
Análisis estadístico para la administración del inventario en empresas comerciales aplicando métodos multivariados

Statistical analysis for inventory management in commercial companies applying multivariate methods

Jonathan Cano Bedoya^a
jonathancano@itm.edu.co

Luz Sánchez Herrera^b
luz.sanchez@udea.edu.co

Estefanía Puerta Uribe^c
estefania.puerta@udea.edu.co

Resumen

Las empresas comerciales se dedican a la compra y venta de productos terminados, así que, parte del éxito financiero está en el control y la administración del inventario. El análisis de las ventas requiere que se estudie la rotación, identificando aquellos artículos que generan más movimiento y ganancias, esto con la finalidad de implementar planes estratégicos de mejoramiento en la inversión para la consecución de la mercancía. Por otra parte, las compañías que cuentan con varios sitios de venta requieren de un control sobre sus tiendas en términos de gestionar los inventarios. La operación crece y se vuelve compleja ya que comprende diferentes clases de productos y diferentes sedes de venta, se genera mucha información que necesita ser analizada. La estadística provee diferentes técnicas para el análisis de este tipo de datos, las cuales se utilizaron para agrupar y clasificar tanto artículos como puntos de venta, analizando la evolución de las series de las cantidades vendidas mediante los modelos SARIMA de series temporales.

Palabras clave: Empresas comerciales, Inventarios, Análisis de conglomerados, Series de tiempo, Tablas de tres vías.

Abstract

Commercial companies are dedicated to the purchase and sale of finished products, so financial success depends on the control and management of the inventory. Analysis sales requires that turnover rate be studied, identifying those items that generate more movement and profits, its purpose of implementing strategic plans of improvement the investment for the procurement of the merchandise. On the other hand, companies that have many points of sale require control over their stores in terms of managing inventories. The operation grows and becomes complex for the variety of products and many points of sale, a lot of information is generated that needs to be analyzed.

^a Ingeniero Financiero y docente del Instituto Tecnológico Metropolitano. Estadístico de la Universidad de Antioquia

^b Estudiante de estadística Universidad de Antioquia

^c Docente Estadística Instituto de Matemáticas de la Universidad de Antioquia

Statistics provides different methods for data analysis, which were used to group and classify both products and points of sale, analyzing the evolution of the time series of quantities sold using SARIMA models.

Keywords: Commercial companies, Inventories, cluster analysis, time series, three-way tables

1. Introducción

La administración y gestión del inventario es una actividad esencial que realizan las empresas debido a la dinámica que este genera, en uno o varios almacenes donde se manufacturan y o comercializan las mercancías. El sistema de abastecimiento debe ser riguroso en el control de entradas y salidas de las unidades. En el caso de las empresas que solo se dedican a comercializar (compra y venta de mercancías no fabricadas por ellas), y que concentran la operación en un solo almacén, el control es una tarea sencilla en cierto sentido, al momento de ampliarse a varios puntos de venta la labor crece, la distancia geográfica, por ejemplo, conlleva a que los niveles de ventas pueden ser diferentes en cada una, y esto incrementa el gasto en fletes, seguros y otras erogaciones. Es necesario retroalimentar los planes de compra por el volumen de las operaciones que se deben controlar. Se utiliza la técnica multivariada del análisis de conglomerados para la clasificación de los productos y con ello se analiza algunos de estos con series de tiempo usando modelos que estudien la dinámica de las ventas. Finalmente, se analizan las sedes o tiendas mediante tablas de tres vías con la finalidad de ayudar a que las compañías con un gran volumen de artículos y puntos de ventas, puedan analizar la información que generan los inventarios de una manera más eficiente.

2. Planteamiento del problema

En el caso a desarrollar se toma como referente una empresa que ofrece sus servicios en municipios del departamento de Antioquia, con catorce sedes, se omite la clase de productos que maneja solo con finalidad de que sea un referente para todo tipo de empresa que maneja inventarios en grandes volúmenes y diferentes puntos de operación. Los productos estudiados son de alta rotación y poco perecederos. El problema radica en que la empresa cuenta con un sistema que registra las entradas y salidas de inventario, pero no ha logrado realizar un estudio de los 2187 productos que maneja en cada almacén y comparar el comportamiento de las ventas entre las sedes. Es necesario encontrar una herramienta que permita hacer un análisis completo de los ingresos, así como unidades vendidas por artículos y tiendas. Esta situación ayuda a las empresas a solucionar inconvenientes en la asignación de recursos económicos para la compra de los artículos que se disponen a vender en cada almacén, al igual que la distribución del capital operacional a los almacenes, es decir, saber a cuáles productos y sedes invertir más o de otro modo, generar estrategias de crecimiento en ventas y posicionamiento en el mercado, al mejorar falencias en los sitios de operación.

3. El análisis de los inventarios

Uno de los enfoques administrativos en las organizaciones es la planeación, la cual al aplicarse al análisis de inventarios debe darse desde un plan de mejoramiento continuo. Algunos beneficios de planear son:

- Aclarar, amplificar y determinar objetivos.
- Pronosticar.
- Establecer condiciones y suposiciones.
- Establecer políticas, procedimientos y métodos de desempeño.
- Anticipar posibles problemas futuros.
- Modificar los planes a la luz de los resultados de control.
- Establecer un plan general de logros, enfatizando la creatividad para encontrar medios nuevos y mejores de desempeñar en el trabajo.

(Bolaños, Cortés y Delgado, 2011)

Otra de las áreas que se une al proceso de administración y gestión de inventario en las empresas, es el control interno, el cual mantiene unas premisas para la aplicación de un control responsable:

- Mantener el capital mínimo invertido.
- Reducir costos financieros ocasionados por mantener cantidades excesivas de inventario.
- Reducir el riesgo de fraudes, robos o daños físicos.
- Evitar que se dejen de realizar ventas por falta de mercancías.
- Evitar o reducir las pérdidas resultantes de baja de precios.
- Reducir el costo de la toma del inventario físico anual.

(Velásquez, 2012)

En el campo de las empresas comerciales, la gestión del inventario se enmarca en las finanzas de la compañía, pues los beneficios económicos giran en torno al buen trabajo que se realice con los productos (Garrido y Cejas, 2017). La creación de valor, el posicionamiento en el mercado, la fidelización del cliente son aquellos puntos en los que se deben promover estrategias de mejora, se requiere conocer bien la situación de los productos y de las sedes, debido a que el incumplimiento en el servicio provocaría la pérdida de un cliente o también de los proveedores. Este trabajo ilustra la aplicación de metodologías estadísticas para analizar la situación de estudio en el proceso de la gestión de inventario para unidades de negocio dedicadas a la actividad de comercialización.

