

ANÁLISIS PREDICTIVO DE VENTAS RETAIL MEDIANTE CIENCIA DE DATOS

Presentado por:

- Arion Bardo Arianda
- Luisa Fernanda Vázquez Razo
- Edgar Cabrera Velazquez
- Juan Francisco Osorno Martínez
- Sebastián Rodríguez Martínez
- Victor Manuel Ortíz Feregrino

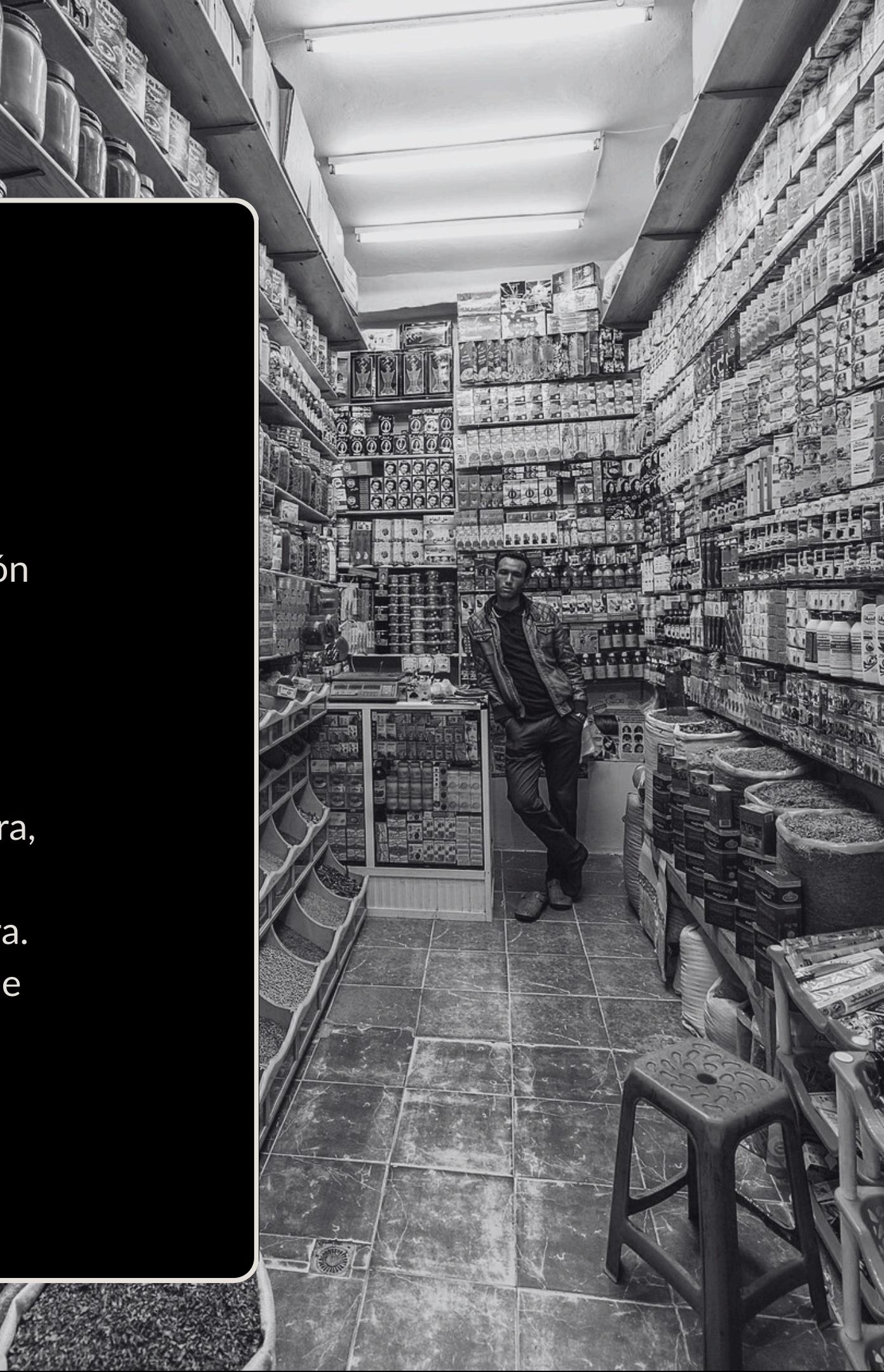




CONTEXTO DEL PROYECTO

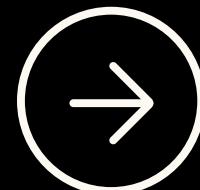
Entender el comportamiento del cliente se vuelve una tarea fundamental. Las empresas ya no pueden basar sus decisiones únicamente en la intuición o en la experiencia pasada, sino que necesitan apoyarse en información concreta. Por eso nuestro propósito es:

Analizar un conjunto de datos de ventas minorista con el fin de **identificar tendencias importantes, patrones de compra, factores que explican variaciones en el gasto de los clientes y relaciones entre distintos elementos del comportamiento de compra.** Con los resultados obtenidos, sería posible proponer **estrategias** que mejoren el rendimiento general de una tienda minorista.





RETAILX & DATASET OVERVIEW



RetailX: Empresa Analizada

RetailX es una tienda retail ficticia que vende productos de consumo en tres grandes categorías: Electronics, Clothing y Beauty. Funciona como una tienda departamental típica, donde cada compra queda registrada con información del cliente y del producto, lo que la vuelve ideal para practicar análisis de datos reales simulados.

Naturaleza del Dataset

El proyecto utiliza el Retail Sales Dataset de Kaggle, con 1000 registros de ventas simuladas. Aunque es un dataset sintético, sus patrones y distribuciones se parecen mucho a los de una tienda real, permitiendo explorar tendencias de venta, comportamiento de clientes y variaciones en el gasto.

Información Registrada en Cada Venta

Cada transacción incluye datos como: fecha de compra, edad del cliente, categoría del producto, cantidad adquirida, precio por unidad y total pagado. Con estas variables se pueden analizar temporadas de alta venta, segmentos de clientes y factores que impactan el monto total de compra.

IMPORTANCIA EN EL ANÁLISIS DE DATOS

DETECTAR CAMBIOS

Detectar cómo cambian las ventas a lo largo del tiempo, ya sea por temporadas, promociones o situaciones externas.

ENCONTRAR PATRONES

Encontrar patrones que ayuden a dividir a los clientes en grupos con características similares.

ANALIZAR DATOS

Analizar qué productos tienen precios más sensibles para los consumidores.

REVISAR DATOS ANORMALES

Reconocer qué categorías de productos registran mayores ventas y cuáles tienen menor movimiento.

OBJETIVO Y PREGUNTAS DE ANÁLISIS



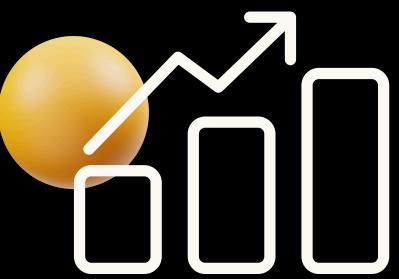
Objetivo general

Realizar un análisis integral del conjunto de datos de ventas de RetailX para identificar patrones de compra, tendencias principales y relaciones entre variables como edad del cliente, categoría de producto, precio, cantidad y monto total. Con estos hallazgos se busca apoyar la toma de decisiones en el negocio retail.



Preguntas sobre clientes y gasto

- ¿La edad del cliente influye en cuánto gasta?
- ¿Qué categorías de producto generan mayores ventas?
- ¿Qué tipo de clientes realizan las compras de mayor monto?



Preguntas sobre tiempo y comportamiento

- ¿En qué meses se registra mayor actividad de compra?
- ¿Existen patrones de comportamiento que permitan agrupar a los clientes en segmentos?
- ¿Cómo cambian las ventas a lo largo del tiempo?

METODOLOGÍA DE CIENCIA DE DATOS



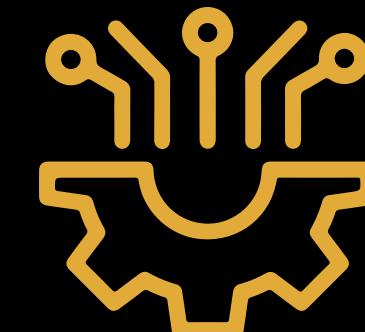
Entendimiento
del problema



Data Wrangling



Modelación
con Machine
Learning



Visualizaciones finales



ENTENDIMIENTO DEL PROBLEMA

CONTEXTO DEL NEGOCIO

RetailX registra cientos de ventas al día en distintas categorías de producto. Tiene mucha información disponible, pero no cuenta con un análisis sistemático que le permita aprovecharla para tomar decisiones.

PROBLEMA CENTRAL

Los datos se almacenan, pero no se transforman en conocimiento. La empresa no sabe con claridad qué factores influyen en el monto de compra, qué clientes generan más ingresos ni qué productos son más relevantes.

RETOS ACTUALES

No existe segmentación de clientes, se desconocen las temporadas fuertes de venta y no hay métricas claras para evaluar el desempeño de las categorías. Esto limita la planeación de inventario y de estrategias comerciales.

