

# Estrategia de análisis de datos multivariados para caracterizar Mipymes

## Strategy of variable data analyses to characterize mipymes

Valery José Lancheros Suárez<sup>1\*</sup>, Ángel León González Ariza<sup>2</sup>

<sup>1</sup> *Departamento de Ingeniería Mecánica, Universidad de Córdoba, Montería – Colombia. \* valancheros@hotmail.com*

<sup>2</sup> *Departamento de Ingeniería Industrial, Universidad del Norte, Barranquilla – Colombia. agonzalez@uinorte.edu.co*

*Fecha de recepción del artículo:*

*: Fecha de aceptación del artículo:*

### Resumen

Este artículo contiene una estrategia de análisis de datos multivariados para caracterizar Mipymes. Se aplicó en el sector alimenticio de la ciudad de Montería, pero puede emplearse en otros sectores económicos en los que se desee conocer sus características. Uno de los beneficios de esta aplicación, es poder comparar las empresas con sus homólogas a nivel nacional e internacional, para tener argumentos que permitan determinar su permanencia en el tiempo a pesar de enfrentarse a los grandes retos que se avecinan en materia de estrategias comerciales como tratados de libre comercio con otros países y las demás políticas que el Estado colombiano tenga proyectadas. Osorio, B. (2009) [1].

Se seleccionó la industria alimenticia (Código CIU D150000) como objeto de estudio de esta investigación por su importancia y representatividad, ya que en la base de datos de la Cámara de Comercio de Montería aparecen 12.497 empresas registradas en el departamento de Córdoba, de las cuales 6.307 tienen domicilio en la ciudad de Montería y un buen número de ellas están relacionadas con dicho sector.

El punto de partida de este trabajo es un estudio del grado de difusión e implementación de normas técnicas relacionadas con sistemas de gestión organizacionales en las Mipymes del sector alimenticio de Montería, que presenta un diagnóstico del estado de las empresas de este

sector desde la óptica de la calidad, en donde los datos recolectados se procesaron a través de métodos estadísticos descriptivos, por lo que se consideró importante desarrollar una segunda parte correspondiente a la aplicación del análisis de datos multivariados donde se utilizó la Estadística Inferencial. Sánchez, J. (2007) [2].

### Palabras clave

Métodos multivariados, caracterización, Mipymes, sector alimentos, diagnóstico.

### Abstract

This article contains a strategy of variable data analyses to characterize Mipymes, was applied in the nutritional sector of the city of Montería, but it can be used in other economic sectors in which it is desired to know its characteristics. One of the benefits of this application, is to be able to compare the companies with its counterparts at national and international level, to have arguments that allow to determine their permanence in the time in spite of facing the great challenges that are approached in the matter of commercial strategies as Free Trade Agreement with other countries and the other policies that the Colombian State has projected. Osorio, B. (2009) [1].

The nutritional industry was selected (Code CIU D150000) as object of study of this investigation

by their importance and representativeness, since in the data base of the Chamber of Commerce of Montería, appear 12497 companies registered in the department of Cordoba, of which 6307 have address in the city of Montería and a good number of them is related to this sector.

The departure point of this work is a study of the diffusion degree and implementation of practical standards related to organizational Systems of Management in the Mipymes of the nutritional Sector of Montería, that presents a diagnosis of the state of the companies of this sector from the optics of the quality, where the collected data were processed through descriptive Statistical methods, reason why it was considered important to develop one second part corresponding to the application of the variable data analysis where Inferencial Statistics was used. Sánchez, J. (2007) [2].

## Key Words

Variable methods, characterization, mipymes, sector foods, diagnosis.

## Introducción

En nuestra región y específicamente en nuestras localidades o municipalidades, no se está lo suficientemente preparado para afrontar el impacto producido por los rápidos cambios que traen consigo los procesos actuales que se caracterizan por la dinámica con que se desarrollan. Montería, como municipio y ciudad, no es la excepción a esta preocupación, puesto que es una ciudad que no ha podido entrar a cumplir de manera eficiente el papel de satisfacer las necesidades básicas de la población, a pesar de contar con riquezas naturales y un recurso humano de cualidades excepcionales.

La ciudad se encuentra afectada por la problemática de su limitado desarrollo económico, social, político y tecnológico, y esto incide en la calidad de vida y en una manifiesta diversidad de problemas de toda índole, pero que se pueden expresar para el tema que nos ocupa en:

- Falta de un diagnóstico de las Mipymes de los diferentes sectores económicos de la ciudad de Montería, que permitan plantear alternativas de solución a sus problemas.
- La carencia de información sistemática de las Mipymes del municipio, para organizarse en gremios y conocer la forma de enfrentar su problemática.
- La falta de políticas, programas y planes que le permitan una gestión responsable y posibilite un desarrollo de la comunidad.

En la base de datos 2009 de la Cámara de Comercio de Montería, en Córdoba se encuentran registradas 12.497 empresas de diferentes sectores económicos, de las cuales 6.307 tienen domicilio en la ciudad de Montería. De estas empresas no se tienen cifras y datos consolidados ni resultados de estudios que muestren entre otras cosas, información relacionada con las características de cada sector económico, de tal manera que no es posible establecer si por sus características, son consideradas competitivas y si en efecto eso mismo les ha permitido perdurar en el tiempo.