4. Descripción de las ventas

Se tienen 2187 artículos registrados en el inventario, algunos son similares en términos del tipo de producto, el uso final y otras características. Algunas variables como el precio y las cantidades vendidas permiten también clasificarlos, describir cada uno es una tarea compleja, ya que el interés está en que éstos tengan mayores de ventas tanto en ingresos

como en unidades porque son los que permiten que la entidad sea rentable. En la tabla 1 se muestra que en las existencias del 2018 hay productos con un costo unitario de \$252 (pesos colombianos) y otros de \$325.794, precios entre \$420 y \$542.990, de algunos se venden en todo el año solo 3 unidades y otros hasta 14.177 unidades, algunos con costos anuales de \$51.156 y otros de \$65.899.680, y en ingresos totales de \$85.260 hasta \$69.069.240. Esto muestra la existencia de productos con costos bajos, medios y altos, lo mismo para los ingresos, y es necesario captar la atención en estos rangos.

Tabla 1: Rango de precios, costos y cantidades de los productos año 2018

Estadístico	Costo Unitario	Precio Unitario	Cantidades vendidas	Costo total	Ingresos
Mínimo	\$252	\$420	3	\$51.156	\$85.260
Máximo	\$325.794	\$542.990	14.177	\$65.899.680	\$69.069.240

La figura 1 muestra el comportamiento de las unidades vendidas mensuales, la información es la suma de todas las sedes, se nota que los meses de enero, mayo, junio y julio y final de año noviembre y diciembre presentan crecimientos en las ventas. Esto evidencia tiempos en que los inventarios rotan con mayor frecuencia.

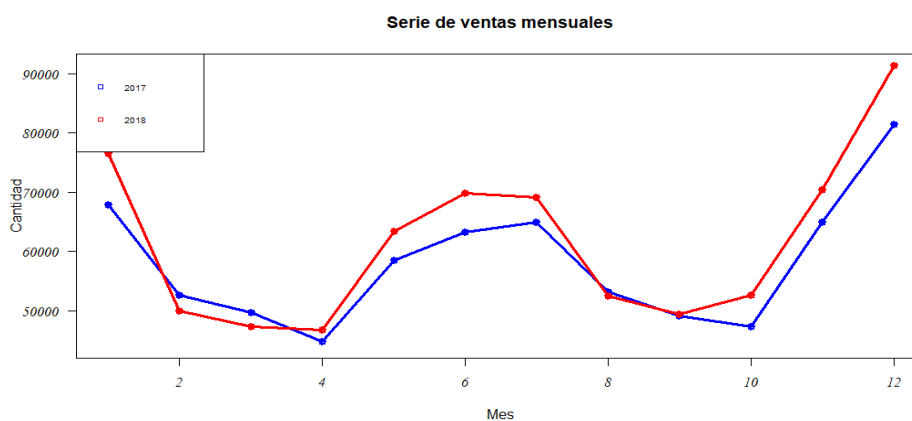


Figura 1: Ventas mensuales en cantidades

En la figura 2 se muestran diagramas de cajas y bigotes de las cantidades vendidas por mes en el año 2018 (el comportamiento es similar para el 2017), las estrellas representan productos con altos volúmenes de venta. Desde el punto de vista estadístico son datos atípicos, sin embargo, por la dispersión que presenta la variable cantidad, en este estudio tienen todo el sentido. Estos productos deben ser identificados para analizar la utilidad

que le dejan a la empresa. Se observa también el mismo patrón de comportamiento de la figura 1, grades volúmenes de venta en enero y diciembre.

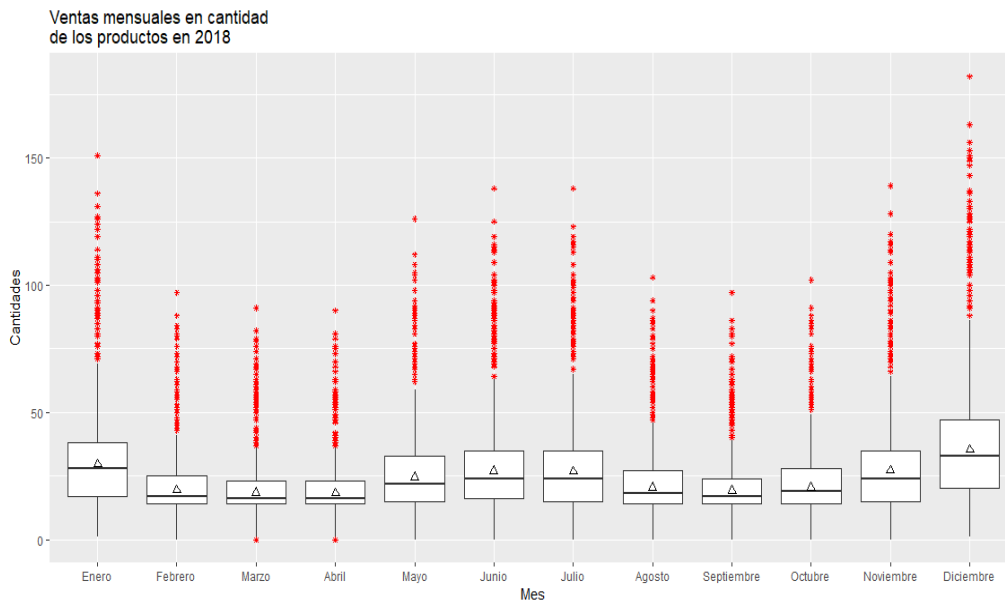


Figura 2: Diagrama de cajas y bigotes de las ventas en cantidades mensuales año 2018.

La figura 3 muestra la dinámica de las ventas (en unidades) de cada mes, se observa que la mayoría de los productos tienen frecuencias de venta entre 0 y 50, también se nota una cola larga a la derecha, donde son pocos los artículos que tienen volúmenes de venta superior a las demás.

