

Análisis de Patrones de Ventas

Luis Alan Morales Castillo A01659147

Paulina Díaz Arroyo A010295932

Rodrigo Jiménez Ortiz A01029623



Problema

Necesidad

Mejorar la asignación de inventario, optimizar rotación de productos y anticipar demanda.

Impacto actual

Exceso de stock en algunas zonas, rupturas de stock en otras, y variación significativa en ticket y frecuencia por método de pago.

Objetivo

Detectar tendencias y estacionalidad por zona, categoría y sector para decisiones de abasto, operaciones y comercial.

Objetivos Generales

Integrar análisis exploratorio, segmentación de productos y modelos predictivos para:

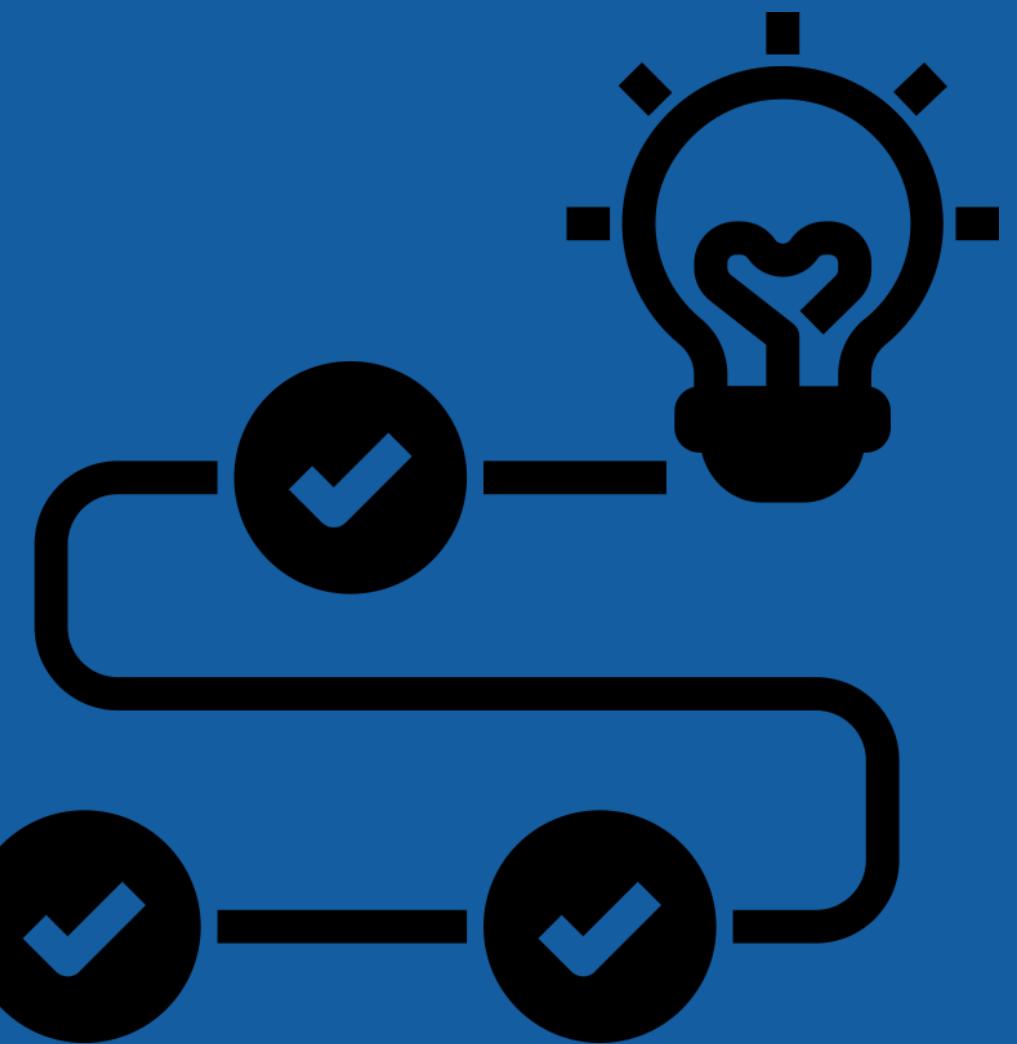
- 01 Detectar tendencias, estacionalidad y patrones clave de demanda.
- 02 Clasificar productos según comportamiento de venta y rotación.
- 03 Predecir ingresos con modelos multivariados y de boosting.
- 04 Generar lineamientos accionables para abasto, surtido, pricing y promociones.