Lo descrito en el párrafo anterior condujo a un grupo de docentes de la Universidad de Córdoba a querer contribuir con el desarrollo de la región, quienes pensaron en ofrecer capacitación y consultoría en los temas presentados en una línea de acción del Plan de Desarrollo de Montería, pero para realizarlo había que consultar el estado actual del comercio en la ciudad con el fin de identificar claramente las necesidades inmediatas de los empresarios de la capital cordobesa en el aspecto antes mencionado. Lancheros, V. (2008) [3]. En tal sentido es válido preguntarse: ¿Es posible, a través de la aplicación del análisis de datos multivariado caracterizar las Mipymes del sector alimenticio de la ciudad de Montería?

Los aspectos antes mencionados indujeron a desarrollar este trabajo de investigación con el objetivo de aplicar una estrategia de análisis de datos multivariados, mediante la caracterización de las Mipymes del sector alimenticio de la ciudad de Montería, orientada al diagnóstico del

estado actual del sector. Para tal fin se realizó una revisión bibliográfica de los diferentes métodos multivariados que podrían usarse en la estrategia, su delimitación de acuerdo a algunas herramientas informáticas y los tipos de variables que finalmente permitirían tener las características planteadas para la caracterización.

## Marco teórico y estado del arte

El estado del arte relacionado con las Mipymes de la ciudad de Montería es mínimo ya que los trabajos que se encuentran se caracterizan por presentar información nacional referida al panorama mundial de la certificación frente a normas de gestión, tal como lo expresa un trabajo realizado por el Icontec denominado el estudio ISO 2005. También es conocido otro resultado del Centro Nacional de Productividad, que consiste en determinar el Impacto de la Certificación de Sistemas de Gestión de la Calidad en las Empresas Colombianas, el cual corresponde a la primera investigación realizada en Colombia para establecer lo ocurrido directamente en las empresas por lo relacionado con este hecho.

## Métodos multivariados y su clasificación

Es el conjunto de métodos estadísticos cuya finalidad es analizar simultáneamente conjuntos de datos multivariados en el sentido de que hay varias variables medidas para cada individuo u objeto estudiado. Su razón de ser radica en un mejor entendimiento del fenómeno a estudiar obteniendo información que los métodos estadísticos univariantes y bivariantes son incapaces de conseguir. Figueras, S. (2000) [4].

Los análisis multivariados empleados en el estudio fueron:

### *Análisis de Componentes Principales (ACP)* *Gurrea, M. (s.f.). [5]*

El Análisis de Componentes Principales (ACP) es una técnica estadística de síntesis de la información, o reducción de la dimensión (número

de variables). Es decir, ante un banco de datos con muchas variables, el objetivo será reducirlas a un menor número perdiendo la menor cantidad de información posible. En notación abreviada la forma de calcular cada componente se define por  $y_i = a'X = \mathbf{X}$ . El total de componentes está dado por:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{1,p} \end{bmatrix} \begin{bmatrix} X_1 & X_1 & \dots & X_{1,p} \\ X_2 & X_2 & \dots & X_{2,p} \\ \vdots & \vdots & \dots & \vdots \\ X_{n1} & X_{n2} & \dots & X_p \end{bmatrix}$$

Cabe destacar que en algunos casos para la extracción de los componentes es necesario verificar la normalización del vector  $a$ , el cual está dado por  $a_i^2 a_i = 1$  o en su defecto la ortogonalidad dada por  $a_i^2 a_j = 0$ .

Un aspecto clave en ACP es la interpretación de los factores, ya que esta no viene dada *a priori*, sino que será deducida tras observar la relación de los factores con las variables iniciales (habrá, pues, que estudiar tanto el signo como la magnitud de las correlaciones).

Es de suma importancia en el análisis de componentes principales, estimar la calificación de las variables de cada uno de ellos, lo cual se puede realizar por  $y_{ji} = \hat{a}(x_r - \mu)$  para  $j = 1, 2, \dots, p$  y  $r = 1, 2, \dots, N$ .

### *Análisis Factorial Olmos, S. (2004). [6]*

El Análisis Factorial (AF) tiene como objetivo encontrar una estructura más simple reduciendo la dimensionalidad de las variables sin perder información. Para simplificar el análisis de los datos se reduce el número de variables a un pequeño número de índices o factores. Algunas diferencias entre estas dos técnicas son que las componentes principales están definidas como una combinación lineal de las variables originales y no están basadas en un modelo estadístico particular y por lo tanto no se requiere el cumplimiento de supuestos previos.

En el AF las variables están expresadas como una combinación de factores, está basado en un modelo

especial y requiere el cumplimiento de distintos supuestos. Por otra parte, mediante el ACP se busca explicar una gran parte de la varianza total, mientras que con el AF se enfatiza el estudio en las relaciones entre las variables explicadas con las covarianzas o correlaciones. El AF resulta apropiado cuando el objetivo consiste en encontrar un grupo de variables similares, altamente correlacionadas, y postular que esas similitudes provienen del hecho de que estas son variables «latentes o factores» que actúan en forma particular sobre el proceso estudiado.