La tabla 2 ilustra el comportamiento de las catorce sedes ubicadas en fila, y en cada columna se encuentran los meses del año 2018. Por ejemplo, la sede 1 en el mes 1 (enero) representa el 10.8% de las ventas respecto al total de las sedes, y durante los 12 meses es el porcentaje promedio de 10.6% (promedio de toda la fila), cifra que en el 2017 fue del 10.7%, es decir, que no hubo cambio sustancial.

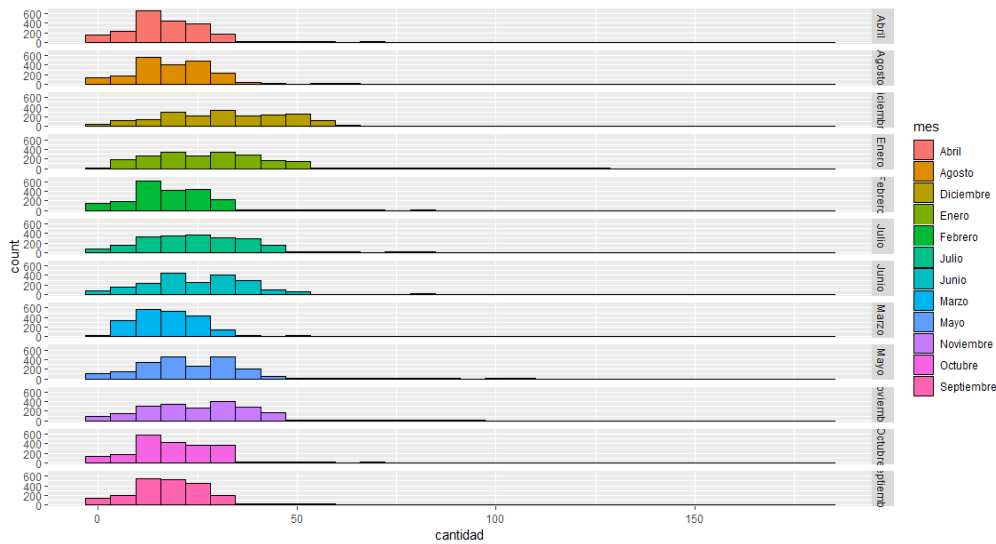


Figura 3: Histograma de las ventas en cantidades mensuales año 2018

Tabla 2: Participación en las ventas 2018 por sede en %

Sede	M-1	M-2	M-3	M-4	M-5	M-6	M-7	M-8	M-9	M-10	M-11	M-12
1	10.8	10.9	11.3	10.3	10.8	10.2	11.0	11.0	10.2	10.6	10.8	10.2
2	9.8	9.4	10.3	9.4	10.0	10.0	9.5	7.6	10.1	9.9	9.4	9.6
3	9.0	8.9	8.6	8.8	8.4	8.8	9.1	9.4	8.8	9.0	9.2	9.4
4	8.6	8.4	8.6	8.7	8.7	8.9	8.5	8.8	8.4	8.8	8.4	9.0
5	8.2	8.2	8.3	7.7	8.2	8.3	8.1	8.5	8.2	8.5	8.5	8.5
6	6.8	7.1	7.1	7.2	6.9	6.8	7.1	6.7	7.0	6.3	6.7	6.6
7	6.5	6.6	6.8	6.8	6.6	6.6	6.6	6.3	6.9	6.1	6.6	6.6
8	6.3	6.7	6.4	6.5	6.2	6.4	6.3	6.8	6.6	6.0	6.5	6.0
9	6.1	5.6	6.2	6.2	6.0	6.0	6.5	6.0	6.1	6.0	5.8	6.1
10	5.7	6.0	5.7	5.4	5.7	5.6	5.9	6.6	6.1	5.8	5.2	6.0
11	5.6	5.5	5.1	5.9	5.7	5.8	5.9	6.2	5.9	5.8	5.5	5.3
12	5.4	5.6	4.9	5.6	5.5	5.7	5.2	5.2	5.7	5.6	5.9	5.5
13	5.6	5.2	5.5	5.8	5.7	5.3	5.2	6.0	5.2	5.8	5.9	5.7
14	5.4	5.9	5.1	5.6	5.5	5.5	5.1	4.9	4.9	5.6	5.6	5.5

La figura 4 muestra el porcentaje de ventas de cada sede en los años 2017 y 2018, representados en el tamaño de los círculos. La sede 1 se mantuvo igual en ambos años, siendo la de mayor desempeño. La 14 tuvo un buen desempeño en el 2017 pero al 2018 sus ventas bajaron. Las tiendas 2, 3 y 5 son muy parecidas en cada periodo, aumentando las ventas de un año a otro, de igual manera para 7, 8, 11 y 14, son parecidas en los

años, pero sus ventas disminuyeron. Lo anterior permite detallar de manera descriptiva las ventas por tienda y compararlas entre ellas, dado el caso en una asignación de recursos para recompra de inventario, se le debería asignar más a la primera, porque compromete a la compañía en términos de rotación de inventarios, y las de menores porcentajes necesitan planes estratégicos u otros como una fusión o cambio de posición geográfica para mejorar las ventas.

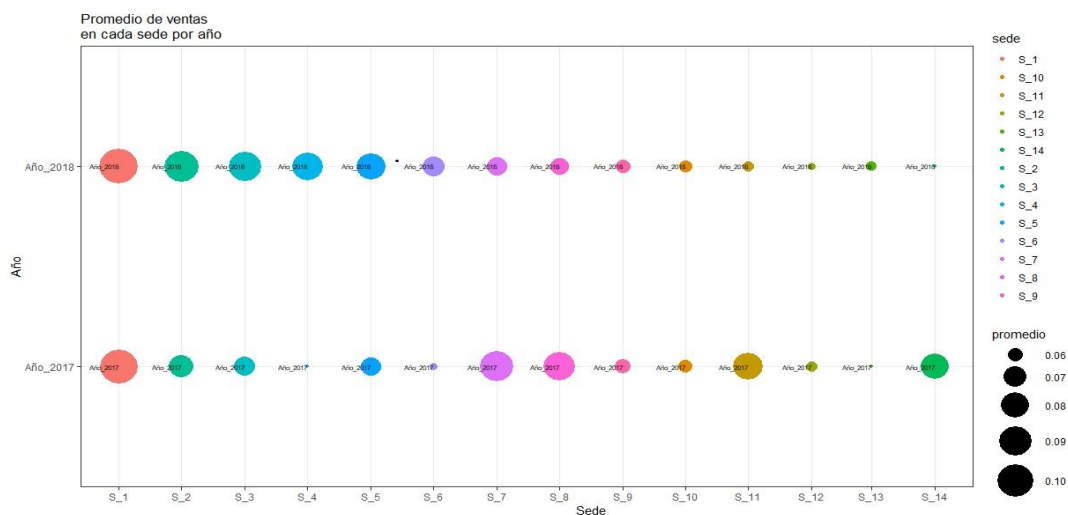


Figura 4: Promedio porcentual de las ventas en cada sede por año. Elaboración propia

