno, R. (2006): «Tipología municipal de la Comunidad de Madrid 2006», *Consejería de Economía e Innovación Tecnológica. Instituto de Estadística de la Comunidad de Madrid*, Madrid.
- Everitt, B. S. (1993): *Cluster Análisis*, Wiley, New York.
- González López-Valcárcel (1991): *Análisis Multivariante. Aplicación al Ámbito Sanitario*, SG Editores, S. A., Barcelona.
- Gutiérrez Jáimez, R.; Abad Montes, F.; Román Román, P., y Torres Ruiz, F. (2000): «Tipificación, caracterización y clasificación de áreas territoriales andaluzas mediante técnicas estadísticas multivariantes», *Instituto de Estadística de Andalucía*, Sevilla.
- Hair, J. F.; Anderson, R. E.; Tatham, R. L., y Black, W. C. (2000): *Análisis Multivariante*, Prentice Hall Iberia, Madrid.
- Herrero Prieto, L. C.; Figueroa Arcila, V., y Sanz Lara, J. A. (2007): «Las disparidades territoriales en Castilla y León: estudio de la convergencia económica a nivel municipal», *Revista de Investigación Económica y Social de Castilla y León*, 9: 15-170.
- Hill, E. W.; Brennan, J. F., y Wolman, H. L. (1998): «What is a central city in the United States? Applying a statistical technique for developing taxonomies», *Urban Studies*, 35, 1935-1969.
- Hon, T. Y.; Poon, C. C., y Woo, K. Y. (2005): «Regional Distribution of Foreign Direct Investment in China, A Multivariate Data Analysis of Major Socioeconomic Variables», *The Chinese Economy*, 38, 56-87.
- Informe Económico de Andalucía (2006): Junta de Andalucía.
- INE (1999 y ss.): *Indicadores Sociales 1999 y siguientes*, INE, Madrid.
- Johnson, D. E. (2000): *Métodos multivariados aplicados al análisis de datos*, Internacional Thomson Editores, Madrid.
- Kaiser, H. F. (1960): «The Application of Electronic Computers to Factor Analysis», *Educational and Psychological Measurement*, vol. 10: 141-151.

- López Ruiz, E. (2005): «Un modelo para el análisis de las desigualdades económicas y sociales en un marco territorial: la provincia de Almería», *Revista de Estudios Regionales*, 74: 119-152.
- Morrison, F. D. (1987): *Multivariate Statistical Methods*, McGraw-Hill, New York, Book Co.
- OCDE (2002): *Society at a Glance OECD Social Indicators. Panorama de la société, les indicateurs sociaux de l'OCDE*, París.
- Pena López, J. A., y Sánchez Santos, J. M. (2008): «Disparidades económicas intrarregionales a escala municipal: Evidencia empírica para el caso gallego», *Revista de Estudios Regionales*, 81: 15-43.
- Peña, D. (2002): *Análisis de datos multivariantes*, McGraw Hill, Madrid.
- PNUD (1999): *1999, Human Development Report*, Oxford University Press, New York.
- Rúa Vieites, A.; Peralta Astudillo, M. J.; Fernández Rodríguez, L., y Borrás Palá, F. (2000): «Búsqueda de patrones socioeconómicos en la Unión Europea I y II: Análisis de componentes principales, factorial y conglomerados», *Actas del XXV Congreso Nacional de Estadística e Investigación Operativa*, 513-516.
- Rúa Vieites, A.; Redondo Palomo, R., y del Campo Campos, C. (2003): «Distribución municipal de la realidad socioeconómica gallega», *Revista Galega de Economía*, 12 (2): 1-20.
- Soares, J. O.; Marquês, M. M. L., y Monteiro, C. M. F. (2003): «A multivariate methodology to uncover regional disparities: A contribution to improve European Union and governmental decisions», *European Journal of Operational Research*, 145, 121-135.
- Stimson, R.; Baum, S.; Mullins, P., y O'Connor, K. (2001): «A typology of community opportunity and vulnerability in metropolitan Australia», *Papers in Regional Science*, 80 (1), 45-66.
- Uriel, E. (1995): *Análisis de datos. Series Temporales y Análisis Multivariante*, AC, Madrid.
- Vivo Molina, J. M.; Hermoso Gutiérrez, A., y Cano Guervós, R. (2004): «Factores latentes del desarrollo en los municipios de la región de Murcia», *XII Jornadas de ASEPUMA*.