Metodología

Librerías utilizadas: pandas, matplotlib, seaborn, plotly, sklearn, statsmodels, scipy, numpy, XGBoost.

Flujo del análisis:

1. Carga y limpieza (corrección de formato de precio, tipos de datos).
2. Integración multitablea (merge de ventas, productos, clientes, métodos de pago).
3. Creación de variables clave:
 - Cantidad_dinero = Cantidad × Precio_Unitario
 - Stock actual, porcentaje de venta, demanda acumulada.
4. Normalización, clustering y modelado predictivo.
5. Dashboard final interactivo.



Hallazgos

Hallazgo 1: Análisis Exploratorio (EDA)

- Se identifican diferencias marcadas entre categorías en volumen, precio y comportamiento de venta.
- Correlaciones significativas: Cantidad vendida, Precio Unitario, Stock y Cantidad_dinero.
- Boxplots y distribuciones revelan presencia de productos “estrella”, “estables” y “rezagados”.
- Se observa concentración de ventas en ciertas categorías → oportunidad para diversificar surtido.

Hallazgo 2: Segmentación de Productos (Clustering KMeans k=4)

Variables del clustering: Cantidad_vendida, Porcentaje_venta, Stock, Producto_actual_stock

Resultados:

- Cluster 0: Alta existencia + baja rotación → sobre-stock. Stock promedio muy alto (4,579) vs. ventas bajas (6.1% del stock).
- Cluster 1: Rotación saludable y buena relación venta/stock. % de venta elevado (14.7%) con stock moderado.
- Cluster 2: Productos estables de rotación media. Ventas y stock equilibrados, comportamiento intermedio.
- Cluster 3: Alta demanda + alta rotación → productos estrella. Mayor porcentaje de venta (18.6%) y mejor salida del inventario.

Hallazgos

Hallazgo 3: Segmentación de Clientes (k=6)

Variables: frecuencia de compra, gasto total, ticket promedio y variedad de categorías

- Cluster 0: Clientes Premium + Alta Frecuencia. Frecuencia y gasto muy por encima del promedio. Clientes más rentables, fuerte afinidad y comportamiento diversificado.
- Cluster 1: Compradores de Nicho. Frecuencia moderada pero centrados en pocas categorías específicas → Requieren estrategias personalizadas por categoría.
- Cluster 2: Clientes Estables. Frecuencia y gasto en niveles intermedios. Variedad moderada → Base sólida para planes de retención y upselling.
- Cluster 3: Ocasionales pero Diversos → baja recurrencia, alta variedad → Clientes oportunistas y estacionales, con potencial para aumentar frecuencia.
- Cluster 4: Cazadores de Oferta → precio-sensibles y baja variedad. Menor diversidad de productos y comportamiento centrado en descuentos → Ideal para campañas tácticas (cupones, combos).
- Cluster 5: Nuevos o Dormidos → ticket alto pero poca actividad. Muy baja variedad y pocos clientes, comportamiento irregular → Clientes por activar, clave para estrategias de bienvenida o reactivación.

Hallazgos

Hallazgo 4: Modelos Predictivos

Se ejecutaron dos enfoques:

- Modelos OLS por clúster:
 - Permiten entender variaciones internas de cada segmento.
 - Entrega R^2 y R^2 ajustado detallado.
 - En algunos clústeres, Precio_Unitario y Cantidad son fuertes predictores.
- Modelos globales de Boosting:
 - Gradient Boosting Regressor
 - AdaBoost Regressor
 - XGBoost Regressor

Hallazgo clave:

- Los modelos de boosting superan a la regresión múltiple tradicional en RMSE y R^2 , mostrando mejor capacidad de generalización.



Impacto Esperado

- Optimización del inventario: menor sobre-stock y reducción del riesgo de stock-out.
- Incremento en ingresos: promociones y segmentación más precisas elevan ticket y recompra.
- Planeación inteligente: modelos predictivos mejoran la estimación de demanda e ingresos.
- Eficiencia operativa: mejor alineación entre ventas, compras y logística.
- Marketing más efectivo: estrategias diferenciadas por tipo de cliente y producto.