Cuando se tiene un vector  $\mathbf{x}$  de respuestas que van desde 1 hasta  $p$  con media  $\boldsymbol{\mu}$ , se puede construir un modelo del Análisis Factorial de la siguiente manera:

$$X_j = \lambda_{j1}f_1 + \lambda_{j2}f_2 + \dots + \lambda_{jm}f_m + \eta_j,$$

para  $j = 1, 2, \dots, p$

El modelo también puede expresarse en forma matricial tal como se indica a continuación:

$$X = \Lambda f + \eta$$

Donde:

$$X = [x_1, x_2, \dots, x_p]$$

$$f = [f_1, f_2, \dots, f_m]$$

$$\eta = [\eta_1, \eta_2, \dots, \eta_p]$$

$$\Lambda = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \dots & \lambda_{1m} \\ \lambda_{21} & \lambda_{22} & \dots & \lambda_{2m} \\ \dots & \dots & \dots & \dots \\ \lambda_{p1} & \lambda_{p2} & \dots & \lambda_{pm} \end{bmatrix}$$

## Análisis de conglomerados (clúster)

De acuerdo con Peña, D. (2002) [7] el propósito del análisis de conglomerados (clúster en terminología inglesa) es agrupar las observaciones de forma que los datos sean muy homogéneos dentro de los

grupos (mínima varianza) y que estos grupos sean lo más heterogéneos posible entre ellos (máxima varianza). De este modo obtenemos una clasificación de los datos multivariante con la que podemos comprender mejor los mismos y la población de la que proceden. Podemos realizar análisis clúster de casos, de variables o por bloques si agrupamos variables y casos. Romero, R. (2003). [8]

Según Pérez, C. (2004) [9] en el análisis de conglomerado se explican las relaciones entre variables y casos de acuerdo a la similitud o distancia entre ellos: si el valor de la distancia es alto, más alejada es su relación. Existen diferentes formas de calcular las distancias, una de ellas es la distancia Euclídea, la cual se indica a continuación:

$$d_{ij} = \sqrt{\sum_{k=1}^t (X_{ik} - X_{jk})^2}$$

## Población

La población objeto de estudio en este trabajo de investigación, corresponde a 97 empresas de alimentos registradas en el 2009 en la base de datos de la Cámara de Comercio de la ciudad de Montería. El sector alimenticio a estudiar, está conformado por empresas que se encargan de:

Fabricación de productos lácteos; fabricación de aceites, grasas vegetales y animales; matanzas de ganado, preparación y conservación de carnes; productos de molinería, productos de panadería; productos alimenticios diversos; elaboración de alimentos preparados para animales; industria de bebidas no alcohólicas, aguas y gaseosas.

## Variables a medir u observar

Para recolectar la información se diseñó una encuesta con nueve ejes temáticos o variables conformados por un grupo de preguntas que siguieron un patrón de operacionalización, descritos a continuación:

- Participación en agremiaciones
- Recurso humano

- Formación académica del personal
- Gestión y administración
- Producción
- Ventas y exportaciones
- Capacitación
- Recurso tecnológico
- Costos de funcionamiento

## Trabajo de campo y resultados

El trabajo de campo inició con el cálculo de una muestra que dio como resultado un total de 43 establecimientos que se distribuyeron por zonas para recolectar la información por medio de la aplicación de una encuesta. La distribución se indica en la Tabla 1.

**Tabla 1.** Distribución de empresas por zonas

Zona	N° de Empresas
Centro	28
Sur	8
Industrial	2
Margen izquierda	4
Vía a Planeta Rica	1

Fuente: Elaboración propia.

## Análisis de la base de datos

El primer análisis se realizó en la base de datos de la Cámara de Comercio de Montería en la que se encontró que hay treinta y nueve (39) empresas con menos de diez (10) empleados, de las cuales cinco (5) tienen un valor comercial de más de quinientos

(500) smlmv y uno con un patrimonio de más de cinco mil (5.000) smlmv. Además se encontraron tres empresas que tienen entre 11 y 50 personas vinculadas y dos de ellas tienen un valor comercial por debajo de los quinientos uno (501) smlmv y una empresa que no clasifica como Mipymes, en razón a que la Ley 905 del 2004 la categoriza como gran empresa ya que tiene más de doscientos trabajadores y cuenta con un patrimonio de más de treinta mil (30.000) smlmv. Al reorganizar la información anterior se obtuvo la Tabla 2.

**Tabla 2.** Número de empresas por categoría

N° de Empresas	Categoría
1	Mediana
5	Pequeña
36	Micro
1	N.A

Fuente: Elaboración propia.

Otro aspecto importante de los primeros análisis fue el de establecer el comportamiento de las empresas cordobesas, durante los años 2007, 2008 y 2009. La Tabla 3 muestra que en los últimos tres años, ha habido un aumento en el número de establecimientos registrados en Cámara de Comercio, sin embargo, el comportamiento de las empresas del sector alimentos decreció ligeramente ya que pasó de 100 a 97 unidades.