5. Clasificación de sedes de venta y productos en conglomerados

En el análisis de la información se tiene un volumen considerable de datos, se dispone de 2187 productos en catorce sedes de venta, con 25 variables, la primera es el precio y las otras 24 son las ventas mensuales durante el periodo 2017-2018. Esto correspondería a analizar 76450 ($2187 \times 14 \times 25$) datos. Uno de los propósitos más frecuentes para la construcción y análisis de conglomerados es la identificación de una estructura natural en los objetos, es decir, el desarrollo de una tipología o clasificación de los objetos (Díaz, González, Henao & Díaz, 2013). Estas estructuras definen unos grupos con elementos similares en término de las variables medidas sobre ellos, existen varios procedimientos de agrupación, si existen grupos muy definidos, todos estos dan resultados similares. Existen varios criterios para elegir el número de grupos, aunque dentro de los grupos, el criterio decisivo en la homogeneidad podría elegir de antemano por el investigador.

Cada grupo o “cluster” (clase a la que el grupo pertenece) sea lo más homogéneo o compacto con respecto a ciertas características, es decir, las observaciones dentro de cada grupo han de ser similares entre sí. Cada grupo debe diferenciarse de los otros demás respecto a las mismas características, es decir, las observaciones de un grupo deben diferenciarse de las observaciones de los otros.

En la formación de los grupos se debe elegir una regla de agrupación: Vecino más cercano (single linkage), Vecino más lejanos (Complete linkage), Promedio no ponderado por par de grupos, Promedio ponderado por par de grupos, Centroide no ponderado por par de grupos, Centroide ponderado por par de grupos, Método de Ward. De igual manera se debe optar una distancia: distancia euclídea, euclídea ponderada, de Mahalanobis, Chebyshev, etc.

En el desarrollo de este trabajo se utilizó el método de conglomeración jerárquico, número crecientes de clases anidadas, con la regla de agrupación del vecino más lejano y la distancia euclídea (de gran uso en estadística), ecuación 1. Se realizaron otros métodos como el método de Ward, vecino más cercano, promedios y los resultados fueron muy similares. El resultado de la conglomeración se observa en la figura 5, el dendrograma muestra la secuencia de formación, eje horizontal, indicando la distancia de unión, este sugiere tomar 4 grupos, pues el salto de agrupación es alto.

$$d_{ij} = \sqrt{\sum_{k=1}^p |x_{ik} - x_{jk}|^{1/2}} \quad (1)$$

(Díaz, Morales, 2012)

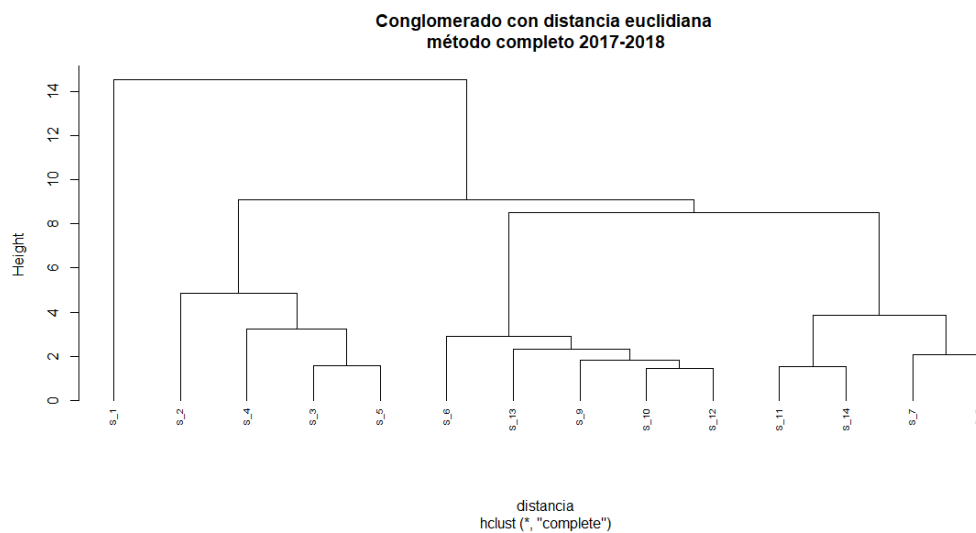


Figura 5: Dendrograma para las ventas en el periodo 2017-2018 por mes en cada sede, método completo

Tabla 3: Conglomerado de las sedes de venta en cuatro grupos

Grupo	Sedes
1	1
2	2, 3, 4, y 5
3	7, 8, 11 y 14
4	6, 9, 10, 12 y 13

Los grupos ubicados en la tabla 3 muestran que el primero solo cuenta con la sede 1, la cual es la sede que más vende, el segundo son las que venden más en segundo lugar y así con los otros dos grupos, de hecho, las sedes del último grupo son las que más se parecen y lo hacen en que son las que tienen menores porcentajes de participación en las ventas totales.

También se realizó una agrupación de los productos, no usando los métodos aglomerativos mencionados, sino el método de *k-medias* el cual sigue los siguientes pasos para la agrupación de los individuos (en este caso productos) (Véliz, 2017, p. 55):

- Definir una muestra de N elementos en p variables numéricas
- Se forman al azar K grupos y se calculan los centroides de cada grupo los cuales son las medias aritméticas de las variables.
- Con la distancia euclidiana se calcula la distancia de cada elemento respecto a los centroides y asignándolos al centroide más cercano, se recalculan los centroides de los nuevos grupos.
- Si la distancia entre los centroides iniciales y los nuevos centroides es pequeña o si se ha completado un número fijo de iteraciones, el proceso termina. De otro modo se repite el paso 2.

