

# Informe Técnico - Análisis Exploratorio de Datos (EDA)

## Proyecto ComercioYA

### 1. Introducción

El presente informe documenta el proceso de Análisis Exploratorio de Datos (EDA) realizado sobre el dataset de la empresa ComercioYA, dedicada a ventas en línea. El objetivo principal fue comprender el comportamiento de los clientes, identificar patrones, relaciones entre variables, valores atípicos y generar información útil para la toma de decisiones estratégicas.

El EDA permitió transformar datos crudos en información comprensible mediante estadísticas descriptivas y visualizaciones, facilitando la interpretación del comportamiento comercial.

### 2. Identificación de Variables

El dataset analizado contiene variables cuantitativas y categóricas:

Variables cuantitativas:

- Monto de compra
- Número de visitas
- Número de devoluciones
- Calificación del cliente
- Edad

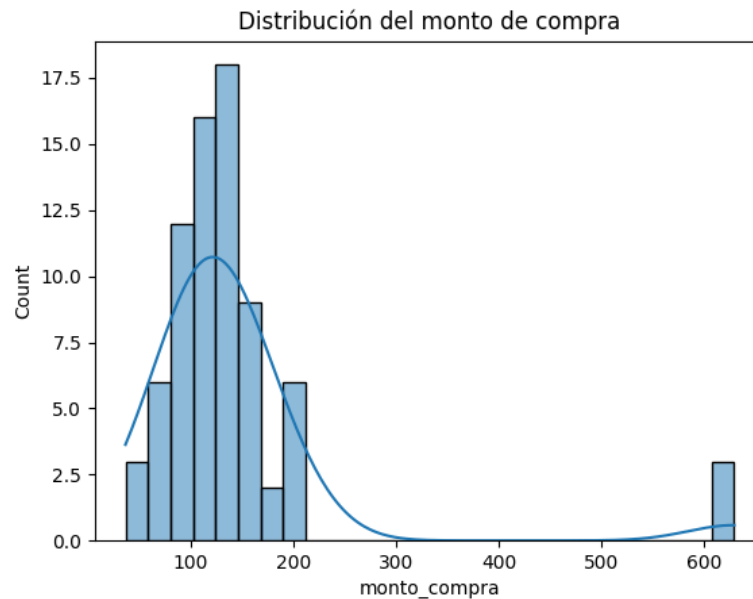
Variables categóricas:

- Género
- Ciudad
- Categoría del producto

### 3. Análisis Estadístico Descriptivo

Se calcularon medidas de tendencia central como media, mediana y moda, además de medidas de dispersión como varianza y desviación estándar. También se determinaron cuartiles y percentiles.

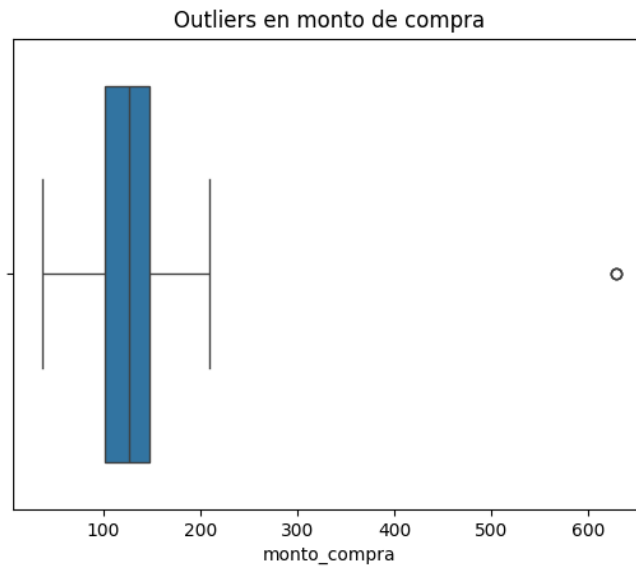
Los histogramas mostraron la distribución de los montos de compra y visitas, mientras que los boxplots permitieron identificar valores atípicos que podrían distorsionar el análisis.



### 4. Detección de Valores Atípicos

Mediante el uso de boxplots y análisis de IQR, se detectaron valores extremos en las variables de monto y visitas. Estos outliers representan clientes con comportamientos inusuales, como compras muy elevadas o visitas excesivas.

