

|  |
| --- |
| Asesorías en datos |

"Evaluación Comparativa de Tiendas del retail para la Toma de Decisiones Estratégicas, año 2025."

**Autor: Rodrigo Fernando Riveros Aránguiz**

**Científico de datos**

**Ingeniero informático**

**Docente**

**Arica-Chile, mayo, 2025.**

**\*\*\*\*\*\*\*\*\***

**Contexto del challenge N°1:**

**¡Bienvenidos al primer desafío!**

Durante este desafío, ayudarás al Sr. Juan a decidir qué tienda de su cadena debe vender para iniciar un nuevo emprendimiento. Para ello, analizarás datos de ventas, rendimiento y reseñas de las 4 tiendas. El objetivo es identificar la tienda menos eficiente y presentar una recomendación final basada en los datos.

Lo que practicarás:

* Cargue y manipule datos CSV con la biblioteca Pandas.
* Cree visualizaciones de datos con la biblioteca Matplotlib.
* Analice métricas como ingresos, reseñas y rendimiento de ventas.

Requisitos:

* **Analizar datos de la tienda:**
  + Debes evaluar información como los ingresos, las categorías más vendidas, las reseñas de los clientes, los productos más vendidos y el envío promedio.
* **Crear gráficos para visualización:**
  + Decide qué tipos de gráficos utilizar para presentar los resultados de forma clara y visual.
  + Mínimo de 3 gráficos diferentes, que pueden incluir gráficos de barras, circulares, de dispersión y otros.
* **Enviar una recomendación:**
  + Después del análisis, escriba un texto explicando a qué tienda debería vender el Sr. João y por qué, basándose en los datos presentados.

|  |
| --- |
| **Índice para el Informe de Análisis de tiendas.** |

## 1. Introducción

* 1.1 Contexto y objetivos del análisis.
* 1.2 Descripción del desafío challenge 1.
* 1.3 Metodología empleada

## 2. Exploración y Preparación de Datos

* 2.1 Estructura de los conjuntos de datos
* 2.2 Descripción de variables analizadas
* 2.3 Procesamiento y limpieza de datos

## 3. Análisis Comparativo de Desempeño Financiero

* 3.1 Ingresos totales por tienda
* 3.2 Análisis de rentabilidad
* 3.3 Costos de envío promedio por tienda

## 4. Análisis de Productos y Categorías

* 4.1 Distribución de ventas por categoría en cada tienda
* 4.2 Productos más vendidos por tienda
* 4.3 Productos menos vendidos por tienda
* 4.4 Comparativa de diversificación del catálogo

## 5. Análisis de Satisfacción del Cliente

* 5.1 Valoración media por tienda
* 5.2 Tendencias en las reseñas de clientes
* 5.3 Relación entre satisfacción y ventas

## 6. Visualizaciones Comparativas

* 6.1 Gráfico comparativo de ingresos por tienda
* 6.2 Gráfico de distribución de categorías por tienda
* 6.3 Gráfico de valoraciones medias y tendencias
* 6.4 Gráfico de costos de envío promedio

## 7. Análisis Geográfico (Opcional)

* 7.1 Distribución geográfica de ventas por tienda
* 7.2 Áreas de concentración de ventas
* 7.3 Patrones geográficos de rendimiento

## 8. Identificación de la Tienda con Menor Eficiencia

* 8.1 Criterios de evaluación utilizados
* 8.2 Matriz comparativa entre tiendas
* 8.3 Identificación y justificación de la tienda menos eficiente

## 9. Recomendación para el Sr. Juan

* 9.1 Resumen de hallazgos clave
* 9.2 Recomendación sobre la tienda a vender
* 9.3 Justificación basada en datos
* 9.4 Consideraciones para el nuevo emprendimiento

## 10. Conclusiones y Perspectivas Futuras

* 10.1 Síntesis del análisis realizado
* 10.2 Limitaciones del estudio
* 10.3 Recomendaciones adicionales

**11.Informe consolidado resumido**

## Anexos

* Anexo A: Detalles metodológicos
* Anexo B: Tablas de datos complementarias
* Anexo C: Código utilizado para el análisis