Al implementar el algoritmo en el software R Project, se decidió tomar 8 grupos, de ellos, tal como lo sugiere la figura 6, se observa la relación entre la distancia y el número de grupos, se hizo el corte en 4 conglomerados porque al pasar a 5 grupos la distancia ya era mayor.

6. Aplicación de las series de tiempo en las ventas

En el análisis de las series de venta de los productos, se analiza dos componentes, el primero para probar estacionaria en media, y si existe una componente estacional, es decir, si hay un patrón en las ventas debido a estaciones, lo cual es útil para comparar las ventas de todos los eneros, los febreros y así para cada mes del año. En el caso analizado las ventas de enero de 2017 y 2018 son altas, las de febrero bajan en ambos años (figura 1). El modelo ARIMA (Autoregressive Integrated Moving Average) aplica para series de tiempos estacionarias que no tienen estacionalidad, modelo descrito en la ecuación (2), pero en la metodología de Box-Jenkins se contempla este mismo modelo incluyendo la componente estacional, esto se conoce como el modelo SARIMA (Seasonal Autoregressive Integrated Moving Average). Si se estudia una serie de tiempo estacional como si fuese una serie solamente estacionaria entonces la serie $\{b_t\}$ no será un proceso ruido blanco, ya que contiene inexplicables correlaciones entre períodos (estacionales).

$$\phi_p(B)(1-d)^d Z_t = \theta_q(B)b_t \quad (2)$$

Sea la expresión

$$\rho_j(s) = \frac{E(b_{t-js} - \mu_b)(b_t - \mu_b)}{\sigma_b^2}, j = 1, 2, 3, \dots$$

la función de autocorrelación para $\{b_t\}$ representando la inexplicable relación entre períodos. Se puede ver que esta relación puede ser representada con un modelo ARIMA reescrito de la siguiente manera

$$\phi_p(B^s)(1-B^s)^D b_t = \theta_q(B^s)a_t \quad (3)$$

Donde,

$$\phi_p(B^s) = 1 - \phi_1 B^s - \phi_2 B^{2s} - \dots - \phi_p B^{ps}$$

y

$$\theta_q(B^s) = 1 - \theta_1 B^s - \theta_2 B^{2s} - \dots - \theta_q B^{qs}$$

son polinomios en B^s sin raíces comunes. Las raíces de los polinomios están por fuera del círculo unitario y $\{a_t\}$ es un proceso ruido blanco con media cero y varianza σ_a^2 .

Combinando (2) y (3) obtenemos el conocido modelo SARIMA de Box-Jenkins (Wei, 2006)

$$\phi_p(B^s)\phi_p(B)(1-B^s)^d \dot{Z}_t = \theta_q(B)\theta_q(B^s)a_t \quad (4)$$

Donde,

$$\dot{Z}_t = \begin{cases} Z_t - \mu, & \text{si } d = D = 0, \\ Z_t, & \text{otro caso,} \end{cases}$$

Entonces los términos $\phi_p(B)$ y $\theta_q(B)$ son los polinomios estacionarios autorregresivos y de medias móviles los cuales describen la parte regular, y $\Phi_P(B^s)$ y $\Theta_Q(B^s)$ los polinomios estacionales que describen la estacionalidad. El modelo dado en la ecuación (4) se denota como $ARIMA(p, d, q) \times (P, D, Q)_s$, donde el subíndice $-s$ se refiere al periodo estacional.

Los datos que representan las unidades vendidas, por lo general, se ajustan a un modelo con componente estacional, se analiza el periodo 2017-2018, escogiendo los cuatro productos con más ganancias de los ocho grupos formados en la parte de los conglomerados. En el análisis de series temporales se realiza primero una exploración de la información, en este caso se observa el comportamiento natural de las series de tiempo que se ilustra en la figura 7, con la evolución de la serie en niveles.

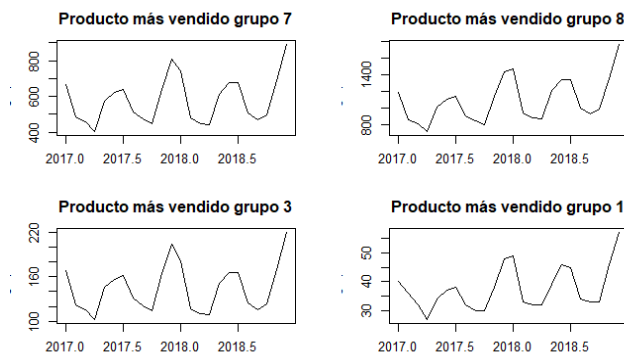


Figura 7: Serie de datos observados

Se desea observar si la serie es estacionaria en media, para esto se utiliza la prueba de Dickey-Fuller aumentada con un nivel de significancia de 0.05. El *valor-p* de la prueba para el producto 7 es 0.1664, así que no se rechaza la hipótesis nula, es decir, existen raíces unitarias que impiden que la serie sea estacionaria, lo que indica que la serie de tiempo se debe diferenciar, aunque esto también se debe a la estacionalidad, ya que la prueba de Dickey-Fuller tiene mejores resultados para las series que no presentan estacionalidad.

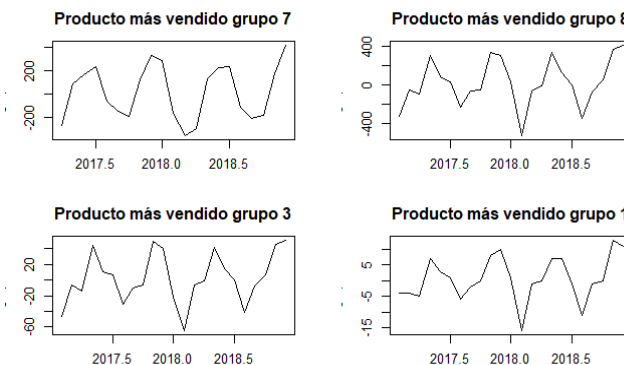


Figura 8: Serie tiempo diferenciada

La series W_t diferenciada (Figura 8.) se observa que hay un patrón que se repite cada periodo de tiempo a lo que se concluye que la serie es estacional es decir, que cada intervalo de tiempo la cantidad de productos vendidos por la empresa suben o bajan de manera similar, con un comportamiento por estaciones.