PREGUNTAS CLAVE INICIALES

¿Qué patrones se esconden en las ventas? ¿Quiénes son los clientes más valiosos? ¿Cómo cambian las compras a lo largo del tiempo? Estas preguntas guían el análisis de datos que se desarrolla en el proyecto.



RELEVANCIA Y JUSTIFICACIÓN DEL ENFOQUE ANALÍTICO

Para una empresa retail, entender cómo y cuándo se venden sus productos es fundamental para sobrevivir y crecer.

Si la empresa no analiza su información histórica, corre el riesgo de tomar malas decisiones, como comprar productos que no se venden, no identificar temporadas de alta demanda, o aplicar promociones que no generan resultados. En cambio, si RetailX usa sus datos de forma correcta, puede identificar sus productos más importantes, conocer mejor a sus clientes y mejorar sus ganancias.



DIAGRAMAS AS-IS

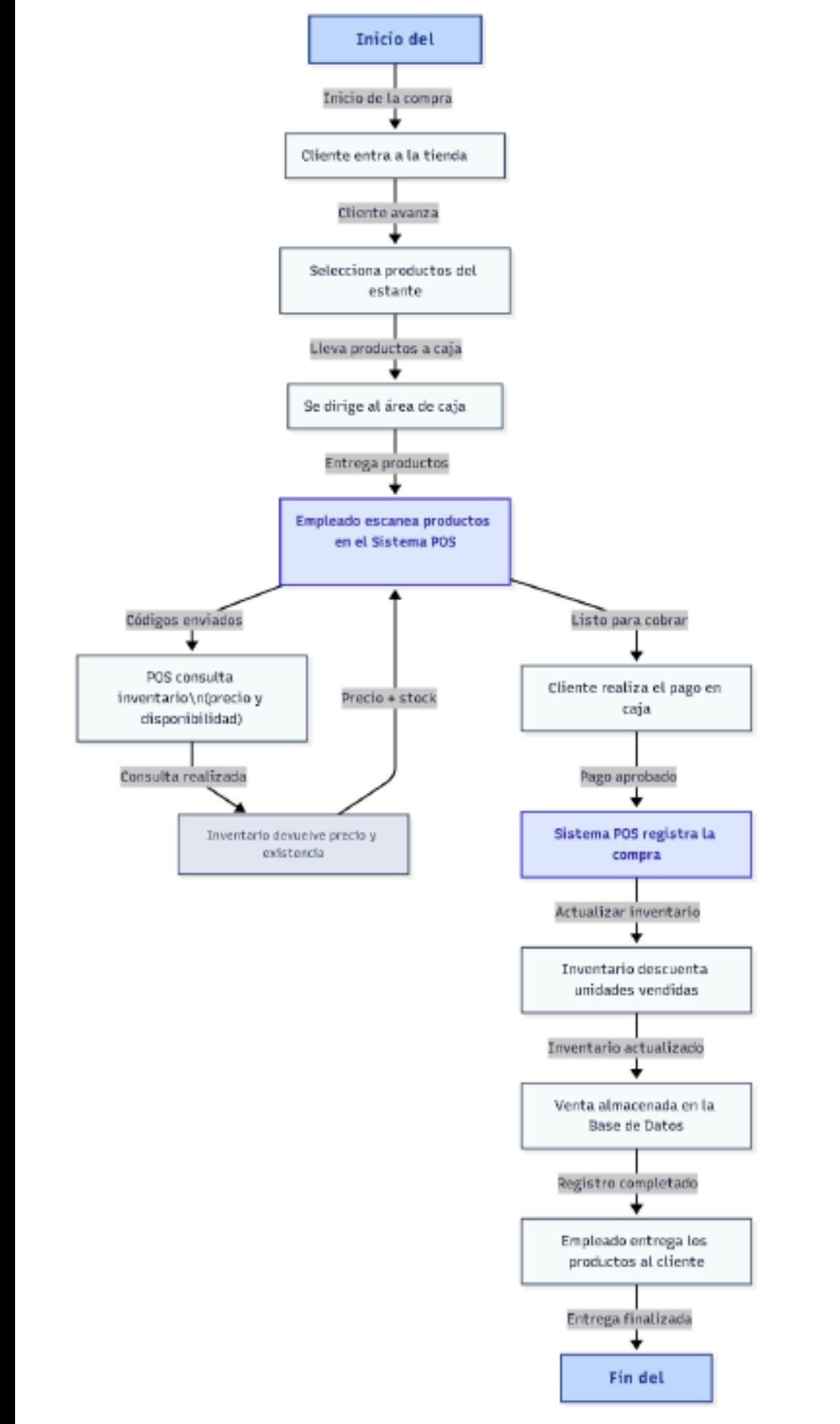
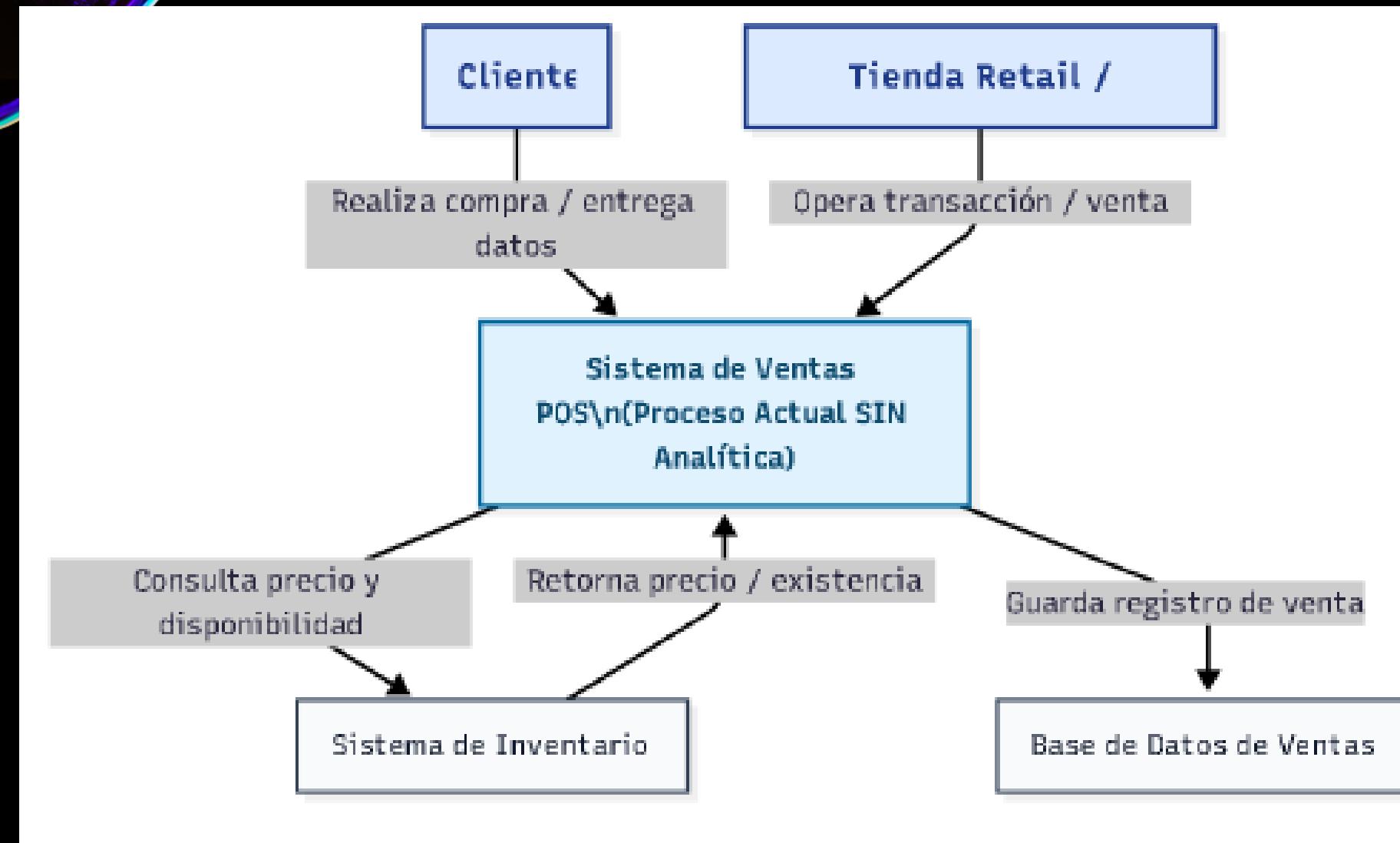
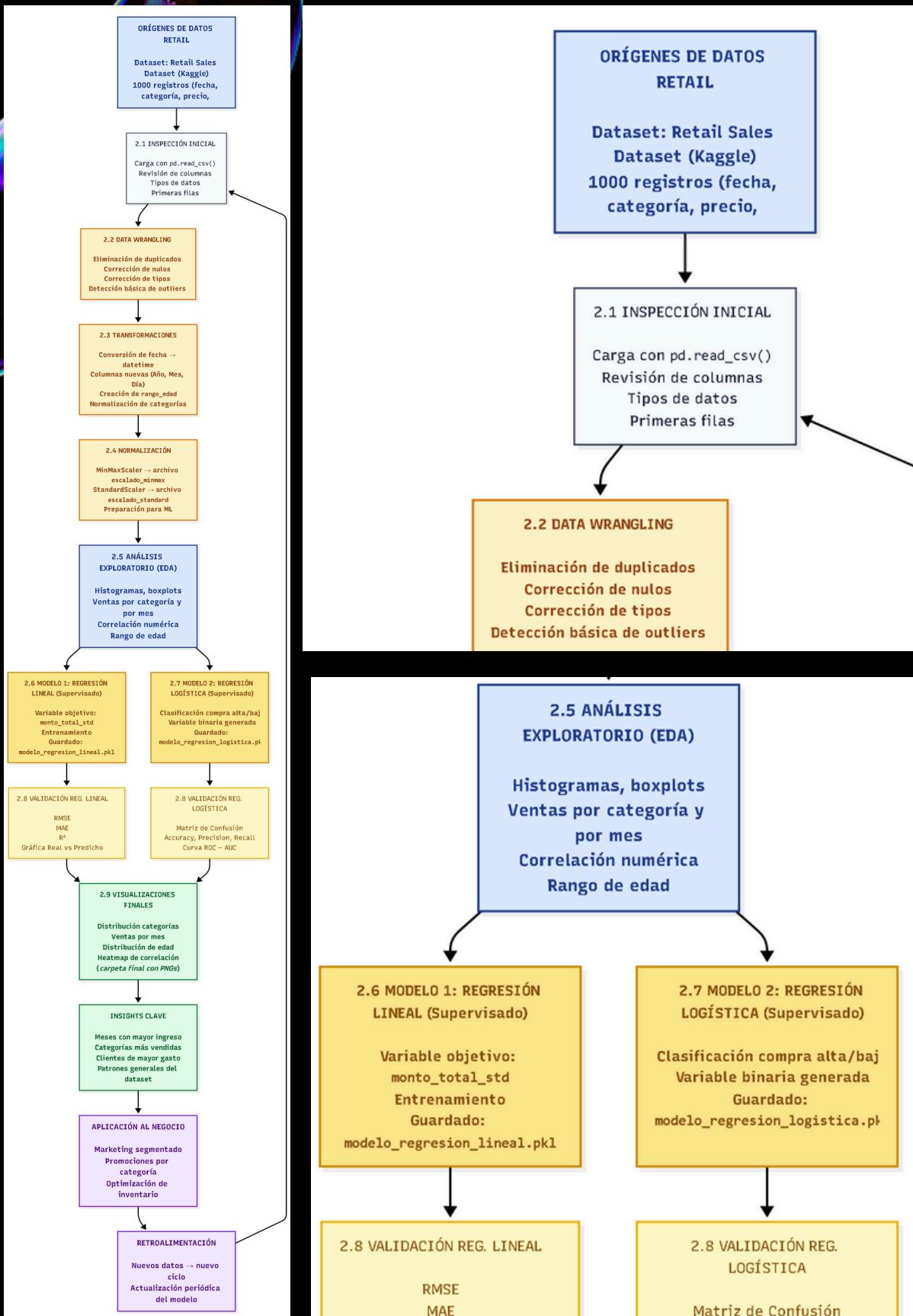