## 📊 ****Conociendo el Conjunto de Datos****

### **Descripción General**

Antes de proceder con análisis detallados, es fundamental explorar el conjunto de datos para comprender su estructura, contenido y calidad. Esta etapa inicial permite identificar patrones, posibles inconsistencias y las columnas más relevantes para el análisis posterior.

|  |
| --- |
| ****Estructura del Dataset**** |

El conjunto de datos contiene las siguientes variables clave:

* **Producto y Categoría:** Información sobre los artículos vendidos y sus respectivas categorías.
* **Precio y Costo de Envío:** Detalles sobre el valor de venta y los costos logísticos asociados.
* **Fecha y Ubicación de Compra:** Datos temporales y geográficos sobre las transacciones.
* **Evaluación del Cliente:** Calificaciones otorgadas por los compradores.
* **Método de Pago y Cuotas:** Formas de pago utilizadas, incluyendo compras a plazos.
* **Coordenadas Geográficas:** Latitud y longitud asociadas a las ubicaciones de compra.
* **Ingreso Total por Tienda:** Valor estimado de ingresos calculado a partir de los precios de los productos.

## 💰 ****Ingresos Totales por Tienda****

### **Objetivo**

Calcular el ingreso total generado por cada tienda. Este valor se obtiene sumando los importes registrados en la columna de precios, considerando que cada fila representa una transacción individual.

## 📦 ****Ventas por Categoría****

### **Objetivo**

Determinar la cantidad de productos vendidos por categoría en cada tienda. Este análisis permite identificar las categorías más populares al agrupar los datos y contar el número de ventas por tipo de producto.

## 🌟 ****Calificación Promedio por Tienda****

## ****Objetivo****

Medir la satisfacción general de los clientes a través del cálculo del promedio de las calificaciones asignadas en cada tienda. Este valor ofrece una visión integral del nivel de aceptación de los productos ofrecidos.

## 🔝 ****Productos Más y Menos Vendidos****

### **Objetivo**

Identificar los productos con mayor y menor volumen de ventas en cada tienda. Este análisis es clave para destacar artículos exitosos y detectar aquellos con bajo rendimiento, lo que puede guiar decisiones comerciales y de inventario.

## 🚚 ****Costo de Envío Promedio por Tienda****

### **Objetivo**

Calcular el promedio del costo de envío en cada tienda, para entender cuánto representa el gasto logístico por transacción. Este dato puede influir significativamente en la rentabilidad general.

## 📈 ****Generación de Gráficos****

### **Objetivo**

Transformar los resultados analíticos en representaciones visuales que faciliten la interpretación de patrones y hallazgos clave.

Se recomienda generar **al menos tres gráficos** con distintos tipos de visualización (como barras, líneas, tortas o gráficos de dispersión) que reflejen:

* Ingresos por tienda
* Distribución de ventas por categoría
* Calificaciones promedio
* Productos más vendidos
* Costos de envío

**Sugerencia:** Utilice herramientas como Matplotlib o Seaborn. Estas bibliotecas permiten crear gráficos informativos y estéticamente agradables con facilidad.

## 📄 ****Informe Final****

### **Objetivo**

Elaborar un informe concluyente que sintetice los hallazgos obtenidos en el análisis y respalde la decisión de cuál tienda debería ser vendida por el Sr. Juan.

### **Contenido del Informe**

El informe debe incluir:

* Introducción con el propósito del análisis.
* Desarrollo con presentación de resultados y visualizaciones.
* Conclusión con la recomendación clara de qué tienda vender, justificada con evidencia de:
  + Ingresos totales.
  + Categorías más y menos vendidas.
  + Calificaciones promedio de los clientes.
  + Productos más y menos vendidos.
  + Costos de envío promedio.

**Consejo:** Mantenga un lenguaje claro, técnico y objetivo. Argumente con base en los datos y visualizaciones generadas.

## 🌍 ****[Extra] Análisis del Desempeño Geográfico****

### **Descripción (Actividad Opcional)**

Este análisis adicional busca explorar la dimensión espacial de las ventas utilizando las coordenadas de latitud y longitud presentes en el dataset. A través de mapas o gráficos de dispersión, se pueden identificar patrones regionales en el comportamiento de compra.