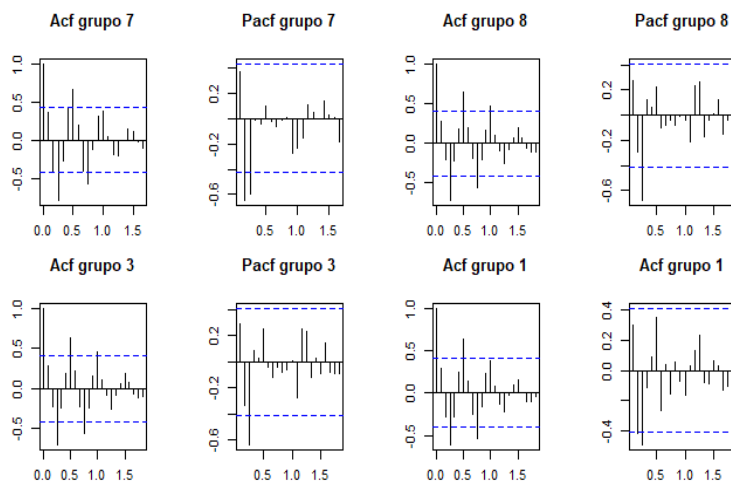


Figura 9: ACF y PACF Serie diferenciada

Antes de usar diferenciación estacional se identifica el periodo estacional de la serie de tiempo (se refiere a qué tiempo se repite el patrón estacional en un año, si es mensual $s=12$). En la figura 9. se muestra en la Función de Autocorrelación (ACF) con la autocorrelación de orden 3, 6, 9, 12 que son estadísticamente diferentes de cero, pues se salen de las bandas de confianza. lo cual indica, que el periodo estacional para esta serie es $s = 3$, es decir, cuatrimestral. La Función de Autocorrelación Parcial (PACF) muestra el proceso como un Ruido Blanco, el cual no se confirma todavía, debido a que la autorrelación parcial de orden 3 es estadísticamente diferente de cero.

Al hacer la diferenciación estacional con periodo $s = 3$, se identificó un modelo autoregresivo de orden 3 ($AR(3)$) para la parte regular, y una media móvil de orden 3 ($MA(3)$) para la componente estacional. En consecuencia, este análisis inicial sugiere que un posible modelo para estos datos es un $SARIMA(3,1,0)(0,1,3)_3$. Se adapta este modelo, junto con algunas variaciones, y se calcula sus valores del Criterio de Información de Akaike (AIC) como se muestra en la tabla 5.

Estos modelos son una opción para todos los productos estudiados, incluso en la tabla 5 los modelos para el elemento seleccionado del grupo 1, también aplica para los productos que pertenecen a ese grupo debido a la clasificación que se hizo anteriormente.

Una de las posibilidades que brindan las series de tiempo es hacer pronóstico de las ventas de los productos, para que se pueda tener un plan de ventas en cada sede. En la figura 10 se muestran los pronósticos de los productos seleccionados en cada grupo,

éste tiene el mismo comportamiento estacional con una tendencia creciente. Dependiendo del tipo de producto que se venda, sea perecedero o no, un bien de consumo o material, un bien necesario o de lujo no siempre la tendencia tiene que ser creciente.

Tabla 5: Modelos de ajuste

Grupo 7			Grupo 8		
N°	Modelo	AIC	N°	Modelo	AIC
1	$SARIMA(3, 1, 0)(0, 1, 3)_3$	239.95	1	$SARIMA(3, 1, 0)(0, 1, 3)_3$	266.09
2	$SARIMA(3, 1, 0)(0, 1, 4)_3$	237.5	2	$SARIMA(3, 1, 0)(0, 1, 4)_3$	263.84
3	$SARIMA(3, 1, 0)(0, 1, 3)_3$	240.95	3	$SARIMA(3, 1, 0)(0, 1, 3)_3$	266.36
4	$SARIMA(4, 1, 0)(0, 1, 4)_3$	239.5	4	$SARIMA(4, 1, 0)(0, 1, 4)_3$	239.5
5	$SARIMA(3, 1, 0)(1, 1, 1)_3$	241.66	5	$SARIMA(3, 1, 0)(1, 1, 1)_3$	265.83
6	$SARIMA(3, 1, 1)(1, 1, 1)_3$	243.35	6	$SARIMA(3, 1, 1)(1, 1, 1)_3$	268.51

Grupo 3			Grupo 1		
N°	Modelo	AIC	N°	Modelo	AIC
1	$SARIMA(3, 1, 0)(0, 1, 3)_3$	266.09	1	$SARIMA(3, 1, 0)(0, 1, 3)_3$	126.37
2	$SARIMA(3, 1, 0)(0, 1, 4)_3$	263.84	2	$SARIMA(3, 1, 0)(0, 1, 4)_3$	127.89
3	$SARIMA(3, 1, 0)(0, 1, 3)_3$	266.36	3	$SARIMA(3, 1, 0)(0, 1, 3)_3$	127.91
4	$SARIMA(4, 1, 0)(0, 1, 4)_3$	239.5	4	$SARIMA(4, 1, 0)(0, 1, 4)_3$	129.82
5	$SARIMA(3, 1, 0)(1, 1, 1)_3$	265.83	5	$SARIMA(3, 1, 0)(1, 1, 1)_3$	126.69
6	$SARIMA(3, 1, 1)(1, 1, 1)_3$	268.51	6	$SARIMA(3, 1, 1)(1, 1, 1)_3$	126.8

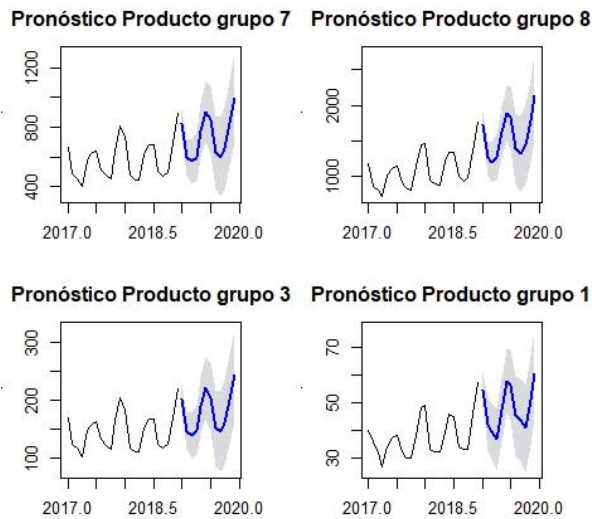


Figura 10: Pronósticos

7. Análisis de tablas de tres vías

En el trabajo de los inventarios para empresas comerciales, se tiene una gran cantidad de información, 2187 productos registrados en cada una de las 14 sedes de venta, es decir, que cada sede tiene 2187 productos analizados en 25 variables (precio y ventas mensuales de 2 años) para un total de 54675 registros, esto multiplicado en todas las sedes, porque éstas cuentan con el mismo número de productos y de variables. Lo que

