

**Análisis factorial y clúster que explica los factores urbanísticos que pueden estar determinando la localización de las empresas con actividades de innovación.  
Análisis para la RMB**

---

R e p o r t

Carlos Marmolejo Duarte  
Juan Eduardo Chica  
Zulema Conto  
Mabel Peña  
Oleguer Ribera  
Haytham Alhubashi  
Andrés Martínez de la Riva

Personal de recerca  
CPSV

Enero de 2012



UNIVERSITAT POLITÈCNICA  
DE CATALUNYA  
BARCELONATECH

Universitat Politècnica de Catalunya

---

Centre de Política de Sòl i Valoracions

UPC-CSPV

CPSV  
CENTRE  
DE POLÍTICA DEL SÒL  
I VALORACIONS

Este report es el resultado del trabajo desarrollado dentro del marco del Seminario de Investigación: ***Análisis de las características urbanísticas de las áreas urbanas con actividades económicas intensivas en conocimiento***, del Master en Gestión y Valoración Urbana 2011-2012.

## ÍNDICE

1. Análisis factorial
2. Análisis clúster

### 3. Análisis factorial

Para el análisis factorial se han introducido los indicadores de las variables estudiados juntamente con algunos datos preexistentes de las empresas estudiadas. En particular se incluye la fecha de constitución de la empresa, el código SIC, los ingresos, la media de empleados del período 2000-2009 y si importa o exporta.

Para el código SIC se da el número uno a las industrias de baja tecnología, el dos para las industrias de alta tecnología, el número tres para los servicios de bajo conocimiento y el número cuatro para los servicios de alto conocimiento.

La totalidad de variables se han factorizado con rotación, y para un número de factores por defecto, que se ajusta a cuatro. Se repite la prueba para un número de factores de dos, de tres y de cinco. Se observa que la pérdida de información representada para los casos de dos y tres factores es notable. Así mismo, para cinco factores la ganancia de información se entiende que no compensa la complicación de lectura de los factores. De esta forma se estima la factorialización en cuatro elementos como la óptima. El grado de extracción en este caso es de casi el 82%. (Tabla 17).

En un primer momento también se añaden las variables de accesibilidad general, y el código CNAE, que finalmente se retiran por considerar que no aportan información suficientemente relevante.

En la tabla adjunta se observa el grado de extracción de cada variable, se puede observar que el grado extraído para las variables definitorias de las parcelas es bastante alto, entorno al 90%, excepto el de las propias topologías que se reduce al 75%. Para las variables definitorias de las empresas el grado de extracción disminuye, en particular para los ingresos donde solo alcanza el 59%. (Tabla 16)

Tabla 16.			Tabla 17					
	Inicial	Extracción	Componente	Sumas de las saturaciones al cuadrado de la extracción		Suma de las saturaciones al cuadrado de la rotación		
				% de la varianza	% acumulado	Total	% de la varianza	% acumulado
Fecha de constitución	1,000	,706	1	46,446	46,446	6,059	43,280	43,280
CODIGO_SIC	1,000	,770	2	17,260	63,705	2,009	14,347	57,627
INGRESOS	1,000	,595	3	11,864	75,569	1,983	14,162	71,788
MEDIA EMPLEADOS 2000-2009	1,000	,732	4	6,291	81,860	1,410	10,072	81,860
Importa / Exporta	1,000	,628	5					
TIPOLOGIAS	1,000	,753	6					
la media de la Accesibilidad en transporte privado	1,000	,914	7					
la media Accesibilidad en transporte público	1,000	,878	8					
Ratio servicios personas / empresas	1,000	,830	9					
Estado de conservacion	1,000	,887	10					
Diversidad de uso en parcela	1,000	,945	11					
Tipo de edificacion	1,000	,932	12					
Calidad de espacios libres	1,000	,939	13					
Interacción social	1,000	,952	14					

A la vista de la matriz de componentes (Tabla 18), se puede analizar los factores resultantes y su composición.