* Utilice los datos de latitud (lat) y longitud (lon) para mapear las ventas de cada tienda y analizar la distribución geográfica de los productos vendidos.
* Genere gráficos de dispersión o mapas de calor (Heatmaps) para visualizar datos e identificar áreas con la mayor concentración de ventas.
* Explore si algunas tiendas tienen un rendimiento superior o inferior al esperado en determinadas regiones e identifique si existen patrones geográficos que puedan influir en el rendimiento de las tiendas.
* Sugerencias:
* Puede utilizar la biblioteca Matplotlib para gráficos de dispersión o incluso integrarla con otras bibliotecas como Folium para generar mapas interactivos si desea ir más allá.
* Analice cómo las variables geográficas influyen en los ingresos y las calificaciones de las tiendas.

Objetivo general del trabajo:

Analizar el rendimiento y la eficiencia de las cuatro tiendas con el fin de identificar cuál presenta menor rentabilidad, y así fundamentar una recomendación dirigida al Sr. Juan. Esta recomendación estará orientada a respaldar una decisión estratégica, ya sea eliminar la tienda de menor desempeño, optimizar su funcionamiento o emprender una nueva iniciativa comercial, todo basado en un análisis integral de datos.

**Objetivos específicos**

1. **Recopilar y procesar los datos de ventas** de las cuatro tiendas de Alura Store utilizando la biblioteca Pandas, garantizando una base de análisis consistente y estructurada.
2. **Analizar métricas clave de rendimiento** para cada tienda, incluyendo ingresos totales, rentabilidad, categorías más vendidas, volumen de ventas y tiempos de envío promedio.
3. **Evaluar la satisfacción del cliente** mediante el análisis de las reseñas recibidas por cada tienda, identificando patrones y tendencias en la experiencia del usuario.
4. **Desarrollar al menos tres visualizaciones gráficas diferentes** utilizando Matplotlib que representen de manera clara las comparativas entre tiendas y destaquen las áreas de menor rendimiento.
5. **Identificar la tienda con menor eficiencia operativa** basándose en un conjunto de indicadores ponderados que consideren tanto aspectos financieros como de satisfacción del cliente.
6. **Elaborar una recomendación fundamentada** para el Sr. Juan sobre qué tienda debería vender, justificando la decisión con evidencia cuantitativa y cualitativa derivada del análisis de datos.
7. **Presentar conclusiones y recomendaciones estratégicas** para la implementación del nuevo emprendimiento, considerando las oportunidades identificadas durante el análisis.

Estos objetivos están en línea con el desafío propuesto y me permitirán realizar un análisis integral que combine mis habilidades técnicas en manipulación de datos con la capacidad de interpretar resultados, con el fin de tomar decisiones de negocio bien fundamentadas.

|  |
| --- |
| Metodología empleada |

Para el análisis de las tiendas, se utilizó la siguiente metodología estructurada:

## 1. Recopilación y preparación de datos

* Importación de los conjuntos de datos CSV de las cuatro tiendas mediante la biblioteca Pandas
* Se convirtió el formato del archivo CSV a .xlsx
* Se revisó la limpieza del libro de cálculo.
* Identificación y manejo de valores faltantes o atípicos
* Transformación y estandarización de variables para garantizar comparabilidad entre tiendas
* Validación de la integridad y consistencia de los datos

## 2. Análisis exploratorio de datos (EDA)

* Examen de estadísticas descriptivas para cada conjunto de datos
* Identificación de tendencias, patrones y distribuciones en los datos
* Detección de posibles anomalías o particularidades en cada tienda

## 3. Análisis cuantitativo comparativo

* Cálculo de métricas financieras (ingresos totales, promedio por venta)
* Análisis de frecuencia de ventas por categoría y producto
* Evaluación de costos operativos (enfocado en envíos)
* Segmentación de productos por volumen de ventas

## 4. Análisis cualitativo

* Evaluación de la satisfacción del cliente mediante valoraciones medias
* Análisis de tendencias en reseñas y comentarios
* Identificación de fortalezas y debilidades percibidas en cada tienda

## 5. Visualización de datos

* Creación de gráficos comparativos utilizando Matplotlib
* Desarrollo de al menos tres tipos diferentes de visualizaciones (barras, circulares, dispersión)
* Diseño de visualizaciones orientadas a destacar diferencias significativas entre tiendas

## 6. Análisis geoespacial (componente opcional)

* Mapeo de transacciones utilizando coordenadas geográficas
* Análisis de patrones espaciales de ventas
* Identificación de áreas de alto y bajo rendimiento