interesa en este punto, es analizar las ventas de las sedes en cada mes, y esto conlleva a que no es sencillo realizar el análisis por producto, sería una tarea poco eficiente. Debido a lo anterior, se recurre al análisis de tres vías (datos longitudinales), en el cual se analiza el mismo grupo de individuos (vía 1), en el mismo grupo de variables (vía 2) en las sedes de venta (vía 3). La metodología multivariada se conoce como el método STATIS (Thioluse, *et al.* 2018). La figura 11 muestra cómo se organiza la información para ser ingresada en el programa R-Project usando la librería *ade4* para el análisis STATIS, la información de cada sede se ubica en una tabla y se hace un despliegue vertical. La esencia con el método STATIS es encontrar una estructura de las tablas, a través de la construcción de unos coeficientes de correlación vectorial que permitan calcular un matriz compromiso o consenso que es común a todas las tablas de datos. La matriz consenso tiene presente la covariabilidad (covarianza vectorial-*Covv*-) de todas las tablas y con ello se compara la varianza vectorial (*Vav*) de una tabla a otra para mirar si es estable, sino, es debido a cambios en el comportamiento de los individuos seleccionados en las variables medidas en cada situación.

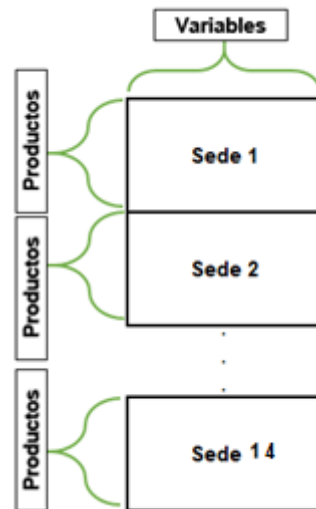


Figura 11: Organización de tablas para análisis de tres vías

El coeficiente vectorial Rv calculado entre pares de tablas es:

$$Rv(\mathbf{X}_k, \mathbf{X}_l) = \frac{Covv(\mathbf{X}_k, \mathbf{X}_l)}{\sqrt{Vav(\mathbf{X}_k)}\sqrt{Vav(\mathbf{X}_l)}} \quad (5)$$

(Thioluse, *et al.* 2018)

Rv también puede calcularse como el promedio de las correlaciones entre todos los pares de variables

$$Rv(\mathbf{X}_k, \mathbf{X}_l) = \frac{1}{p} \sum_{i=1}^p cor(\mathbf{X}_k^i, \mathbf{X}_l^i) \quad (6)$$

(Thioluse, *et al.* 2018)

Con p igual al número de variables y si éstas están estandarizadas entonces, $Vav(\mathbf{X}_k) = p$. \mathbf{X}_k^i y \mathbf{X}_l^i son los vectores columna asociados a las i -ésima variable de la tabla de la figura 7 conocida también como la tabla k porque es conformada por un conjunto de K tablas, donde cada una de estas es la situación considerada para el análisis de tres vías. La matriz \mathbf{Rv} contiene todos los coeficientes de correlación vectorial calculados por cada par de tablas (por cada par de situaciones), es una matriz simétrica. \mathbf{Rv} también es conocida como la Interestructura de la tabla k .

$$\mathbf{Rv} = \begin{bmatrix} Rv(\mathbf{X}_1, \mathbf{X}_1) & \cdots & Rv(\mathbf{X}_1, \mathbf{X}_K) \\ \vdots & \ddots & \vdots \\ Rv(\mathbf{X}_K, \mathbf{X}_1) & \cdots & Rv(\mathbf{X}_K, \mathbf{X}_K) \end{bmatrix} \quad (7)$$

Se toman los valores y vectores propios de \mathbf{Rv} para hacer una descomposición espectral y representar esta Interestructura en un espacio de dimensión reducida. En el programa ade4 de R-Project se tiene la función *pta* que traduce Análisis Parcial Triádico, el cual es uno de los métodos para analizar un objeto de k tablas, este debe tener el mismo número individuos y el mismo número de variables. La K tablas se proyectan a partir de la matriz consenso.

El matriz consenso se calcula como:

$$\mathbf{X} = \sum_{k=1}^K \alpha_k \mathbf{X}_K \quad (8)$$

Con α_k desde $k = 1, 2, \dots, K$ los valores de los coeficientes del primer vector propio de la inter estructura \mathbf{Rv} . Una de las ventajas del método STATIS al comparar k -tablas, es que la interpretación suele hacerse como en el método de Análisis de Componentes Principales (ACP), así los gráficos superiores de la figura 12 y el gráfico inferior derecho se interpretan análogamente al ACP, es útil para comparar las variables, los individuos en consenso. El gráfico inferior izquierdo proyecta en el espacio de dimensión reducida a las 14 tablas (cada una representa una sede). Esta figura pone en un plano *cartesiano* el peso de cada tabla (Tables weights) versus el coseno al cuadrado (Cos2) de cada una, recordando que este indicador se interpreta como la calidad de presentación que tiene en el problema, a mayor valor, se interpreta que la tabla está bien representada en el plano y se puede interpretar que tablas cercanas se parecen.