## Anexo

Cuadro 1

<i>Variables</i>	<i>Rango</i>	<i>Media</i>	<i>CV</i>
<b>ES</b>	1.251,5	113,8	127,9
<b>ASM</b>	1.473,0	508,6	65,2
<b>PT</b>	699.095,0	10.477,2	375,4
<b>DP</b>	1.062.140,6	20.041,9	361,0
<b>P20</b>	28,1	20,1	22,5
<b>P64</b>	43,7	20,0	33,8
<b>IRP</b>	320,6	7,4	362,1
<b>PEX</b>	59,2	5,7	142,7
<b>PIN</b>	33,7	5,0	75,6
<b>PNV</b>	3,2	0,9	47,5
<b>PMAT</b>	1,4	0,4	60,7
<b>TANM</b>	34,0	8,2	51,1
<b>TANH</b>	25,0	3,7	65,0
<b>MES</b>	58,7	18,9	40,4
<b>HES</b>	56,6	18,8	45,1
<b>CEB</b>	208,0	3,7	339,7
<b>CES</b>	136,0	1,9	422,6
<b>CEA</b>	24,0	0,9	167,6
<b>CS</b>	32,0	0,5	360,9
<b>CSCONSPH</b>	2,0	0,1	146,2
<b>VFP</b>	226.597,0	3.142,2	388,9
<b>VFPPH</b>	44,2	34,8	19,2
<b>VT</b>	327.422,0	4.464,1	403,4
<b>OB</b>	767,0	9,2	429,1
<b>LT</b>	207.996,0	2.876,5	398,7
<b>LTPH</b>	70,3	27,6	22,6
<b>ADSL</b>	65.341,0	910,7	422,6
<b>ADSLPH</b>	27,8	4,9	89,8
<b>R</b>	454,0	10,6	362,0
<b>H</b>	91,0	2,0	339,3
<b>HP</b>	100,0	2,0	331,4
<b>HPPH</b>	1,1	0,0	271,8
<b>PHPPH</b>	19,7	0,7	253,6
<b>EHPH</b>	1,1	0,1	198,2
<b>PRM</b>	28.787,0	394,0	434,9
<b>PRH</b>	17.181,0	246,2	416,4
<b>CRM</b>	207.918,0	2.415,0	429,6
<b>CRH</b>	213.688,0	3.032,1	379,9

**Cuadro 1 (Continación)**

<i>Variables</i>	<i>Rango</i>	<i>Media</i>	<i>CV</i>
<b>CRI</b>	37.817,0	320,1	592,1
<b>OA</b>	77,8	22,5	68,0
<b>SC</b>	86.572,0	4.239,3	186,3
<b>SCCH</b>	100,0	34,0	90,6
<b>SA</b>	109,7	38,4	75,3
<b>TEASM</b>	1.457,0	135,2	150,7
<b>TEASMPH</b>	14,2	3,0	83,8
<b>TEASH</b>	841,0	73,1	142,2
<b>TEASHPH</b>	8,2	1,9	85,1
<b>EAE</b>	57.093,0	752,5	423,6
<b>EAEPH</b>	20,0	5,6	35,3
<b>EAE5</b>	17,1	4,8	33,0
<b>EAE20</b>	0,6	0,1	115,9
<b>RT</b>	5.494.288.900,0	56.046.012,5	499,8
<b>RTPH</b>	10.129,2	3.449,6	38,9
<b>RTRPH</b>	8.863,5	2.646,4	45,4
<b>RAEPH</b>	3.117,2	331,2	62,2
<b>RNMD</b>	20.282,1	10.902,9	28,4
<b>LAE</b>	57.968,0	882,8	382,2
<b>LAEPH</b>	20,2	7,6	34,8

**Figura 1**

Dendrograma para los 769 municipios  
Método de Ward  
Distancia Euclídea al cuadrado