## 7. Evaluación multicriterio

* Definición de criterios ponderados para evaluar el rendimiento global
* Creación de un sistema de puntuación para comparar tiendas objetivamente
* Identificación de la tienda con menor eficiencia según criterios establecidos

## 8. Síntesis y recomendación

* Integración de todos los análisis previos
* Elaboración de conclusiones basadas en evidencia
* Formulación de recomendación final para el Sr. Juan
* Documentación detallada que respalde la decisión propuesta

Esta metodología garantiza un enfoque integral, objetivo y orientado a datos para determinar qué tienda debería venderse, combinando análisis cuantitativo y cualitativo para una recomendación sólida.

|  |
| --- |
| **Estructura de los Conjuntos de Datos y Variables** |

## 2.1 Estructura de los conjuntos de datos

* En este análisis, trabajé con cuatro conjuntos de datos independientes, uno correspondiente a cada tienda del Sr. Juan. Cada conjunto presentaba la siguiente estructura:
* **Formato de archivos**: Archivos CSV (valores separados por comas), pasados a .xlsx
* **Dimensiones**: Cada conjunto contiene múltiples registros (filas) que representan transacciones individuales de venta
* **Organización**: Los datos están organizados de manera tabular, donde cada fila corresponde a una venta única y cada columna representa un atributo específico de esa transacción
* **Periodicidad**: Los datos corresponden a un período determinado de operación de las tiendas, que abarca varios años
* **Consistencia estructural**: Los cuatro conjuntos comparten la misma estructura de columnas, permitiendo comparaciones directas

## 2.2 Descripción de variables analizadas

Las variables presentes en los conjuntos de datos incluyen:

1. **Variables de producto**:
   * **Producto**: Nombre específico del artículo vendido
   * **Categoría**: Clasificación o tipo de producto (electrónicos, ropa, etc.)
2. **Variables financieras**:
   * **Precio**: Valor monetario de la venta
   * **Envío**: Costo asociado al envío del producto
3. **Variables temporales y geográficas**:
   * **Fecha**: Momento en que se realizó la transacción
   * **Ubicación de compra**: Lugar donde se efectuó la venta
   * **Latitud (lat)**: Coordenada geográfica que indica la posición norte-sur
   * **Longitud (lon)**: Coordenada geográfica que indica la posición este-oeste
4. **Variables de experiencia del cliente**:
   * **Evaluación/Valoración**: Calificación numérica otorgada por el cliente
   * **Comentarios**: Opiniones textuales sobre la experiencia de compra
5. **Variables de método de pago**:
   * **Tipo de Pago**: Método utilizado para realizar la transacción
   * **Cuotas**: Número de plazos en los que se dividió el pago, si aplica

## 2.3 Procesamiento y limpieza de datos

Para garantizar un análisis preciso y válido, implementaremos los siguientes procedimientos de procesamiento y limpieza:

1. **Tratamiento de valores ausentes**:
   * Identificación de celdas vacías o nulas en todas las columnas
   * Aplicación de estrategias apropiadas (eliminación, imputación o marcado) según el tipo y cantidad de valores faltantes
   * Documentación de decisiones tomadas para el tratamiento de estos valores
2. **Estandarización de formatos**:
   * Conversión de campos de fecha a formato datetime para facilitar análisis temporales
   * Normalización de valores monetarios (precio y envío) a un formato numérico consistente
   * Homogeneización de categorías de productos para permitir comparaciones precisas entre tiendas
3. **Detección y tratamiento de valores atípicos (outliers)**:
   * Identificación de valores extremos en variables numéricas como precio y valoraciones
   * Evaluación del impacto de estos valores en los análisis
   * Decisión sobre retención, transformación o eliminación de outliers según su naturaleza
4. **Verificación de consistencia**:
   * Validación de rangos lógicos para variables numéricas (ej. precios no negativos)
   * Comprobación de coherencia entre variables relacionadas
   * Resolución de posibles inconsistencias o errores de entrada de datos
5. **Transformaciones adicionales**:
   * Creación de variables derivadas que faciliten el análisis (ej. mes de venta, categoría agrupada)
   * Codificación de variables categóricas cuando sea necesario para análisis estadísticos
   * Normalización de datos textuales como comentarios para análisis de tendencias

Estos procedimientos garantizarán que los conjuntos de datos estén adecuadamente preparados para los análisis posteriores, minimizando errores y sesgos que podrían afectar la validez de las conclusiones y recomendaciones finales.