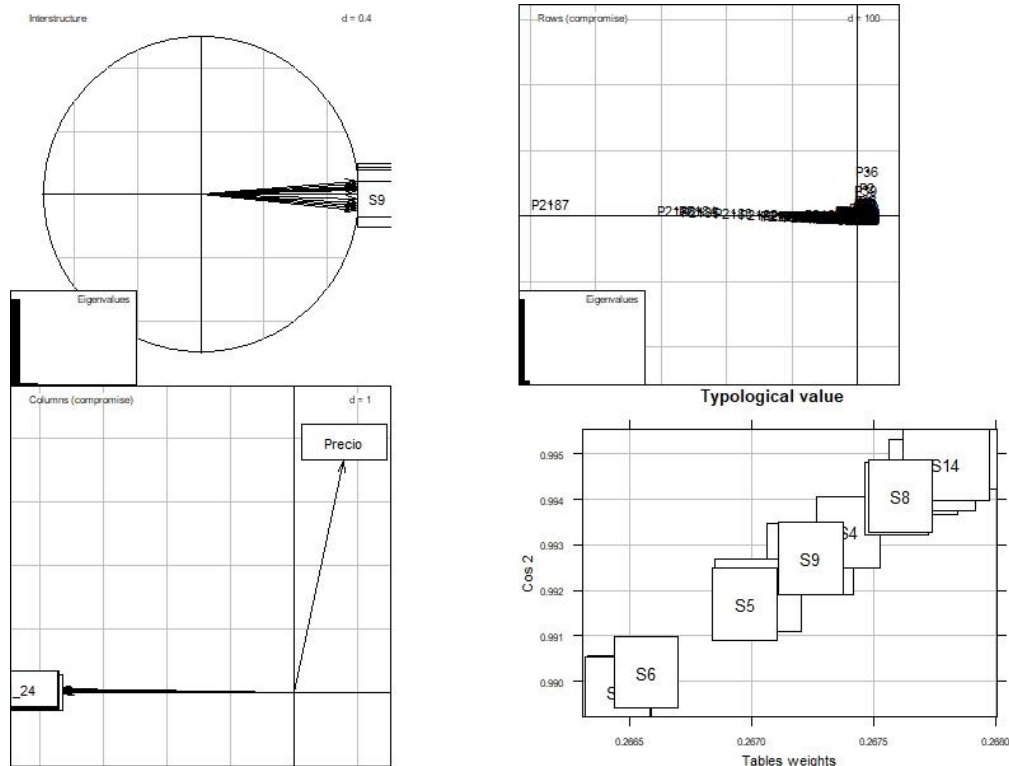


Figura 12: Resultado método STATIS

En el contexto de las sedes, se ha presentado de acuerdo con el gráfico inferior derecho de la figura 8 que hay algunos grupos de sedes, y ocurre que son las mismas planteadas en la tabla 3 del conglomerado para sedes. Este es el resultado que se obtiene del método STATIS y el cual es de gran utilidad para analizar una estructura en datos de tres vías, como le puede ocurrir a las empresas comercializadoras que manejan un número grande de referencias de productos y a la vez que manejan varios sitios de venta. En el caso de estudio, se ha hecho una clasificación de los productos realizando un análisis de conglomerados y en el caso de las sedes, se han estudiado mediante el método de conglomerados y también con el método STATIS herramientas multivariantes en estadística para el análisis de grandes bases de datos de las cuales se desean conocer comportamientos específicos, como por ejemplo el nivel de venta de cada sede, respecto al total de toda la empresa.

8. Conclusiones

Se logró con la técnica multivariada de conglomerados agrupar las sedes de venta de la empresa en 4 grupos, los cuales describían las que mayor participación en las ventas tenían hasta las de menor ubicadas el cuarto grupo. En el caso de los productos por la cantidad inmensa se aprovechó el método no jerárquico de *k medias* para agrupar 2187

artículos en 8 grupos, mostrando los que más cantidades rotaban e ingresos dejan a la empresa en los primeros grupos y en los demás, los de menores cantidades e ingresos.

Las series de tiempo son unas de las herramientas estadísticas que permiten estudiar el comportamiento de las ventas con una utilidad en la cual se pueda descomponer la serie en la componente de tendencia, estacionalidad, cíclico y error, en el estudio de caso analizado las series de venta mostraron una componente estacional muy marcada a pesar de contar con solo 24 datos de cada uno de los productos analizados. El comportamiento de las ventas se modeló con el modelo SARIMA capturando el comportamiento estacional en las ventas. Los productos seleccionados, al ser característicos del grupo al que pertenecían mostraron que los productos de manera particular en cada grupo también muestran un comportamiento bajo los modelos SARIMA.

El método STATIS permite analizar la Interestructura de un conjunto de datos longitudinales con observaciones medidas en tres modos o vías (individuos, variables y situaciones) que, para el caso abordado, los individuos son los 2187 productos, las variables son el precio de venta y las cantidades vendidas en 24 meses (cada mes como una variable), y las situaciones fueron cada una de las 14 sedes. El gráfico en cuestión no muestra bien las tablas por su salida gráfica del programa ade4, se extrae entonces la puntuación factorial de cada tabla ($x = \text{peso}$, $y = \text{coseno al cuadrado}$).

Las técnicas estadísticas utilizadas ayudan a que las empresas puedan mejorar los sistemas de control interno enfocado a la administración y gestión de inventarios, estudiando detenidamente las empresas comerciales las cuales manejan grandes volúmenes de datos y posiblemente diferentes puntos de venta.

Recibido: 2019-02-24

Aceptado: 2019-06-17

Referencias

- Bolaños, D. S. O. Cortes, O. E. M. Delgado, E. A. A. (2011). *Revisión y documentación del control interno para los procesos de inventarios y cartera de la compañía COMFASUAREZ S.A.S.* Universidad de San Buenaventura. Facultad de Ciencias Económicas. Valle del Cauca, Colombia.
- Díaz, L. G. M. Morales, M. A. R. (2012). *Análisis estadístico de datos multivariados.* 1ra ed. Universidad Nacional de Colombia. Bogotá, Colombia.
- Díaz, M. R. González A. L. A. Henao, A. P. (2013). *Introducción al análisis estadístico multivariado aplicado. Experiencias y casos en el Caribe colombiano.* Universidad del Norte. Colombia.
- Garrido, Y. I. B. Cejas, M. M. (2017). La gestión de inventario como factor estratégico en la administración de empresas. *Negotium*, 37(13), 109-129. Venezuela

Thioluse J. Dray, S. A. B. D. Siberchicot, T. J. Pavoine, S. (2018). *Multivariate analysis of ecological data with ade4*.

Velásquez, L. G. (2012). *Elaboración de una cédula como instrumento de gestión de inventario*. (Tesis para obtener título de químico farmacéutico biólogo). Universidad Nacional Autónoma de México. México.

Véliz, C. C. (2017). *Análisis multivariante. Métodos estadísticos multivariantes para la investigación*. Cengage Learning. México.

Wei, W. (2006). *Time Series Analysis. Univariate and multivariate methods*. Estados Unidos: Ed. Pearson.