Recomendaciones

Enfocar inventario en productos del cluster 3

Hacer encuestas post-compra para enriquecer el análisis

Personalizar promociones por tipo de cliente

Aplicar estrategias de retención y reactivación según cluster

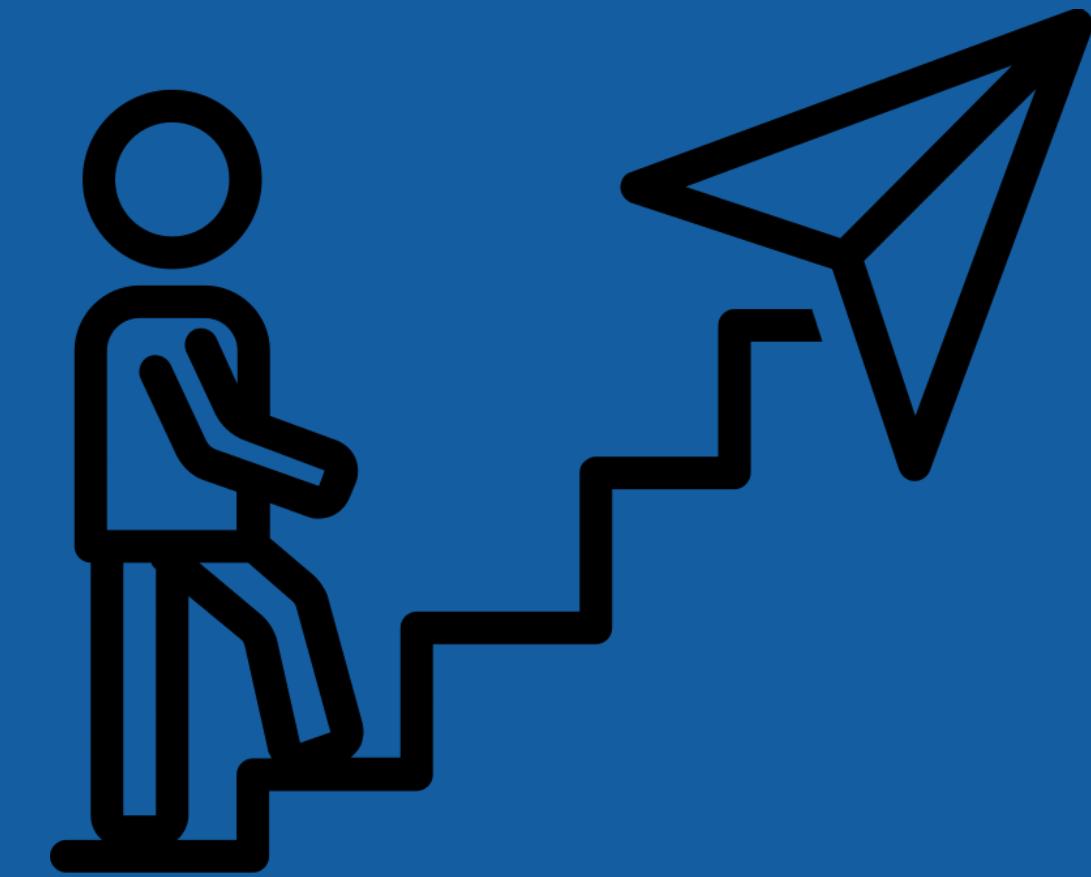
Reentrenar modelos con datos actualizados cada trimestre

Conclusiones y siguientes pasos

El uso integrado de análisis exploratorio, segmentación mediante clustering y modelos predictivos ofrece una visión completa del comportamiento comercial, del stock de productos y de los distintos tipos de clientes. Esta combinación permite entender no solo qué ocurre en términos de ventas, inventario y hábitos de compra, sino también por qué sucede y cómo puede evolucionar en el corto y mediano plazo.

Con esta perspectiva, es posible optimizar la gestión de inventarios al identificar productos con sobre-stock, alta rotación o demanda estable, diseñar estrategias comerciales personalizadas según perfiles reales de clientes y anticipar patrones de demanda mediante modelos más precisos que los enfoques tradicionales.

El refinamiento de los modelos fortalecerá la segmentación y permitirá decisiones más específicas y alineadas con el comportamiento real del mercado.



GRACIAS