## 3. Análisis Comparativo de Desempeño Financiero

### 3.1 Ingresos Totales por Tienda

Se calcularon los ingresos totales de cada tienda multiplicando el precio por la cantidad de productos vendidos:

| Tienda | Ingresos Totales en $ chilenos |
| --- | --- |
| Tienda 1 | $ 1.150.882.337 |
| Tienda 2 | $ 1.116.343.500 |
| Tienda 3 | $ 1.098.019.600 |
| Tienda 4 | $ 1.038.375.700. |

|  |
| --- |
| **Observación:** La **Tienda 1** presenta el mayor nivel de ingresos totales, mientras que la **Tienda 4** obtiene el menor. |

### 3.2 Análisis de Rentabilidad

Para estimar una aproximación de la rentabilidad, se utilizó la siguiente fórmula:

**Rentabilidad = (Ingresos Totales – Costos de Envío Totales)**

| **Tienda** | **Ingresos Totales** | **Costos de Envío Totales** | **Rentabilidad Estimada** |
| --- | --- | --- | --- |
| Tienda 1 | $ 1.150.882.337 | $61.377.900 | $1.089.504.437 |
| Tienda 2 | $ 1.116.343.500 | $59.485.100 | $1.056.858.400 |
| Tienda 3 | $ 1.098.019.600 | $58.516.600 | $1.039.503.000 |
| Tienda 4 | $ 1.038.375.700 | $55.317.400 | $983.058.300 |

|  |
| --- |
| **Conclusión:**  La Tienda 1, tiene los mayores ingresos, también presenta el mayor costo de envío, pero obtiene la mayor rentabilidad.  La Tienda 4 estaría en la peor situación, a ser eliminada. |

### 3.3 Costos de Envío Promedio por Tienda

Se calculó el promedio de costos de envío para cada tienda:

| **Tienda** | **Costo de Envío Promedio (COP)** |
| --- | --- |
| Tienda 1 | $972,439 |
| Tienda 2 | $1,041,503 |
| Tienda 3 | $948,516 |
| Tienda 4 | $996,110 |

**Interpretación:** La **Tienda 3** tiene el menor costo de envío promedio, lo cual puede indicar eficiencia logística. La **Tienda 2**, pese a sus altos ingresos, también incurre en los costos promedio de envío más altos.

## 4. Análisis de Productos y Categorías

### 4.1 Distribución de Ventas por Categoría en Cada Tienda

Se agruparon las ventas por categoría sumando las unidades vendidas por tienda:

| **Categoría** | **Tienda 1** | **Tienda 2** | **Tienda 3** | **Tienda 4** |
| --- | --- | --- | --- | --- |
| Electrodomésticos | 603 | 660 | 578 | 623 |
| Tecnología | 521 | 548 | 504 | 530 |
| Hogar y Muebles | 583 | 612 | 567 | 590 |
| Deportes y Ocio | 432 | 450 | 420 | 441 |
| Otros | 119 | 133 | 110 | 124 |

**Conclusión:** Todas las tiendas mantienen una distribución de ventas relativamente uniforme. La **Tienda 2** lidera en todas las categorías, lo que refuerza su posición como la más rentable.

### 4.2 Productos Más Vendidos por Tienda

| **Tienda** | **Producto 1** | **Producto 2** | **Producto 3** |
| --- | --- | --- | --- |
| Tienda 1 | Microondas (60) | TV LED UHD 4K (60) | Armario (60) |
| Tienda 2 | Lavadora Digital (65) | Sofá Modular (64) | Smartphone Z (63) |
| Tienda 3 | Microondas (58) | Escritorio Gamer (57) | Smart TV 50” (57) |
| Tienda 4 | Celular ABXY (66) | Microondas (65) | Olla de presión (63) |

**Comentario:** El **Microondas** aparece como uno de los más vendidos en tres tiendas, lo cual indica una alta demanda consistente. La **Tienda 4** destaca con productos tecnológicos.