## Análisis de la encuesta

En la encuesta aplicada, se encontró que 29 de las 43 empresas encuestadas son productoras de

**Tabla 3.** Constitución de empresas por año

Año	Total	Ubicación	Total
2007	9401	Montería	5006
		Sector alimentos	91
2008	11139	Montería	5671
		Sector alimentos	100
2009	12497	Montería	6307
		Sector alimentos	97

Fuente: Elaboración propia.

tipo familiar a las que se le adicionan otras seis productoras pero que no fue posible identificar su tipo ya que los encuestados no respondieron. Así mismo, se encontraron dos empresas de servicio tipo asociada, dos que no especificaron la actividad pero respondieron que son tipo familiar y cuatro que no respondieron ni la actividad ni el tipo de empresa.

El primer eje temático correspondiente a participación en agremiaciones, presentó solo una empresa que hace parte de una asociación y ninguna de ellas ha recibido apoyo por parte de gremios o del Estado.

En lo correspondiente a recurso humano se encontró un predominio del empleo directo, ya que de las 43 empresas encuestadas, solo dos de ellas subcontratan empleados. Así mismo se comprobó que cinco de ellas tienen entre 11 y 50 empleados. El personal se encuentra ubicado en su gran mayoría en el área de ventas y producción.

La formación académica presentó que las empresas vinculan personal técnico y bachiller mientras que tres empresas cuentan con tecnólogos y solo dos de las 43 tienen personal profesional. La característica principal es que los bachilleres se encuentran ubicados en todas las áreas de las empresas.

Lo relacionado con la gestión, evidenció que la gran mayoría de las empresas no tienen alianzas estratégicas ya que solamente dos de ellas tienen esta relación con sus proveedores. También se encontró que 36 establecimientos manejan la información de la gestión manualmente y que sus principales problemas en esta parte se deben a la falta de recursos.

En cuanto a la producción se halló que la consecución de la materia prima se hace a nivel nacional e incluso regional y la tecnología usada en el proceso es electromecánica de operación manual, en la mayoría de los casos.

Los resultados en ventas y exportaciones expresaron que la mayoría de las empresas son competitivas, aunque seis de ellas dijeron ser altamente competitivas y tres poco competitivas.

Se encontró también que no hay una sola empresa que exporte sus productos, mientras que siete de ellas venden tanto a nivel local como regional. Así mismo manifestaron desconocer cuánto aportan al producto interno bruto (PIB).

En las Mipymes del sector alimenticio de Montería, se encontró que hay 8 empresas que tienen dependencias encargadas de la capacitación, que el tipo de capacitación que se imparte con mayor frecuencia son cursos de actualización según tendencias y cursos específicos del área de trabajo. Dijeron que cuando realizan las capacitaciones las orientan a público general y son impartidas por sus proveedores, aunque 29 de ellas no han realizado capacitación en los últimos dos años.

De las 43 Mipymes, 38 cuentan con recursos tecnológicos con menos de diez años de uso y cinco de ellas con más de diez años a los que dicen les aplican mantenimiento preventivo pero la implementación de nuevas tecnologías, está limitada por los costos que ellas generan.

En cuanto a los costos de funcionamiento, 16 de ellas manifestaron que está comprendido entre uno y 25 millones de pesos al año y el resto no sabe o no responde, también respondieron que el presupuesto anual es preparado por las áreas de contabilidad y finanzas o por el representante legal.

## **Resultados de los métodos multivariados**

Para realizar la caracterización planteada como objetivo general del presente proyecto de investigación era necesario encontrar las características de las empresas, por medio de la aplicación de técnicas dirigidas por variables, en razón a la necesidad de establecer las relaciones entre las variables medidas en las Mipymes seleccionadas en la muestra.

Según Johnson, D. (2000) [10] las técnicas por variables se encuentran en los análisis realizados sobre las matrices de correlación, el análisis de componentes principales, el análisis por factores, el análisis de

regresión y el análisis de correlación canónica, sin embargo se aplicó análisis de conglomerado en razón a la necesidad de agrupar las variables, a pesar de que este último es un método considerado por el mismo autor como una técnica dirigida por los individuos o unidades experimentales.

En primera instancia se realizó un Análisis de Componentes Principales (ACP) con el fin de presentar una combinación de las nueve variables para explicar la mayor variabilidad de los datos. En esta parte del análisis de datos se consideró importante calcular los Eigenvalores dados por  $|\Sigma - \lambda I| = 0$  para seleccionar aquellos que presentaran valores mayores de 1 para determinar el número de componentes principales.

Una vez se cargaron los datos en Statgraphics, se obtuvieron los resultados mostrados en la Tabla

2, en la que los cuatro primeros componentes presentaron valores mayores de 1 y en conjunto explican el 71,4991% de la variabilidad de los datos originales (Tabla 4).

### Ponderación del componente

Para ponderar los componentes se determinaron los pesos de las variables en cada componente como se muestra en la Tabla 5.

Un ejemplo de la ponderación del primer factor, utilizando los pesos de los componentes, se muestra a continuación:

$$- 0,366105 * CAP + 0,12465 * CFU + 0,4316 * FOR - 0,52376 * GEST - 0,027355 * PAG + 0,0575 * PRO + 0,3710 * REH + 0,22597 * RTE - 0,443897 * VYE$$

**Tabla 4.** Análisis de componentes principales

Componente Número	Eigenvalor	Porcentaje de Varianza	Porcentaje Acumulado
1	2,6756	29,729	29,729
2	1,39058	15,451	45,180
3	1,20047	13,339	58,518
4	1,16827	12,981	71,499
5	0,92023	10,225	81,724
6	0,64537	7,171	88,895
7	0,385971	4,289	93,183
8	0,374555	4,162	97,345
9	0,238952	2,655	100,000

Fuente: Elaboración propia.