Tabla 18. Matriz de componentes					Tabla 19. Matriz de componentes rotados.				
	Componente					Componente			
	1	2	3	4		1	2	3	4
Fecha de constitución	,062	,812	-,108	,179	Fecha de constitución	-,130	,755	,116	-,326
CODIGO_SIC	,335	,655	-,081	,472	CODIGO_SIC	,151	,859	,092	-,018
INGRESOS	-,068	-,521	,364	,431	INGRESOS	-,021	-,208	,050	,741
MEDIA EMPLEADOS 2000-2009	-,035	-,553	,325	,565	MEDIA EMPLEADOS 2000-2009	,020	-,140	-,017	,844
Importa / Exporta	,387	,618	-,167	,261	Importa / Exporta	,235	,728	,068	-,194
TIPOLOGIAS	,845	-,188	,014	-,053	TIPOLOGIAS	,848	-,013	,180	,039
la media de la Accesibilidad en transporte privado	-,877	,185	,329	-,050	la media de la Accesibilidad en transporte privado	-,939	-,108	,141	,016
la media Accesibilidad en transporte público	,935	-,011	,055	-,024	la media Accesibilidad en transporte público	,880	,152	,284	-,017
Ratio servicios personas / empresas	-,908	,025	-,028	,056	Ratio servicios personas / empresas	-,864	-,121	-,258	,044
Estado de conservacion	,474	,353	,728	-,094	Estado de conservacion	,225	,199	,893	,012
Diversidad de uso en parcela	,930	-,231	-,163	-,020	Diversidad de uso en parcela	,971	,017	,023	,017
Tipo de edificacion	,960	-,101	-,029	-,023	Tipo de edificacion	,942	,101	,189	-,002
Calidad de espacios libres	,376	,280	,828	-,186	Calidad de espacios libres	,134	,054	,958	,020
Interacción social	,947	-,180	-,151	,009	Interacción social	,971	,076	,047	,017

Cuando analizamos la matriz de componentes rotados, (Tabla 19), se confirman las intuiciones que observábamos en la matriz sin rotar.

El Factor uno explica muy bien las topologías y la mayoría de sus variables, la accesibilidad, tanto pública como privada, el indicador de servicios, la diversidad de usos, el tipo de edificación y la interacción social. No explica tan bien el estado de conservación ni la calidad de espacios libres y prácticamente no incluye información sobre las características de las empresas. El factor se relaciona negativamente con la accesibilidad de transporte privado y con la ratio de servicios.

El factor dos se refiere fundamentalmente a las características de las empresas, da cuenta del código SIC, seguidamente la fecha de constitución en primer lugar y después de la importación.

El tercer factor tiene carácter de urbanización, explica básicamente las variables de la calidad de los espacios libres y del estado de conservación.

El último factor, el número cuatro también se refiere, básicamente, a las empresas. Explica la media de empleados y los ingresos. Se podría entender que da cuenta del volumen de la empresa.

Podemos renombrar estos factores según los criterios anteriormente destacados como:

- Factor de centralidad urbana.
- Factor de empresas de alta calidad.
- Factor de espacios urbanos de calidad.
- Factor de grandes empresas.

#### 4. Análisis clúster

Según lo observado en el análisis factorial consideramos oportuno realizar un análisis cluster con el objetivo de ver si hay alguna tendencia de agrupación de tipos de empresas según los factores anteriormente anunciado. Para ello realizamos un análisis clúster con cuatro aglomerados de los las empresas en base a los factores anteriores, obteniendo un aglomerado para cada una de las empresas.

Tabla 20. Centros de los conglomerados finales

	Conglomerado			
	1	2	3	4
REGR factor score 1 for analysis 1	,14962	-,48390	-,15307	,97064
REGR factor score 2 for analysis 1	-,43511	-,75693	,44671	,78215
REGR factor score 3 for analysis 1	,12625	-,51373	1,10631	-,63199
REGR factor score 4 for analysis 1	3,50082	-,26662	-,17788	-,19664

Lo que haremos para comprobar el tipo de aglomerados es realizar histogramas para saber que tipo de empresas se sitúan en cada conglomerado, fijándonos en el tipo de empresa y en la tipología en la que se sitúa. Resumidamente podemos observar la distribución de las empresas en base a estos conglomerados.

Tabla 21. Número de casos en cada conglomerado

Conglomerado	1	15,000
	2	100,000
	3	79,000
	4	60,000
Válidos		254,000
Perdidos		,000

Para el conglomerado 1 al que pertenecen 15 empresas obtenemos las siguientes distribuciones por tipo de empresa y por tipología de trama urbana sobre la que se asienta, ilustradas en los tabla 22.