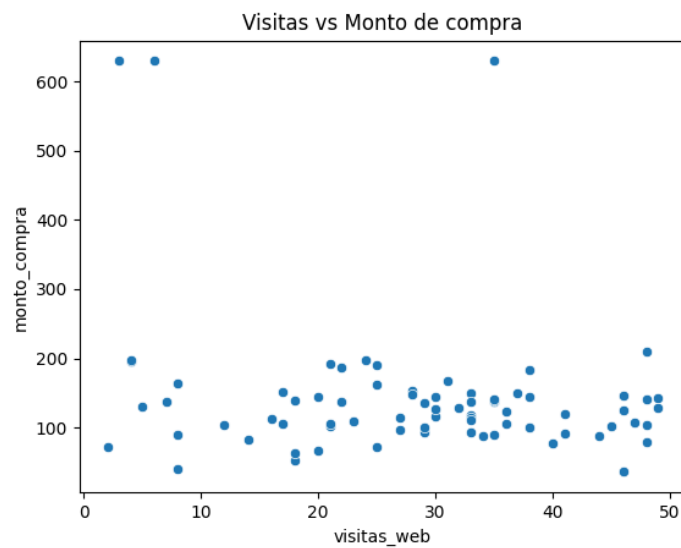
Se decidió conservar estos valores ya que representan casos reales importantes para el negocio.



## 5. Correlación entre Variables

Se construyeron scatterplots y una matriz de correlación para identificar relaciones entre variables numéricas. Se calculó el coeficiente de correlación de Pearson.

Se observó una relación positiva entre el número de visitas y el monto de compra, indicando que clientes más activos tienden a comprar más.



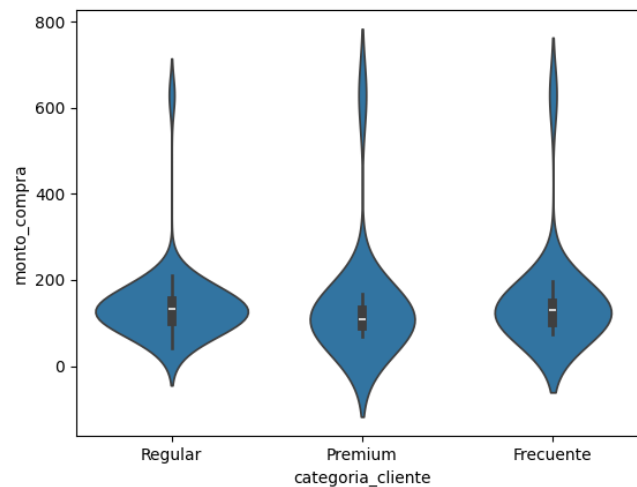
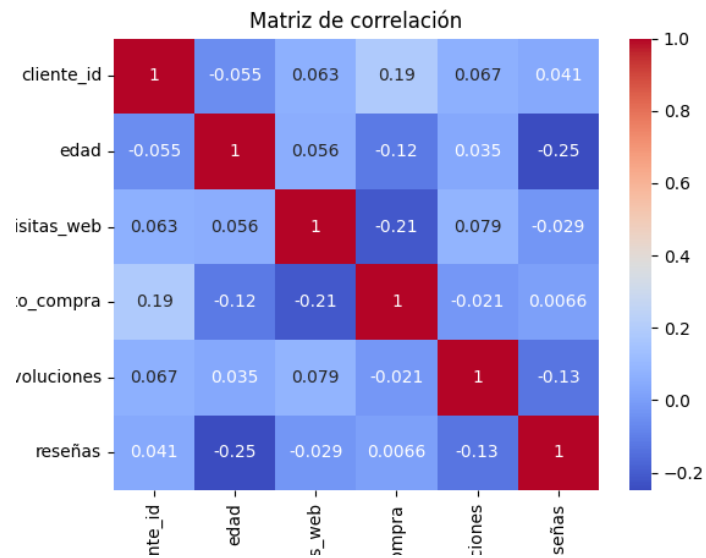
## 6. Regresión Lineal

Se aplicaron modelos de regresión lineal simple y múltiple utilizando la librería statsmodels.

Los resultados mostraron que la variable visitas es un predictor significativo del monto de compra.

## 7. Visualizaciones Avanzadas

Se utilizaron gráficos como pairplot, violinplot, jointplot, heatmap y FacetGrid para analizar múltiples variables simultáneamente.



## 8. Conclusiones y Recomendaciones

El análisis exploratorio permitió identificar patrones claros de comportamiento del cliente. Se recomienda utilizar estos hallazgos para diseñar campañas de marketing segmentadas y mejorar las estrategias de retención de clientes.