**Tabla 5.** Pesos de los componentes

	Componente 1	Componente 2	Componente 3	Componente 4
CAP	-0,3661	0,119122	0,565561	0,205187
CFU	0,12465	-0,433559	0,361312	0,340613
FOR	0,43160	0,311818	0,069714	0,167703
GEST	-0,5237	0,17588	0,027226	-0,136731
PAG	-0,0273	0,306088	0,156727	-0,685604
PRO	0,05752	-0,510734	0,462041	-0,304182
REH	0,37101	0,444429	0,241675	0,246566
RTE	0,22597	0,211679	0,470314	-0,241063
VYE	-0,4438	0,26983	0,162316	0,33701

Fuente: Elaboración propia.

## Resultados del Análisis Factorial (FA)

Con el Análisis Factorial se explicó la mayor parte de la variabilidad en las variables, en un número pequeño de factores, al igual que la correlación entre ellas. Para poder realizar el análisis se verificó que los datos de cada variable fueran linealmente independientes y que tuvieran una distribución normal multivariada. Las nueve variables que fueron cargadas en Statgraphics, arrojaron la Tabla 6 en la que se extrajeron cuatro factores que tuvieron eigen-valores mayores o iguales que 1,0. En conjunto ellos explican el 71,4991% de la variabilidad en los datos originales.

## Estadísticas de la selección

En la Tabla 7 se muestra la matriz inicial de cargas con las correlaciones entre variables y factores. Los valores de las correlaciones que en su mayoría se encuentran alejados de 0 y 1,

permiten determinar que es correcto el número de factores del análisis.

Haciendo uso del *software* SPSS se encontró el valor del determinante de la matriz de correlaciones que corresponde a 0,107, con el fin de establecer si podía realizarse el análisis factorial, ya que si es cercano a 1, no debe continuarse con el análisis debido a que las variables no están correlacionadas entre sí.

Para la validación del análisis se verificó que los valores de la varianza específica de las variables no fueran negativos, y que el valor de la comunalidad no sea mayor de 1, en razón a que cuando esto sucede, no tendría sentido realizar este análisis. En la Tabla 8 se evidencia el cumplimiento de lo expresado.

## Estadísticos de rotación

Debido a que el análisis factorial debe hacerse sobre los factores rotados, se aplicó el método de rotación

Tabla 6. Análisis de factores

Factor Número	Eigenvalor	Porcentaje de Varianza	Porcentaje Acumulado
1	2,6756	29,729	29,729
2	1,39058	15,451	45,180
3	1,20047	13,339	58,518
4	1,16827	12,981	71,499
5	0,92023	10,225	81,724
6	0,64537	7,171	88,895
7	0,385971	4,289	93,183
8	0,374555	4,162	97,345
9	0,238952	2,655	100,000

Fuente: Elaboración propia.

Tabla 7. Matriz de cargas antes de rotar

	Factor 1	Factor 2	Factor 3	Factor 4
CAP	-0,598847	0,140472	0,619662	0,221779
CFU	0,203906	-0,51126	0,395875	0,368156
FOR	0,705985	0,367705	0,0763828	0,181264
GEST	-0,856729	0,207403	0,0298306	-0,14778
PAG	-0,044746	0,360948	0,171719	-0,74104
PRO	0,094095	-0,60227	0,50624	-0,32878
REH	0,606881	0,524083	0,264794	0,266504
RTE	0,369636	0,249619	0,515304	-0,26055
VYE	-0,726094	0,318192	0,177843	0,364262

Fuente: Elaboración propia.

**Tabla 8.** Comunalidad y varianza

Variable	Comunalidad	Varianza Específica
CAP	0,811518	0,188482
CFU	0,595227	0,404773
FOR	0,672312	0,327688
GEST	0,799732	0,200268
PAG	0,710921	0,289079
PRO	0,735961	0,264039
REH	0,784108	0,215892
RTE	0,532369	0,467631
VYE	0,792774	0,207226

Fuente: Elaboración propia.

**Tabla 9.** Matriz de cargas del factor rotado

	Fact 1	Fact 2	Fact 3	Fac r 4
CAP	0,87567	-0,0222	0,20992	0,01236
CFU	0,02217	0,12688	0,58338	-0,4881
FOR	-0,2807	0,76330	-0,0969	-0,0386
GEST	0,64184	-0,5038	-0,2326	0,28247
PAG	0,0068	0,02102	0,01721	0,84269
PRO	-0,0907	-0,1524	0,83115	0,11687
REH	-0,0306	0,87970	-0,0960	-0,0078
RTE	0,01707	0,51885	0,3184	0,40183
VYE	0,82825	-0,1491	-0,2703	-0,1067

Fuente: Elaboración propia.

de Varimax, que arrojó los resultados mostrados en las Tablas 9 y 10, en donde se observa que los valores de la varianza específica de las variables no son negativos, y la comunalidad de las variables no son mayores de 1.