DIAGRAMA TO-BE

El diagrama TO-BE muestra cómo se vería el sistema una vez que se integra un proceso formal de Ciencia de Datos. Aquí los datos ya no solo se guardan, sino que se transforman y analizan para generar información útil para la toma de decisiones.

Los Elementos serían:

- La base de datos.
- El proceso de Data Wrangling.
- Los modelos de Machine Learning (Regresión Lineal y Logística).
- Un módulo de visualización.
- La capa de insights, que resume los resultados clave.
- Un ciclo de retroalimentación.



DATA WRANGLING Y ESTRUCTURA DEL PROYECTO

Origen del dataset

Se trabaja con un dataset de ventas minoristas (RetailX) que registra transacciones con información de cliente, producto y monto de compra. Este conjunto de datos sirve como base para explorar patrones de venta y comportamiento de los clientes.

Limpieza y preparación de datos

Se eliminaron registros duplicados, se corrigieron tipos de datos, se trajeron valores nulos y se validaron los rangos de las variables numéricas. Con esto se garantizó que el análisis se realizara sobre información consistente y confiable.



Estructura del proyecto

El proyecto se organizó en carpetas y notebooks separados para: exploración de datos, limpieza, generación de variables y modelado. Esta estructura facilita seguir el flujo de trabajo y reutilizar el código.

Variables derivadas y dataset final

A partir de las columnas originales se generaron nuevas variables (como fecha desglosada o categorías de gasto), y se construyó un dataset final listo para análisis estadístico y modelos de Machine Learning.

DATA WRANGLING

LIMPIEZA DE DATOS

En esta etapa se dejó listo el dataset para poder aplicar transformaciones, EDA y modelos, trabajando siempre sobre las columnas con nombres ya estandarizados en español:

Con todos estos pasos, el dataset quedó:

- Sin duplicados.
- Sin valores nulos en columnas críticas.
- Con tipos de datos correctos (especialmente fecha).
- Con nombres de columnas estandarizados en español.
- Con estadísticas numéricas revisadas para detectar posibles valores atípicos.

- fecha
- genero
- edad
- categoria_producto
- cantidad
- precio_unitario
- monto_total



VARIABLES DERIVADAS

- Se descompuso la fecha de compra en año, mes y día para analizar la evolución de las ventas en el tiempo.
- Se creó una variable de rango de edad (jóvenes, adultos, adultos mayores) para facilitar comparaciones entre grupos.
- Se generó el monto total por compra a partir de cantidad y precio unitario.
- Se añadió un indicador de compra alta/baja según el monto total (por encima o por debajo de un umbral definido).

SEGMENTACIÓN DE CLIENTES

- Los clientes se agruparon por rango de edad para identificar qué grupo consume más.
- Se analizaron segmentos según categoría de producto (por ejemplo, quienes compran más Electronics vs Clothing).
- Se identificaron clientes con alto nivel de gasto para posibles estrategias de fidelización.
- Esta segmentación permite diseñar campañas y decisiones específicas para cada grupo de clientes.

TRANSFORMACIÓN – RESULTADOS

```
# Vista general de algunas columnas clave
df[['id_transaccion',
   'fecha',
   'id_cliente',
   'genero',
   'edad',
   'rango_edad',
   'categoria_producto',
   'cantidad',
   'precio_unitario',
   'monto_total',
   'nivel_compra',
   'es_compra_alta']].head()
```

[14]

	id_transaccion	fecha	id_cliente	genero	edad	rango_edad	categoria_producto	cantidad	precio_unitario	monto_t
0	1	2023-11-24	CUST001	MALE	34	Adulto	Beauty	3	50	150
1	2	2023-02-27	CUST002	FEMALE	26	Adulto	Clothing	2	500	1000
2	3	2023-01-13	CUST003	MALE	50	Mayor	Electronics	1	30	30
3	4	2023-05-21	CUST004	MALE	37	Adulto	Clothing	1	500	500
4	5	2023-05-06	CUST005	MALE	30	Adulto	Beauty	2	50	100

Los resultados fueron:

Conversión correcta de tipos de datos, especialmente la columna fecha.

- Creación de variables temporales.
- Segmentación de clientes por rango_edad.
- Normalización de texto en genero, categoria_producto e id_cliente.
- Verificación de consistencia entre cantidad.
- Definición de variables de negocio que se usarán para el modelo de clasificación.
- Generación del archivo final retail_sales_transformado.csv, que servirá como entrada para la Normalización (2.4), el EDA (2.5) y los modelos de Machine Learning.



Transformación

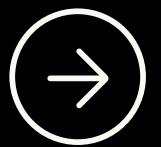


NORMALIZACIÓN DEL DATASET Y CONSTRUCCIÓN FINAL

La normalización es una fase fundamental dentro del proceso de preparación de datos, especialmente cuando se trabajará con modelos de Machine Learning. Muchos algoritmos son sensibles a las diferencias de escala entre variables, lo que puede provocar resultados incorrectos o modelos sesgados.

El objetivo principal es garantizar que todas las variables numéricas tengan una escala comparable para evitar que atributos con valores grandes, como monto_total, dominen el comportamiento del modelo frente a valores pequeños como cantidad.