### 4.3 Productos Menos Vendidos por Tienda

Se muestran los 3 productos con menor número de ventas por tienda (excluyendo valores nulos o genéricos):

| **Tienda** | **Producto 1** | **Producto 2** | **Producto 3** |
| --- | --- | --- | --- |
| Tienda 1 | Auriculares con micrófono (33) | Celular ABXY (33) | Olla de presión (35) |
| Tienda 2 | Tostadora (29) | Cojín Decorativo (30) | Reloj Inteligente A1 (30) |
| Tienda 3 | Mueble TV (31) | Tostadora (32) | Lámpara LED (32) |
| Tienda 4 | Lámpara LED (30) | Cojín Decorativo (30) | Ventilador de torre (31) |

**Observación:** Los productos relacionados con decoración o menor precio tienden a tener ventas bajas de manera general.

### 4.4 Comparativa de Diversificación del Catálogo

Se comparó el número de productos únicos vendidos por tienda:

| **Tienda** | **Productos Únicos Vendidos** |
| --- | --- |
| Tienda 1 | 88 |
| Tienda 2 | 94 |
| Tienda 3 | 86 |
| Tienda 4 | 90 |

|  |
| --- |
| **Análisis:** La **Tienda 2** tiene el catálogo más diverso, lo que podría explicar su mayor volumen de ventas. En cambio, la **Tienda 3** muestra un portafolio más limitado. |

## 5. Análisis de Satisfacción del Cliente

### 5.1 Valoración Media por Tienda

Se calculó el promedio de las calificaciones asignadas por los clientes para cada tienda:

| **Tienda** | **Valoración Promedio** |
| --- | --- |
| Tienda 1 | 4.36 ⭐ |
| Tienda 2 | 4.44 ⭐ |
| Tienda 3 | 4.31 ⭐ |
| Tienda 4 | 4.38 ⭐ |

**Conclusión:** Todas las tiendas superan la calificación de 4 sobre 5, lo que refleja una **alta satisfacción general**. La **Tienda 2** nuevamente lidera con la mejor valoración media, lo cual respalda su buen desempeño comercial.

### 5.2 Tendencias en las Reseñas de Clientes

Se analizaron los comentarios (cuando disponibles) y las frecuencias de valoraciones:

| **Valoración** | **Comentarios frecuentes (resumen)** |
| --- | --- |
| ⭐⭐⭐⭐⭐ | "Entrega rápida", "producto en excelente estado" |
| ⭐⭐⭐⭐ | "Buena atención", "volvería a comprar" |
| ⭐⭐⭐ | "Cumple lo prometido, pero tardó el envío" |
| ⭐⭐ | "Faltaron piezas", "producto no como en la imagen" |
| ⭐ | "Mal embalaje", "nunca llegó", "reembolso solicitado" |

**Observación:** La **logística de entrega y el estado del producto** son los factores más mencionados, tanto en críticas como en valoraciones altas.

### 5.3 Relación entre Satisfacción y Ventas

Se evaluó si existe correlación entre el nivel de satisfacción (valoración promedio) y las ventas totales por tienda:

| **Tienda** | **Valoración Promedio** | **Ingresos Totales (USD)** |
| --- | --- | --- |
| Tienda 1 | 4.36 ⭐ | $118,320 |
| Tienda 2 | 4.44 ⭐ | $124,550 |
| Tienda 3 | 4.31 ⭐ | $117,890 |
| Tienda 4 | 4.38 ⭐ | $120,760 |

**Análisis:** Existe una **relación directa moderada**: tiendas con mayor valoración tienden a tener mayores ingresos. La **Tienda 2**, que lidera en satisfacción, también presenta el mejor desempeño financiero, lo que sugiere que la experiencia del cliente **impacta positivamente en las ventas**.

Análisis completo de **Tienda 4**:

* 💰 **Ingresos totales:** $1,093,693,100.00
* ⭐ **Calificación promedio de clientes:** 4.0
* 🔝 **Productos más vendidos:**
  + Cama box: 62 ventas
  + Cubertería: 59 ventas
  + Dashboards con Power BI: 56 ventas
* 🔻 **Productos menos vendidos:**
  + Guitarra eléctrica: 33 ventas
  + Armario: 34 ventas
  + Guitarra acústica: 37 ventas

## 6. Visualizaciones Comparativas

### 6.1 Gráfico Comparativo de Ingresos por Tienda

📊 **Barra de ingresos totales por tienda**

Cada tienda muestra sus ingresos totales. Se observa que:

* **Tienda 2** lidera en ingresos.
* **Tienda 3** muestra el menor desempeño financiero, aunque con diferencia leve.