**Tabla 10.** Comunalidad y varianza

Variable	Comunalidad	Varianza específica
CAP	0,811518	0,188482
CFU	0,595227	0,404773
FOR	0,672312	0,327688
GEST	0,799732	0,200268
PAG	0,710921	0,289079
PRO	0,735961	0,264039
REH	0,784108	0,215892
RTE	0,532369	0,467631
VYE	0,792774	0,207226

Fuente: Elaboración propia.

## Coefficientes y valores del factor

Un instrumento para calcular las puntuaciones factoriales corresponde a la matriz de coeficientes del factor.

Después de hallar los factores rotados se procede a encontrar las puntuaciones factoriales, tomando los coeficientes de calificación del factor y multiplicándolos con los valores estandarizados, para identificar las Mipymes con mayor puntuación en cada factor, con el fin de establecer un grupo de ellas con las mismas características.

A continuación se indican los coeficientes de los factores para cada variable.

**Tabla 11.** Coeficientes del factor

	<b>Factor 1</b>	<b>Factor 2</b>	<b>Factor 3</b>	<b>Factor 4</b>
CAP	0,531554	0,10618	-0,016988	0,247938
CFU	0,162241	0,075552	0,392862	-0,16453
FOR	0,231346	0,430553	-0,109701	-0,11059
GEST	-0,019689	-0,38298	-0,031646	0,201541
PAG	0,833513	-0,18713	-0,517484	-4,01258
PRO	-0,892154	-0,39568	0,366484	0,703863
REH	-1,28248	0,705868	-3,11983	3,60042
RTE	-1,29076	0,573999	0,376896	1,90515
VYE	-1,47852	-0,29597	0,448513	-4,92289

Fuente: Elaboración propia.

## Interpretación de los factores

Haciendo uso de la Tabla 11, se extrajeron las variables con correlaciones por encima de 0,2 en los componentes:

En el primer factor se encuentran las variables CAP, GEST y VYE con acción directa o positivas (0,876; 0,642 y 0,828, respectivamente), lo que indica que hay un conjunto de Mipymes que se destacan por capacitar al personal y su capacidad de gestión como sus elementos de éxito en ventas y exportaciones.

En el segundo factor se encuentran las variables FOR, REH y RTE con acción directa o positivas (0,763; 0,879 y 0,518, respectivamente), lo que indica que hay un conjunto de Mipymes que se destacan porque su recurso humano tiene la cualificación necesaria y el recurso tecnológico requerido, mientras que por tener acción inversa con GEST, evidencian que a pesar de su fortaleza en buenos recursos, hay malos resultados en la gestión.

En el tercer factor se encuentran las variables CFU y PRO con acción directa o positivas (0,583 y 0,831, respectivamente), lo que indica que hay un conjunto de Mipymes que se destacan porque su producción les permite sufragar sus costos de funcionamiento.

En el cuarto factor se encuentran las variables PAG y RTE con acción directa o positivas

(0,843 y 0,402, respectivamente), lo que indica que hay respuestas altamente coincidentes en la variable participación en agremiaciones además de contar con buen recurso tecnológico, mientras que por tener acción inversa con CFU, evidencian que tienen dificultades con los costos de funcionamiento.

## Validación del análisis

En el análisis factorial, es importante saber si las correlaciones de variables pueden ser explicadas por otra variable, para saber si esto es posible se debe medir la adecuación muestral para lo cual se puede usar el *KMO* (Kaiser-Meyer-Olkin) que corresponde a un número entre 0 y 1. Los valores menores de 0,5 indican que no debe utilizarse el análisis factorial con los datos muestrales que se están estudiando. Para el caso de estudio se encontró un *KMO* de 0,604.

Otra prueba que se realizó es el test de esfericidad de Barlett, el cual confirmó si la matriz de correlaciones es la identidad, lo que significa que las intercorrelaciones entre las variables son cero, el chi-cuadrado aproximado obtenido en esta prueba fue 85,281. Cuando se obtienen valores altos de chi-cuadrado, o en su defecto un determinante bajo, significa que hay variables con correlaciones altas, por tanto cuando el estadístico toma un valor alto se rechaza la hipótesis nula de que la matriz de correlaciones es la idéntica.

## Resultados del análisis de conglomerados (Clúster)

El análisis para las 9 variables, se realizó aplicando el método del vecino más cercano y la distancia euclidiana cuadrada, esto con el fin de establecer la cercanía de las variables CAP, CFU, FOR, GEST, PAG, PRO, REH, RTE y VYE y la forma de agruparlas.

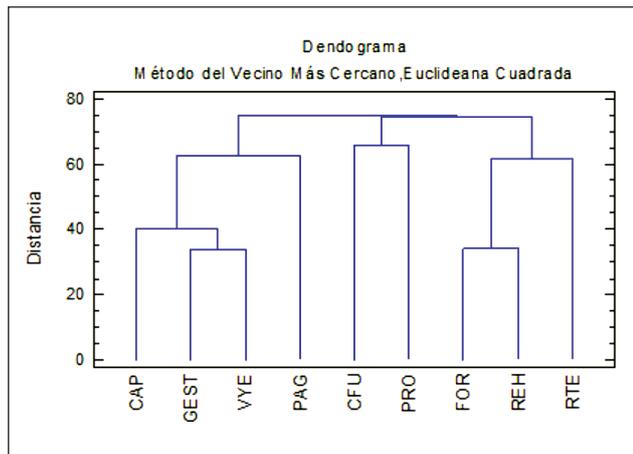


Figura 1. Dendrograma con todas las variables.