NORMALIZACIÓN – RESULTADOS

```
df_normalizado = pd.concat([df, df_minmax, df_standard], axis=1)

df_normalizado.head()
```

Python

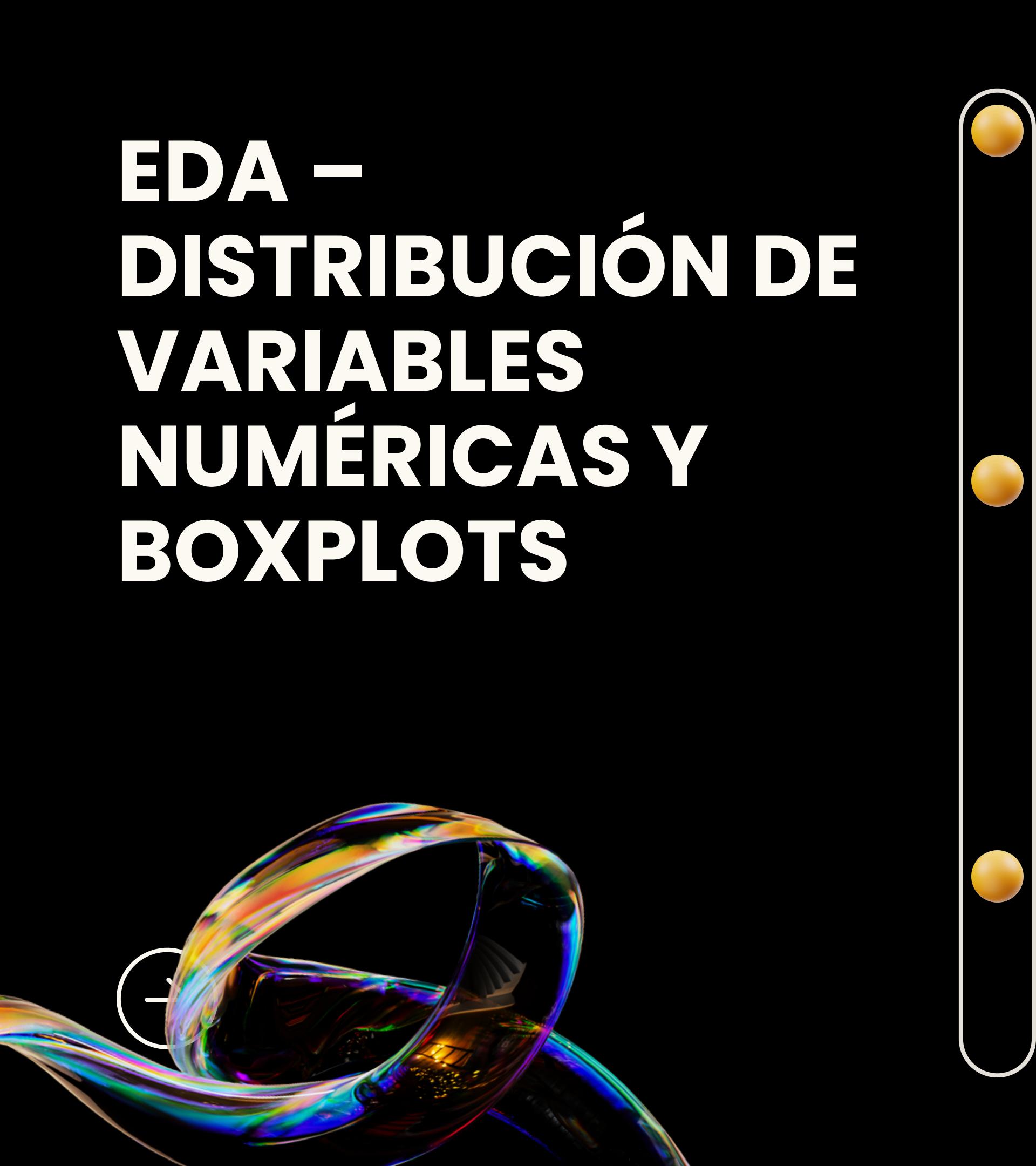
	id_transaccion	fecha	id_cliente	genero	edad	categoria_producto	cantidad	precio_unitario	monto_total	anio	...	nivel_compra	es_compra_fija
0	1	2023-11-24	CUST001	MALE	34	Beauty	3	50	150	2023	...	Baja	Fija
1	2	2023-02-27	CUST002	FEMALE	26	Clothing	2	500	1000	2023	...	Alta	Fija
2	3	2023-01-13	CUST003	MALE	50	Electronics	1	30	30	2023	...	Baja	Fija
3	4	2023-05-21	CUST004	MALE	37	Clothing	1	500	500	2023	...	Media	Fija
4	5	2023-05-06	CUST005	MALE	30	Beauty	2	50	100	2023	...	Baja	Fija

5 rows × 26 columns

La normalización permitió que todas las columnas numéricas quedaran en escalas comparables, evitando sesgos en los algoritmos de Machine Learning. Esta etapa garantiza que el comportamiento de los modelos no dependa del rango original de los datos.

Con esta fase completada, el dataset está completamente listo para continuar hacia:

- Análisis Exploratorio (EDA)
- Modelo de Regresión Lineal
- Regresión Logística
- Validación de modelos



EDA – DISTRIBUCIÓN DE VARIABLES NUMÉRICAS Y BOXPLOTS

Variables numéricas analizadas

Se estudiaron principalmente las variables numéricas del dataset: edad del cliente, cantidad comprada, precio unitario y monto total de la compra. Estas variables permiten describir el comportamiento general de las ventas y del gasto de los clientes.

Distribución de las variables

Mediante histogramas se observó cómo se distribuyen estas variables: rangos más frecuentes de edad, cantidades típicas compradas y niveles de precio más comunes. Esto ayuda a detectar si los datos están concentrados, dispersos o sesgados hacia ciertos valores.

Boxplots y valores atípicos

Con boxplots se identificaron valores atípicos en monto total, cantidad y precio, así como la mediana y la variabilidad de cada variable. Estos gráficos permiten comparar rápidamente la dispersión y localizar compras inusualmente altas o bajas que pueden influir en el análisis.

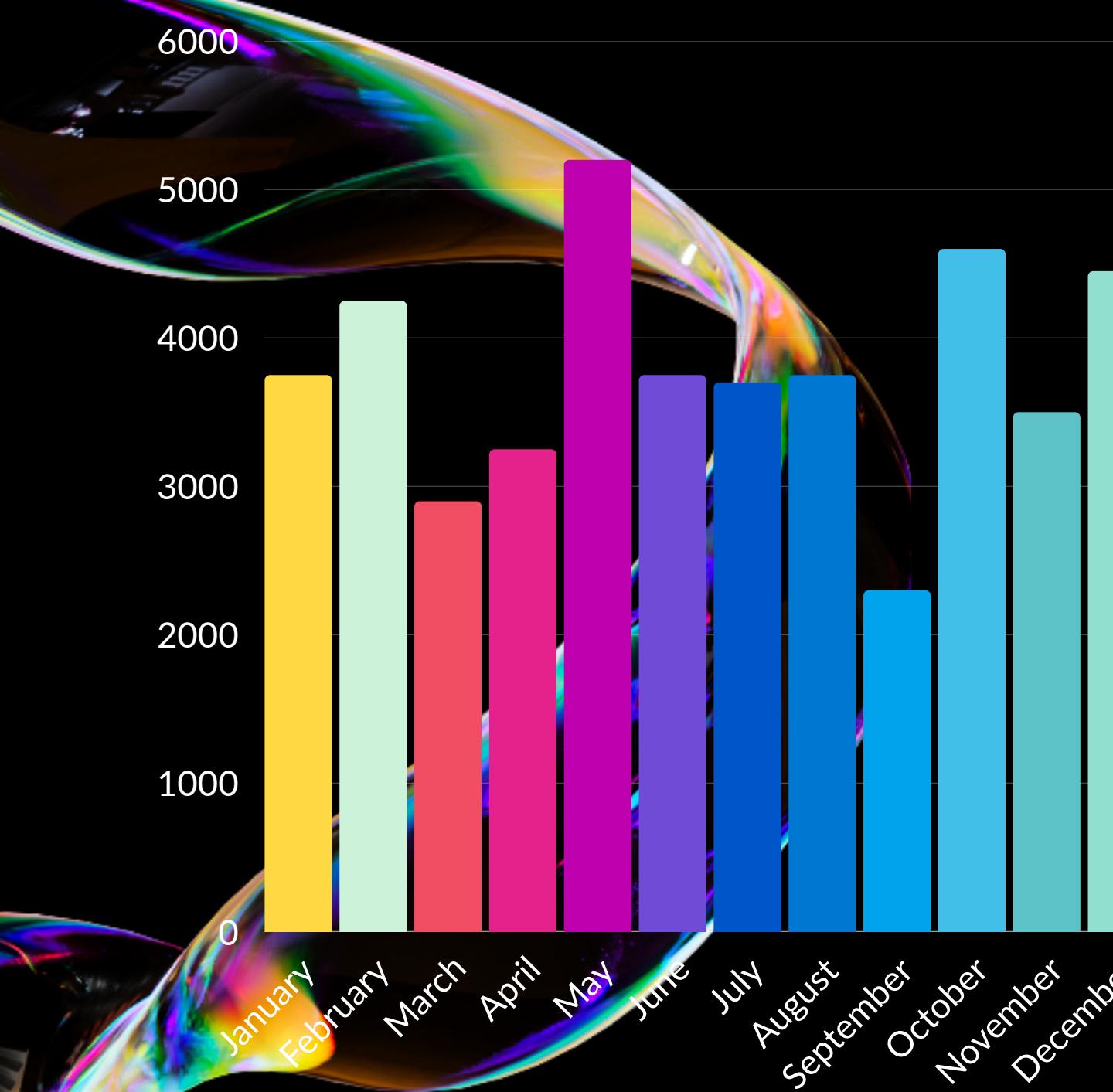
EDA - VENTAS POR CATEGORÍA



Interpretación:

- A. Las categorías Clothing y Electronics concentran el mayor volumen de ventas.
- B. Beauty también tiene participación relevante, pero ligeramente menor.