Conglomerado 1.	CODIGO_SIC				Total	%
	Industria baja tecnología	Industria alta tecnología	Servicios bajo nivel	Servicios alto nivel		
TIPOLOGIAS						
Industrial antigua	1	0	0	0	1	7%
Industrial nueva	2	1	1	0	4	27%
Ensanche	4	0	1	1	6	40%
Ensanche de nueva creación	2	1	0	0	3	20%
Regeneración urbana	0	0	0	1	1	7%
Grand Total	9	2	2	2	15	100%
%	60%	13%	13%	13%	100%	

En el conglomerado 1 se sitúan básicamente empresas de industria de bajo nivel pero contradictoriamente situadas en el ensanche. Si analizamos los datos de número de empleados y facturación obtenemos datos que puede explicar esta situación.

Estas 15 empresas suman 25.012 empleados del total de los 45.538 de los recontados dentro de la base de datos y facturan según los datos acumulados un total de 3.919.331.287 euros de los 6.758.690.317. En ambos casos suponen el 55% en número de empleados y el 58% en el caso de la facturación, teniendo en cuenta que estas 15 empresas representan apenas el 6% del total podemos asimilar que este conglomerado se relaciona directamente con el factor de grandes empresas, además observando la tabla 20, este conglomerado absorbe mucho de las empresas dentro de este factor.

Analizando el conglomerado 2 obtenemos lo siguiente,

TIPOLOGIAS	CODIGO_SIC				Total	% Codigo
	Industria baja tecnología	Industria alta tecnología	Servicios bajo nivel	Servicios alto nivel		
Industrial antigua	24	3	8	5	40	40%
Industrial nueva	19	2	5	0	26	26%
Polo dependiente	1	0	0	0	1	1%
Ensanche	24	1	1	0	26	26%
Regeneración urbana	7	0	0	0	7	7%
Total	75	6	14	5	100	100%
% Tipologías	75%	6%	14%	5%	100%	

Según lo analizado en la tabla 23 podemos deducir que estas empresas se pueden considerar dentro de un conglomerado que agrupa empresas de baja tecnología, con cierta tendencia a situarse en zonas industriales o en todo caso de ensanche.

Se caracterizarían estas empresas por ser relativamente antiguas, siendo su media de fundación en 1978 y con un número de empleados medio/bajo con 123 empleados de media.

Podemos asimilar este conglomerado a empresas dedicadas al sector industrial o en contraste con el factor 2 que es al que se contrapone en la tabla 20, podemos denominarlo de baja calidad.

Si nos centramos en el conglomerado número 3 observando la tabla 24:

TIPOLOGIAS	CODIGO_SIC				Total	% Cod.
	Industria baja tecnología	Industria alta tecnología	Servicios bajo nivel	Servicios alto nivel		
Polos autónomos	4	0	2	9	15	19%
Industrial nueva	0	0	0	3	3	4%
Polo dependiente	4	0	3	21	28	35%
Ensanche de nueva creación	6	0	4	8	18	23%
Regeneración urbana	2	0	4	9	15	19%
Total	16	0	13	50	79	100%
% Tipologías	20%	0%	16%	63%	100%	

Este conglomerado agrupa a empresas que se sitúan mayoritariamente en polos dependientes, pero no de manera clara, ya que se equidistribuyen en otras tipologías como regeneración urbana, ensanche de nueva

creación y polos autónomos. Cabe resaltar que no hay ninguna empresa del conglomerado 3 dentro de la tipología de ensanche. Por otra parte en cuanto al tipo de empresa destacan las de servicios de alto nivel. Se puede concluir que en este conglomerado se agrupan las empresas de alta tecnología situadas que se sitúan en ámbitos urbanos de alta calidad.

Por último analizando el conglomerado número 4:

TIPOLOGIAS	CODIGO_SIC				Total	%
	Industria baja tecnología	Industria alta tecnología	Servicios bajo nivel	Servicios alto nivel		
Industrial antigua	0	0	0	1	1	2%
Casco antiguo	3	0	2	5	10	17%
Ensanche	2	2	7	38	49	82%
Total	5	2	9	44	60	100%
%	8%	3%	15%	73%	100%	

El conglomerado número cuatro agrupa a empresas fundamentalmente de servicios de alto nivel situadas en el ensanche en mayor medida y en el casco antiguo. Se puede asimilar estas empresas al factor de centralidad urbana que se detectó en el análisis factorial.

Como resumen exponemos la tabla 26 con los valores conglomerados renombrados:

Tabla 26. Número de casos en cada conglomerado		
Conglomerado	Grandes empresas	15,000
	Empresas industriales	100,000
	Servicios de alto nivel en espacios urbanos de alta calidad	79,000
	Servicios de alto nivel con centralidad urbana	60,000
Total		254,000