En el dendrograma de la Figura 1 se observa que las primeras variables en formar un conglomerado fueron gestión (GEST) y ventas y exportaciones (VYE), seguidas casi simultáneamente con formación (FOR) y recursos humanos (REH), así mismo se aprecia que son las variables que tienen distancias más cercanas.

Al grupo conformado por GEST y VYE se unió capacitación (CAP) y luego participación en agremiaciones (PAG), mientras que a FOR y REH se le agregó recursos tecnológicos (RTE) quienes se unieron al grupo conformado por costos de funcionamiento (CFU) y producción (PRO) que son variables con mayor distancia en el dendrograma. Lo anterior sugiere conformar dos grupos tal como se muestra en la Tabla 12.

En el primer clúster se observa la interacción de las variables que sugieren que estar agremiados,

Tabla 12. Clúster para variables

Clúster de Variables	
<i>Primer Clúster</i>	Capacitación Gestión Ventas y exportaciones Participación en agremiaciones
<i>Segundo Clúster</i>	Costos de funcionamiento Producción Formación Recurso humano Recurso tecnológico

Fuente: Elaboración propia.

capacitar al personal y realizar una buena gestión, son indicadores de éxito en las ventas y las exportaciones. La Tabla 13 evidencia que este clúster tiene una participación del 44,44% del total de variables.

El segundo clúster expresa que dotando de recurso tecnológico al personal con buena formación académica, mejora la productividad de la empresa, garantizando a su vez que se puede asumir sin problema los costos de funcionamiento. La Tabla 16 evidencia que este clúster tiene una participación del 55,56% del total de variables.

Tabla 13. Resumen de conglomeración

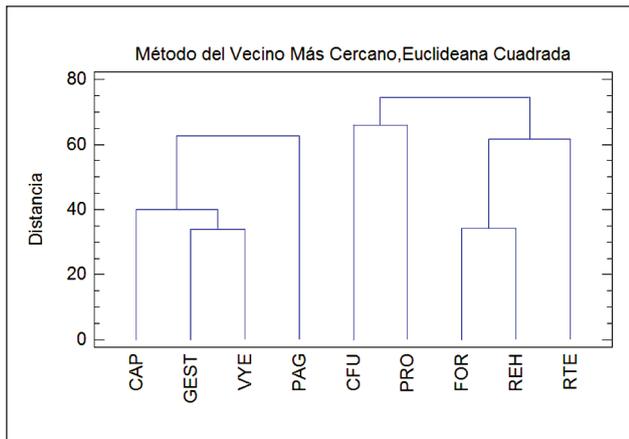
Conglomerado	Miembros	Porcentaje
1	4	44,44
2	5	55,56

Fuente: Elaboración propia.

El dendrograma obtenido con los dos clúster se muestra en la Figura 2.

## Análisis de los resultados de los métodos multivariados

El análisis de los componentes principales permitió establecer que las Mipymes del sector alimenticio de la ciudad de Montería, aciertan en la vinculación del personal con la formación



**Figura 2.** Dendrograma

necesaria en los cargos, ya que su personal tiene la cualificación requerida en los cargos que desempeñan, que también desconocen lo relacionado con los costos de funcionamiento y su articulación con los procesos productivos. El método también permitió establecer la relación de las variables costos de funcionamiento con las ventas, evidenciando que las ventas realizadas permiten costear su funcionamiento.

Con el análisis factorial se pudo determinar que las capacitaciones realizadas han permitido en su momento mejorar los resultados en ventas y la gestión, como también hacer evidente la necesidad de articular con el recurso humano la formación y las ayudas tecnológicas como también sugerir su participación en agremiaciones con el fin de compartir los recursos tecnológicos.

Con el análisis de conglomerados se pudo construir dos grupos de variables que mostraron en primer lugar la relación existente entre capacitación, gestión, ventas y agremiaciones y en segundo lugar la interacción que tienen las empresas entre recurso humano con formación dotado de recurso tecnológico para lograr resultados en la producción para poder costear su funcionamiento. Este método también permitió establecer que no hay homogeneidad entre las Mipymes de tal manera que no es posible agruparlas por su poca similitud como para conformar un clúster.

## Conclusiones

La realización de la descripción de las características generales de los métodos multivariados, y en particular los que se encuentran disponibles en Statgraphics, permitió establecer que el tipo de técnica que se utilizó, es un factor de gran importancia al momento de seleccionar el método. En tal sentido este ejercicio investigativo hizo uso de la metodología de análisis de datos multivariados para caracterizar las Mipymes de la ciudad de Montería. Es pertinente resaltar que la revisión bibliográfica sobre la temática objeto de la investigación, muestra que en la ciudad de Montería se carece de herramientas técnicas, metodológicas y conceptuales para abordar caracterizaciones del sector comercio. Se trata de un campo por explorar, en el cual esta investigación tiene mucho que aportar como experiencia piloto.