EDA - VENTAS POR MES



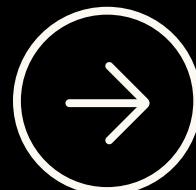
Interpretación:

A. El mes con mayores ventas fue mayo, seguido de octubre y diciembre.

Meses como marzo y septiembre muestran menor actividad.



EDA – VENTAS POR RANGO DE EDAD Y CONCLUSIONES S DEL EDA



Ventas por rango de edad

Se agruparon los clientes en rangos de edad (por ejemplo, jóvenes, adultos y adultos mayores) para analizar cómo varía el monto de compra entre grupos. Esto permitió identificar qué segmento concentra la mayor parte de las ventas y quiénes realizan, en promedio, las compras de mayor valor.

Comportamiento de compra por segmento

Al comparar los rangos de edad, se observaron diferencias en la preferencia de categorías de producto y en la frecuencia de compra. Algunos grupos muestran mayor interés por ciertas categorías, mientras que otros realizan compras más espaciadas pero de mayor monto.

Conclusiones generales del EDA

El análisis exploratorio permitió: detectar patrones de gasto por edad, identificar meses y categorías con mayores ventas, encontrar valores atípicos en los montos de compra y entender mejor la estructura del dataset. Estos hallazgos sirven como base para el modelado y para proponer estrategias comerciales más enfocadas.



MODELACIÓN CON MACHINE LEARNING – PANORAMA GENERAL

En la modelación con machine learning se busca que los modelos aprendan patrones a partir de la información del dataset para poder hacer predicciones o clasificaciones de manera automática. Los modelos que utilizamos son:

Regresión Lineal: se enfoca en predecir valores numéricos continuos. En este proyecto, se usa para estimar el monto_total de una compra a partir de variables como la edad del cliente, la cantidad adquirida y el precio por unidad.

- Regresión Logística: se utiliza para clasificar resultados en dos categorías. En nuestro caso, permite identificar si una compra debe considerarse como alta o no alta.



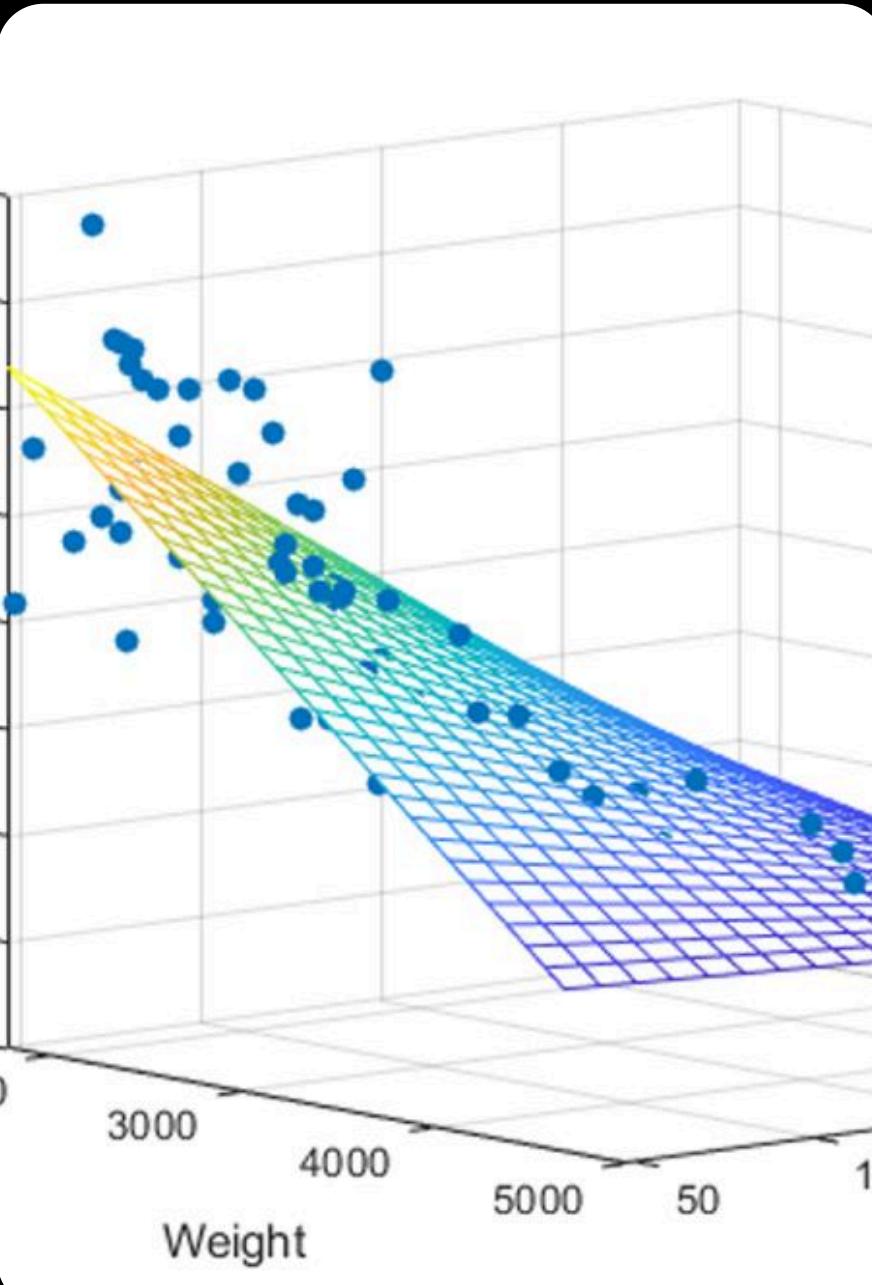
MODELO 1: REGRESIÓN LINEAL

Objetivo del modelo

Se implementó un modelo de Regresión Lineal para predecir el monto total de la compra a partir de distintas características del cliente y del producto. La idea es estimar cuánto podría gastar un cliente dado su perfil y la información de la venta.

Variable de entradas

Como variables explicativas se utilizaron, entre otras: edad del cliente, categoría del producto (codificada), cantidad adquirida y precio unitario. Estas variables se estandarizaron para mejorar la estabilidad del modelo



Proceso de entrenamiento

El dataset se dividió en conjuntos de entrenamiento y prueba. Con los datos de entrenamiento se ajustaron los coeficientes del modelo y se verificó que no hubiera problemas graves de multicolinealidad ni valores extremos que distorsionaran los resultados.

Salida e interpretación

El modelo genera un valor numérico predicho de monto total para cada transacción. Estos resultados permiten comparar el gasto esperado con el gasto real y sirven como base para evaluar qué factores influyen más en el importe de las compras.

ALGUNOS DE LOS CODIGOS USADOS

Definicion de las variables “X”, “Y”

```
# Selección de columnas predictoras (variables independientes)
# Usamos las versiones estandarizadas (_std) generadas en 2.4_Normalización
X = df[['edad_std', 'cantidad_std', 'precio_unitario_std']]

# Variable objetivo (monto total estandarizado)
Y = df['monto_total_std']

X.head(), Y.head()

[18] Python
```

... (edad_std cantidad_std precio_unitario_std)
0 -0.540565 0.429265 -0.685123
1 -1.125592 -0.453996 1.688464
2 0.629489 -1.337258 -0.790615
3 -0.321180 -1.337258 1.688464
4 -0.833078 -0.453996 -0.685123,
0 -0.546704
1 0.971919
2 -0.761098
3 0.078611
4 -0.636035
Name: monto_total_std, dtype: float64)