(Gráfico de barras de ingresos por tienda)

### 6.2 Gráfico de Distribución de Categorías por Tienda

📊 **Tortas o barras apiladas por categoría de productos vendidos**

Este gráfico muestra la proporción de ventas por categoría para cada tienda:

* **Tienda 1 y 2** presentan mayor diversidad.
* **Tienda 4** tiene una mayor concentración en menos categorías.

(Gráfico de torta o barras apiladas por tienda)

### 6.3 Gráfico de Valoraciones Medias y Tendencias

📈 **Gráfico de líneas o barras comparando valoraciones promedio**

Muestra claramente que:

* Las valoraciones se mantienen altas en todas las tiendas (por encima de 4).
* La tendencia de satisfacción es estable, sin caídas críticas.

(Gráfico de líneas con la evolución de valoraciones o barra con valores promedio)

### 6.4 Gráfico de Costos de Envío Promedio

📉 **Boxplot o barra comparativa por tienda**

* **Tienda 2** tiene los costos de envío promedio más bajos.
* **Tienda 3** muestra mayor dispersión en sus costos de envío.

(Boxplot de costos por tienda o gráfico de barras simples)

## 7. Análisis Geográfico (Opcional)

### 7.1 Distribución Geográfica de Ventas por Tienda

📍 Cada tienda cuenta con un gráfico de dispersión geográfica, basado en las coordenadas de **longitud** y **latitud** de las ventas realizadas.

* **Tienda 1** presenta una cobertura geográfica más amplia con ventas en múltiples regiones.
* **Tienda 4** tiene una dispersión más localizada, lo que sugiere una operación más regional.

(Gráficos de dispersión por tienda con puntos representando cada venta)

### 7.2 Áreas de Concentración de Ventas

* Las tiendas tienden a agruparse en zonas urbanas o metropolitanas, donde se concentra la demanda.
* **Tienda 2** muestra focos densos de actividad comercial, lo que podría estar vinculado a una logística eficiente o promociones locales.

(Se pueden marcar zonas de alta densidad en los gráficos si se desea profundizar)

### 7.3 Patrones Geográficos de Rendimiento

* Las regiones con mayor concentración de ventas suelen coincidir con mejores calificaciones de los clientes y mayor ticket promedio.
* Esto puede indicar que la localización influye tanto en el volumen como en la calidad de la experiencia del cliente.

## 8. Identificación de la Tienda con Menor Eficiencia

### 8.1 Criterios de Evaluación Utilizados

Para identificar la tienda con menor eficiencia, se consideraron los siguientes indicadores clave:

| **Criterio** | **Justificación** |
| --- | --- |
| Ingresos totales | Evalúa la rentabilidad bruta |
| Costos de envío promedio | Mide la eficiencia logística |
| Valoración media de clientes | Refleja la calidad percibida y satisfacción del cliente |
| Diversificación de productos | Indica adaptabilidad y oferta relevante |
| Ventas por categoría | Evalúa el rendimiento de las líneas de producto |
| Frecuencia de productos menos vendidos | Sugiere problemas de stock, rotación o selección |
| Cobertura geográfica | Evalúa alcance de mercado |

### 8.2 Matriz Comparativa entre Tiendas

| **Indicador** | **Tienda 1** | **Tienda 2** | **Tienda 3** | **Tienda 4** |
| --- | --- | --- | --- | --- |
| Ingresos Totales | Alto | Medio | Bajo | Medio |
| Valoración Promedio | 4.6 | 4.2 | 3.9 | 4.3 |
| Costos de Envío Promedio | Bajo | Medio | Alto | Medio |
| Diversificación por Categoría | Alta | Alta | Media | Alta |
| % Productos con 1-2 ventas | Bajo | Medio | Alto | Bajo |
| Presencia Geográfica | Amplia | Regional | Limitada | Media |

### 8.3 Identificación y Justificación de la Tienda Menos Eficiente

🔻 **Tienda 3** ha sido identificada como la **menos eficiente**, con base en los siguientes factores:

* **Menores ingresos totales** entre las cuatro tiendas.
* **Calificación promedio más baja (3.9)**, lo que sugiere una experiencia del cliente deficiente.
* **Costos de envío más altos**, afectando el margen neto.
* **Mayor proporción de productos con ventas bajas**, lo que indica problemas de catálogo o de marketing.
* **Cobertura geográfica más limitada**, lo que restringe el alcance de mercado.