Al caracterizar las Mipymes del sector alimenticio de la ciudad de Montería, reportados en la Cámara de Comercio, se pudo establecer que para el periodo 2009, predominaban las microempresas, con una participación del 83,72% del total de las Mipymes. De igual manera se constató que en el mismo periodo el 67,44% de estas, se clasifican en productoras de tipo familiar, lo que es muy notable en la sociedad cordobesa que se caracteriza por la integración familiar.

El diseño de un componente informático para el procesamiento de la encuesta permitió obtener una matriz de datos codificados compatible con los paquetes estadísticos Statgraphics y SPSS, constituyéndose en una herramienta fundamental para el análisis de la información reportada por las Mipymes, objeto de estudio.

Como los valores del determinante (0,107) de la matriz de correlaciones y el *KMO* (0,604) están dentro de parámetros permisibles, se puede garantizar con seguridad la aplicación del Análisis de Componentes Principales y el Análisis Factorial que a su vez permiten explicar en un 71,5% la correlación de los datos en 4 componentes o factores, de donde se puede deducir que en las Mipymes del sector alimenticio de Montería el recurso humano posee la misma cualificación tal como lo expresa el

primer componente conformado por las variables Formación y Recursos Humanos; mientras que en el segundo factor están correlacionadas las variables Recursos Humanos, Formación y Participación en Agremiaciones, indicando que un buen número de Mipymes respondieron de forma parecida en estas variables, indicando situaciones similares entre ellas en este aspecto; el tercer componente, Capacitación, Recurso Tecnológico, Producción y Costos de Funcionamiento; mientras que en el cuarto componente, predominaron las variables Costos de Funcionamiento y Ventas y Exportaciones, en donde se ratifica lo antes expresado.

La aplicación del Análisis de Conglomerados permitió reducir la explicación obtenida en el Análisis de Componentes principales y el Análisis Factorial de 4 factores a 2 grupos de variables, uno conformado por gestión, ventas y exportaciones, capacitación y participación en agremiaciones con diferencias despreciables en las respuestas presentadas por las Mipymes y el otro grupo conformado por formación, recursos humanos, recursos tecnológicos, producción y costos de funcionamiento, lo que expresa claramente la similitud en las respuestas de las Mipymes encuestadas. En donde es evidente que estas Mipymes se caracterizan por no ofrecer una buena atención a los clientes por la formación de sus empleados, los nichos de mercado que atienden, la infraestructura física y tecnológica montada y la vulnerabilidad de ellas.

El Análisis de Conglomerados para las unidades experimentales (Mipymes), expresa que no hay posibilidades de agrupar las Mipymes en grupos con características similares, esto se debe a que se encontraron clúster triviales es decir, constituidos por una sola empresa, lo que indica una falta de factores comunes entre ellas, demostrando que no hay características similares entre las 43 empresas encuestadas que permita agruparlas en un gremio con gran afinidad.

## Referencias

1. Osorio, B. (2009). *Caracterización de las micro, pequeñas y medianas empresas en el Departamento de*

*Córdoba*. Montería: Ediciones Universidad del Sinú, pp. 8-60.

2. Sánchez C., J.J. (2007). “Algunas aproximaciones al problema de financiamiento de las Pymes en Colombia”, en: *Scientia et Technica*, Año XIII, No. 34, mayo, pp. 321-322.

3. Lancheros, Valery; Hernández, Helman y Robles, Juana (2008). “Sistemas de Gestión en el sector de elaboración de productos alimenticios y de bebidas”, en: *Revista Ingeniería*, Universidad Distrital Francisco José de Caldas, Vol. 13, No. 2.

4. Figueras, S. (2000). “Introducción al Análisis Multivariante. de Estadística”, en: <http://www.5campus.com/leccion/anamul>. Consultado el 12 de octubre de 2009.

5. Gurrea, M. T. (s.f.). “UOC, Análisis de componentes principales, Proyecto e-Math”, en: [http://www.uoc.edu/in3/emath/docs/Componentes\\_principales.pdf](http://www.uoc.edu/in3/emath/docs/Componentes_principales.pdf). Secretaria de Estado de Educación y Universidades MECD. Consultado en octubre de 2009.

6. Olmos, S., & Di Renzo, M. (2004). “INTA”, en: [http://www.inta.gov.ar/ediciones/2004/biotec/parte6\\_cap1.pdf](http://www.inta.gov.ar/ediciones/2004/biotec/parte6_cap1.pdf). Consultado en octubre de 2009.

7. Peña, Daniel (2002). *Análisis de datos multivariados*. Madrid - España: Editorial MacGraw Hill, pp. 133-158.

8. Romero, R.V. (2003). *Javeriana*. Recuperado en enero de 2010, de Capítulo 10: análisis de conglomerados (clúster): <http://tic.javeriana.edu.co/apps/Manuales/R/CursodeestadisticaconR.pdf>

9. Pérez, Cesar (2004). *Técnicas de análisis multivariante de datos. Aplicaciones con SPSS*. Madrid - España: Editorial Pearson educación, pp. 121-154.

10. Dallas, Johnson (2000). *Métodos multivariados aplicados al análisis de datos*. México: Thomson Editores S.A., pp. 93-396.