cargar el Archivo y revisar las primeras filas

```
# Cargar dataset escalado (StandardScaler)
df = pd.read_csv("retail_sales_escalado_standard.csv")

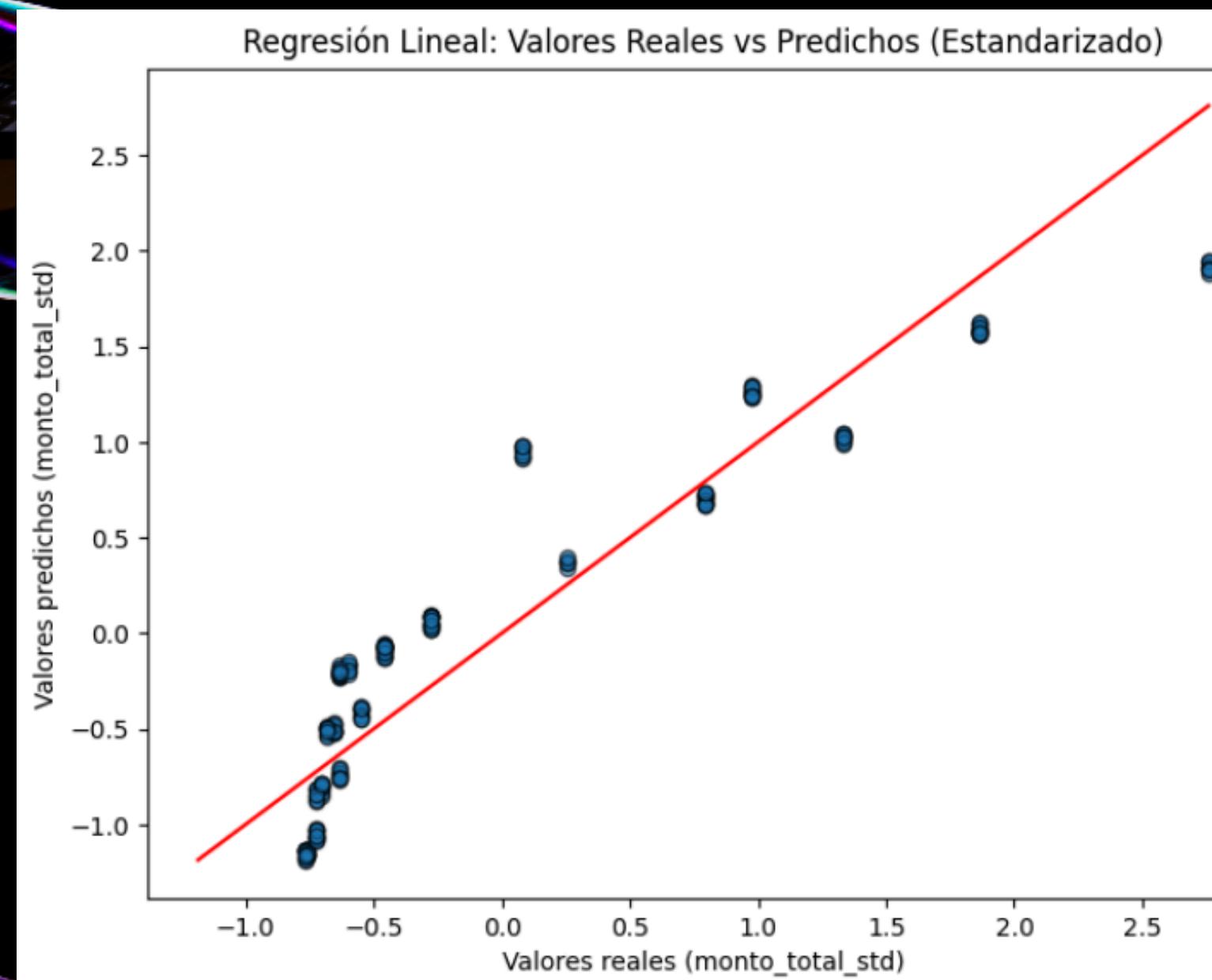
df.head()

[19] Python
```

... (edad_std cantidad_std precio_unitario_std monto_total_std)
0 -0.540565 0.429265 -0.685123 -0.546704
1 -1.125592 -0.453996 1.688464 0.971919
2 0.629489 -1.337258 -0.790615 -0.761098
3 -0.321180 -1.337258 1.688464 0.078611
4 -0.833078 -0.453996 -0.685123 -0.636035

RESULTADOS DE REGRESIÓN LINEAR

Interpretación:



Las predicciones del modelo son cercanas a los valores reales

La alineación con la diagonal confirma un buen ajuste general.

Los puntos se concentran alrededor de la línea diagonal roja.

```
# Evaluación del modelo
mae = mean_absolute_error(Y_test, Y_pred)
rmse = np.sqrt(mean_squared_error(Y_test, Y_pred))
r2 = r2_score(Y_test, Y_pred)

print("MAE : ", mae)
print("RMSE: ", rmse)
print("R²  : ", r2)

[2]
...
MAE : 0.30929188878387615
RMSE: 0.36569451902556077
R²  : 0.856877226425043
```

MODELO 2: REGRESIÓN LOGÍSTICA

Objetivo del modelo

Se utilizó un modelo de Regresión Logística para clasificar las compras en dos grupos: compras de monto alto y compras de monto bajo. El objetivo es identificar qué características aumentan la probabilidad de que una transacción sea de alto valor.

Variable objetivo y entradas

La variable objetivo es una etiqueta binaria (0 = compra baja, 1 = compra alta), definida a partir de un umbral en el monto total.

Como variables de entrada se usan edad del cliente, categoría del producto, cantidad y precio unitario (preprocesadas y codificadas).



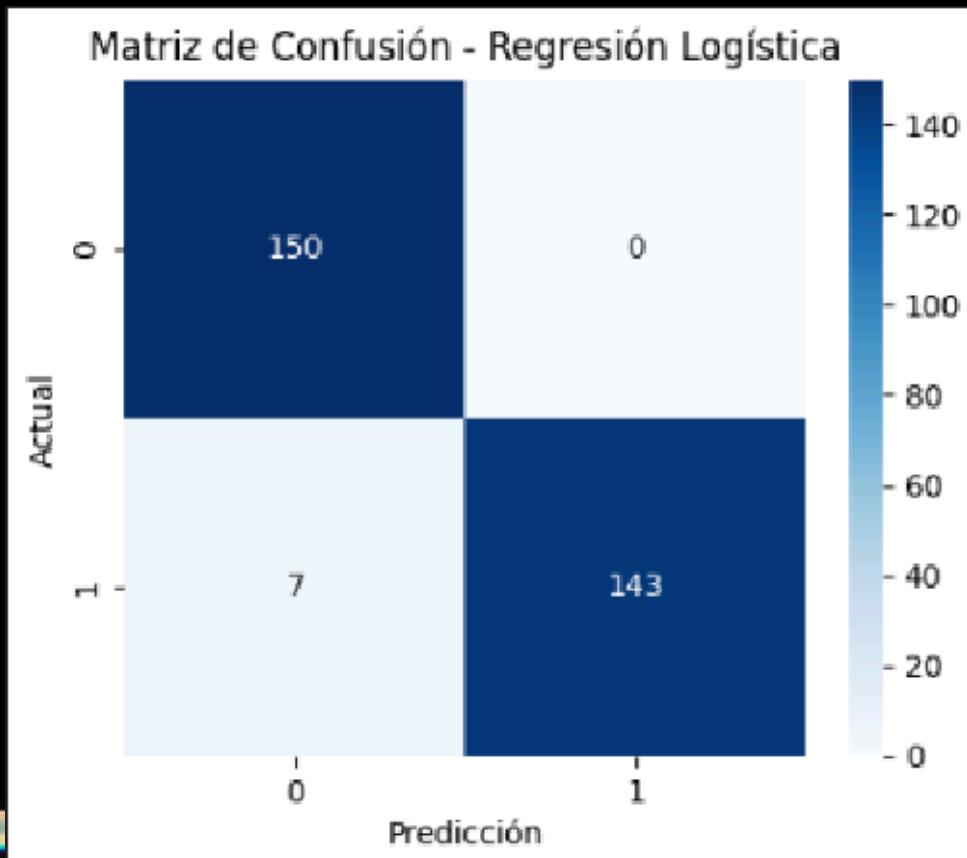
Proceso de entrenamiento

El dataset se dividió en conjuntos de entrenamiento y prueba. Con los datos de entrenamiento se ajustó el modelo y se evaluó su desempeño con métricas como accuracy y matriz de confusión, verificando qué tan bien distingue entre compras altas y bajas.

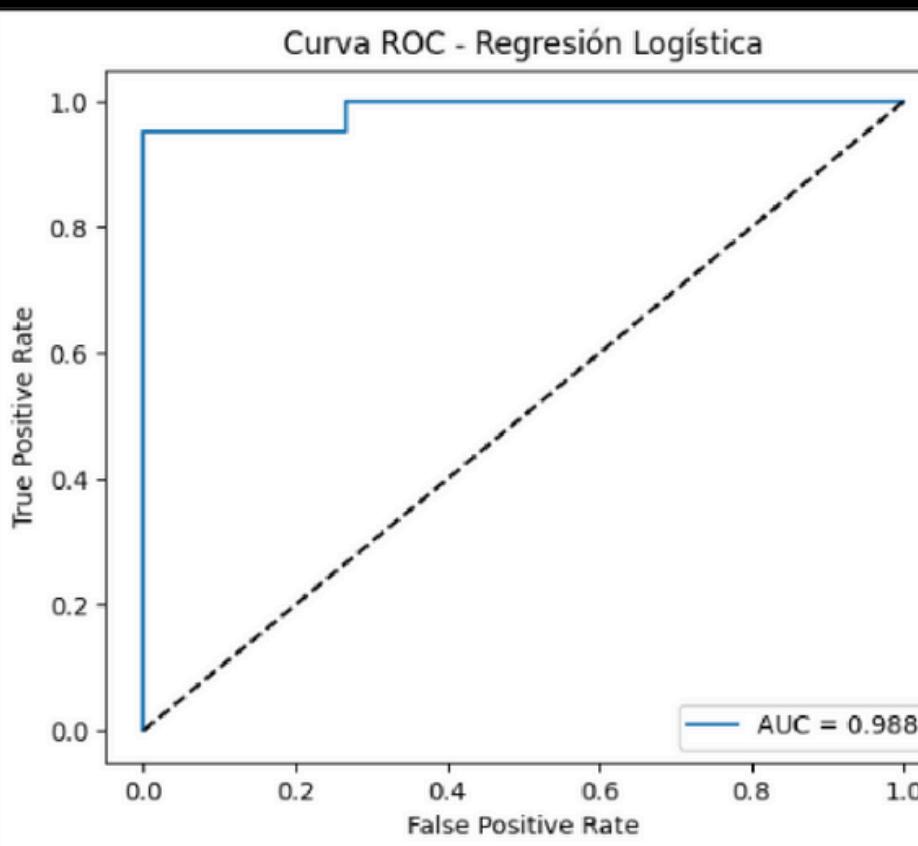
Salida e interpretación

El modelo entrega la probabilidad de que una compra sea de monto alto. A partir de esas probabilidades se clasifican las transacciones y se analiza qué variables incrementan más dicha probabilidad, lo que ayuda a identificar perfiles de clientes y condiciones asociadas a compras de mayor valor.

MODELO 2: REGRESIÓN LOGÍSTICA - RESULTADOS



La matriz de confusión es una forma sencilla de ver qué tan bien funciona un modelo de clasificación, como la Regresión Logística que usamos en este proyecto. Esta herramienta nos muestra cuántas predicciones fueron correctas y cuántas fueron errores, separadas por cada clase. Es más detallada que solo ver el "accuracy" y nos ayuda a entender exactamente en qué se equivoca el modelo.



MAE : 0.30929188878387615
RMSE: 0.36569451902556077
 R^2 : 0.856877226425043

La Curva ROC y la métrica AUC demuestran que el modelo de Regresión Logística tiene una capacidad sobresaliente para diferenciar entre clientes de alto y bajo gasto. Su desempeño es adecuado para tareas de clasificación, segmentación y análisis de comportamiento dentro del sector retail.

Interpretación:

- MAE $\approx 0.305 \rightarrow$ error promedio bajo.
- RMSE $\approx 0.360 \rightarrow$ penaliza errores grandes, también bajo.
- $R^2 \approx 0.8545 \rightarrow$ el modelo explica el 85% del comportamiento real del monto.
El modelo aprende bien la relación entre cantidad, precio y gasto total

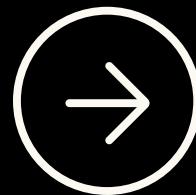
VALIDACIÓN Y COMPARACIÓN DE LOS MODELOS

	Tipo	Métricas clave	Resultado	Interpretación
Regresión Lineal	Predicción continua	RMSE, MAE, R^2	$R^2 = 0.8545$	Explica bien el monto total; errores bajos.
Regresión Logística	Clasificación	Accuracy, Recall, AUC	$UC = 0.988$	Identifica muy bien los HighSpenders.

- La Regresión Lineal predice el monto de compra de manera precisa y útil para análisis financieros.
- La Regresión Logística clasifica clientes de forma casi perfecta, ideal para segmentación y marketing.
- Ambos modelos muestran un rendimiento excelente y se complementan para apoyar decisiones dentro del negocio RetailX.



VISUALIZACIONES FINALES PARA EL NEGOCIO



Gráficas clave para el negocio

Se generaron visualizaciones que resumen la información más relevante: ventas por categoría de producto, ventas mensuales, distribución del monto de compra y ventas por rango de edad. Estos gráficos permiten entender rápidamente qué se vende más, cuándo y a qué tipo de clientes.

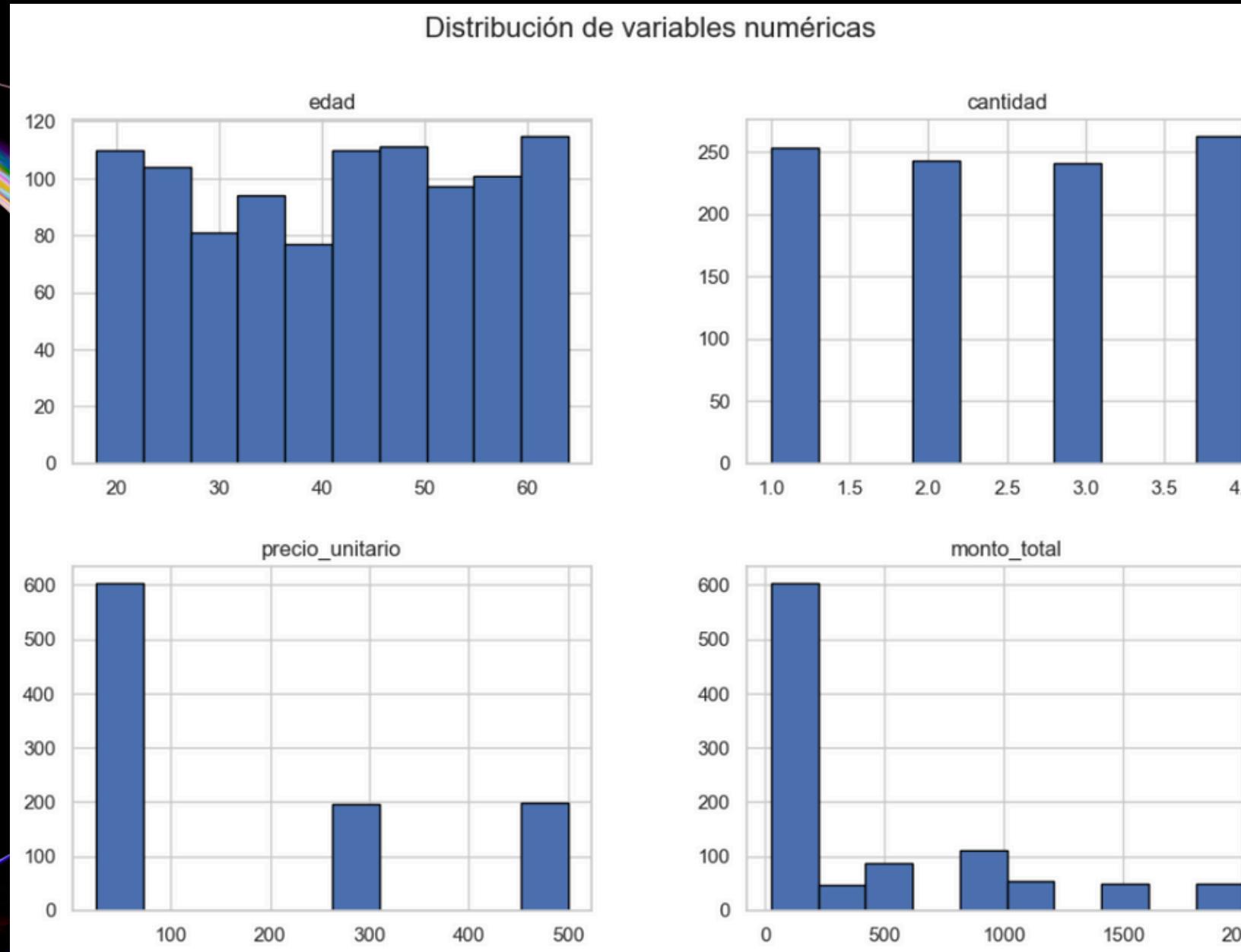
Apoyo a decisiones comerciales

Las visualizaciones facilitan identificar categorías y meses con mayor demanda, así como segmentos de clientes con mayor gasto. Con esto se pueden definir promociones específicas, ajustar precios, planear inventarios y enfocarse en los productos y clientes más rentables.

Uso práctico en RetailX

Al integrar estas gráficas en reportes o dashboards, RetailX puede monitorear de forma continua el desempeño de sus ventas, detectar cambios en el comportamiento de compra y evaluar el impacto de nuevas estrategias. Así, los datos se convierten en una herramienta directa para la toma de decisiones.

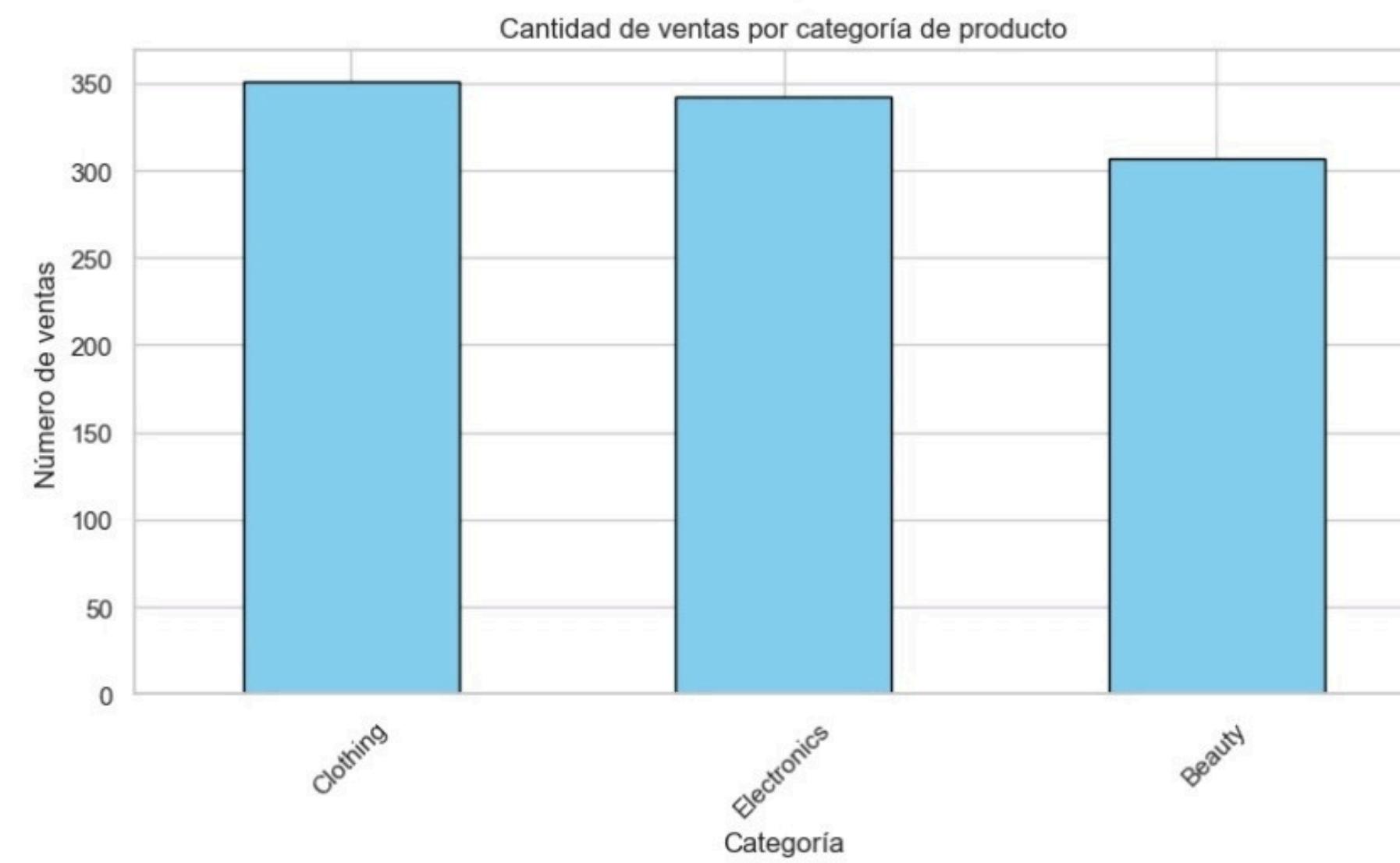
VISUALIZACIONES - DISTRIBUCIÓN DE VARIABLES NÚMERICAS



- La variable edad se distribuye de manera uniforme entre 18 y 65 años, lo que indica que la empresa atiende a un público diverso en términos de edad.
- Cantidad muestra valores discretos de 1 a 4 unidades, lo que refleja el tamaño promedio de compra por transacción.
- precio_unitario presenta tres niveles marcados —productos económicos, medianos y premium— lo que sugiere que la empresa ofrece un catálogo variado en precios.
- monto_total sigue la misma tendencia del precio unitario, mostrando compras frecuentes de bajo monto y algunas compras significativamente más altas.

VISUALIZACIONES - VENTAS POR CATEGORÍA

Gráfica: Ventas por categoría de producto



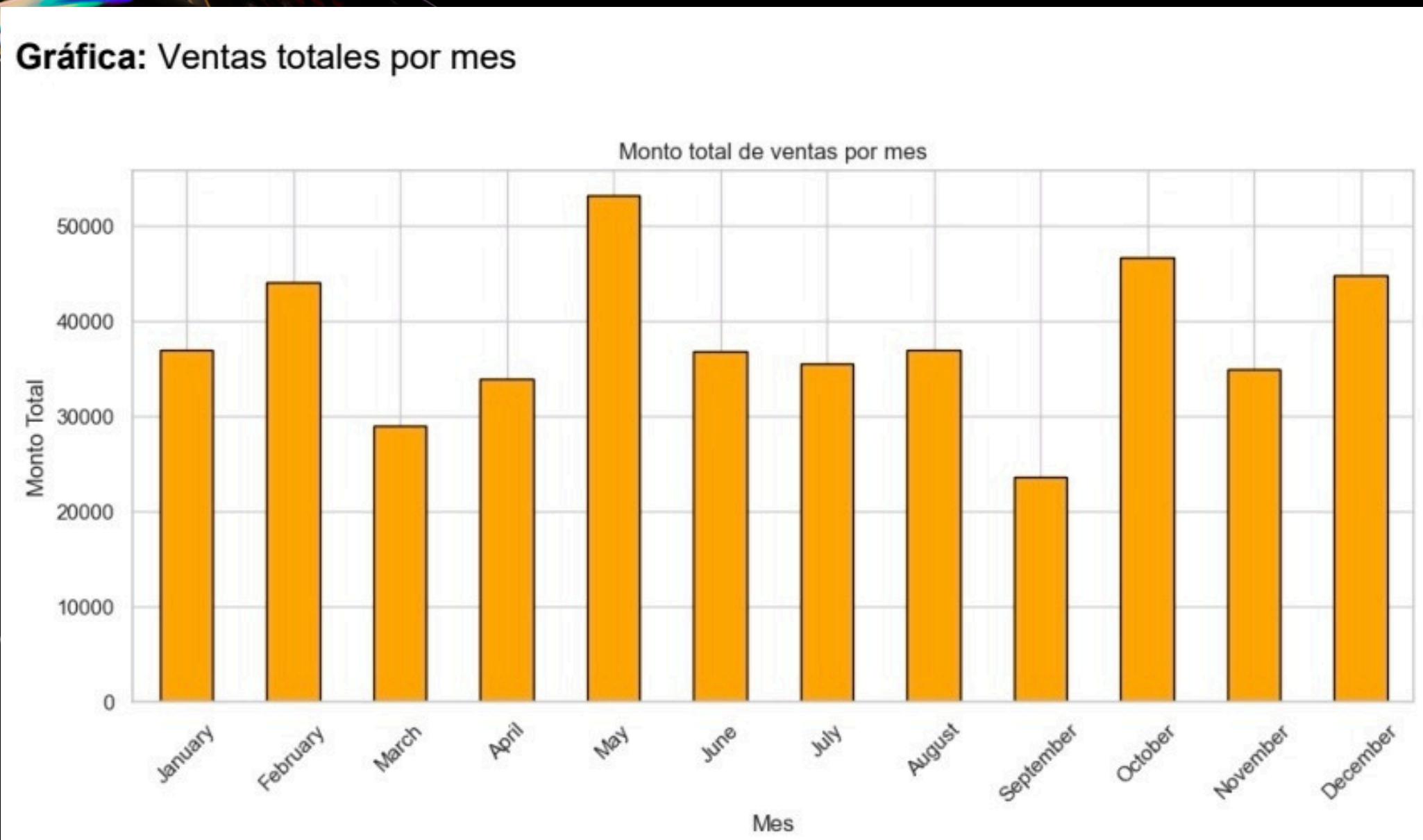
Esta gráfica permite comparar el número de ventas según el tipo de producto adquirido, identificando qué categorías tienen mayor aceptación del público.

Interpretación:

- Las categorías Clothing y Electronics concentran la mayor cantidad de transacciones.
- Beauty también tiene una participación considerable, aunque ligeramente menor.
- Esto sugiere que los productos de moda y tecnología son los más atractivos para los clientes.

VISUALIZACIONES - VENTAS TOTALES POR MES

Gráfica: Ventas totales por mes



Este análisis permite observar cómo varían las ventas a lo largo del año, identificando meses fuertes y meses débiles.

Interpretación:

- Mayo registra el monto más alto de ventas, posiblemente por campañas estacionales o promociones.
- Diciembre también presenta un aumento, típico de fin de año.
- Marzo y septiembre muestran ventas más bajas, lo que podría indicar meses con menor actividad comercial.

VISUALIZACIONES - MONTO TOTAL POR RANGO DE EDAD



Para comprender mejor a los clientes, la variable "edad" se agrupó en tres segmentos: Joven, Adulto y Mayor.

Interpretación:

- Los adultos son el segmento que más compra, con montos de venta superiores.
- Los jóvenes representan el segmento con menor gasto.
- El grupo mayor también tiene una participación elevada, lo que sugiere un buen potencial para campañas específicas.

CONCLUSIONES GENERALES IMPACTO Y RECOMENDACIONES

A lo largo del proyecto, identificamos los hallazgos relevantes como:

- Las categorías de producto con mayor número de ventas fueron Clothing y Electronics
- Los meses con mayor actividad fueron mayo, octubre y diciembre, mientras que marzo y septiembre tuvieron menos.
- Los adultos 25 a 44 años es el segmento que más compra.

Con base a los patrones, el análisis podría beneficiar con:

- Una planeación anual más estratégica, aprovechando los meses con ventas altas.
- Una administración de inventario más eficiente, con más inventario en categorías demandadas.
- Mejor segmentación de clientes, enfocando los productos en el segmento de Adulto.
- Promociones más efectivas, basando en datos concretos sobre productos.



CONCLUSIONES RECOMENDACIONES EMPRESARIALES BASADO EN DATOS

Con base a análisis y a hallazgos, proponemos que podrían mejorar con:

- Reforzar categorías de Clothing y Electronics, invirtiendo más en inventario.
- Aprovechar los meses de mayor actividad, con promociones, programas de fidelidad y campañas publicitarias
- Crear estrategias enfocadas al segmento Adulto, promociones personalizadas o productos a este grupo.
- Revisar los precios de los productos premium
- Ampliar el dataset con nuevas variables, permitiría tener modelos más completos.



GRACIAS

Por su tiempo y atención