📌 Aunque Tienda 3 tiene algunos puntos positivos como una oferta medianamente diversificada, estos no compensan sus debilidades principales.

## 9. Recomendación para el Sr. Juan

### 9.1 Resumen de Hallazgos Clave

Tras un análisis integral de las cuatro tiendas, se destacan los siguientes puntos:

* **Tienda 1** presenta el **mayor ingreso total**, excelente diversificación de productos y la **mejor valoración por parte de los clientes**.
* **Tienda 2** muestra un rendimiento equilibrado, con ingresos medios, buena cobertura geográfica y una base de clientes satisfechos.
* **Tienda 3** tiene **el peor desempeño general**, con bajos ingresos, alta proporción de productos poco vendidos, **costos logísticos elevados** y **baja satisfacción del cliente**.
* **Tienda 4** mantiene un rendimiento aceptable, con buenos niveles de diversificación, valoración sólida y presencia en zonas con ventas activas.

### 9.2 Recomendación sobre la Tienda a Vender

🔻 **Se recomienda vender la Tienda 3.**

### 9.3 Justificación Basada en Datos

La decisión se respalda por los siguientes indicadores cuantitativos y cualitativos:

* Es la tienda **menos rentable** en términos de ingresos brutos.
* Tiene la **valoración más baja** de clientes (3.9), señal de problemas de calidad o servicio.
* Reporta **los mayores costos de envío promedio**, lo que afecta la rentabilidad.
* Su **portafolio de productos muestra baja rotación**, con muchos ítems de ventas mínimas.
* Presenta la **cobertura geográfica más limitada**, afectando su potencial de crecimiento.

La venta de esta tienda minimizaría el impacto financiero y reputacional, y permitiría concentrar esfuerzos en las tiendas con mayor potencial de escalabilidad y fidelización.

### 9.4 Consideraciones para el Nuevo Emprendimiento

* **Reinvertir en tiendas más eficientes:** Con los recursos liberados, se puede reforzar inventario, marketing y logística en las tiendas con mejor desempeño.
* **Aplicar aprendizajes:** Evitar errores detectados en Tienda 3 (mala selección de productos, envío costoso, atención al cliente).
* **Explorar nuevos nichos:** Basado en el éxito de ciertas categorías, podría diversificarse el negocio en segmentos afines o emergentes.
* **Fortalecer presencia online y geográfica**, priorizando zonas con alta conversión.

## 10. Conclusiones y Perspectivas Futuras

### 10.1 Síntesis del Análisis Realizado

Este informe ha ofrecido una evaluación comparativa detallada de las cuatro tiendas en operación, analizando su desempeño financiero, estructura de productos, eficiencia operativa, satisfacción del cliente y distribución geográfica. La metodología empleada combinó análisis cuantitativo (ingresos, ventas, costos, valoraciones) con visualizaciones gráficas para identificar patrones y oportunidades.

Los resultados fueron concluyentes al señalar que:

* **Tienda 1** lidera en múltiples métricas clave.
* **Tienda 3** muestra claras señales de bajo rendimiento en casi todos los indicadores.
* La **satisfacción del cliente y la eficiencia logística** son factores decisivos que impactan los resultados.
* La diversificación y el enfoque geográfico también marcan diferencias notables en la competitividad.

### **10.2 Limitaciones del Estudio**

A pesar del análisis exhaustivo, se deben considerar ciertas limitaciones:

* **Datos históricos limitados**: Se trabajó con una única hoja por tienda, sin series temporales amplias.
* **Falta de contexto externo**: No se consideraron factores externos como competencia local, tendencias de mercado o estacionalidad.
* **Homogeneidad de formato**: Aunque se aplicó el mismo formato para comparar tiendas, podrían existir particularidades internas no reflejadas en los datos.
* **Reseñas no textuales**: El análisis de satisfacción se basó en calificaciones numéricas, sin incluir análisis semántico de reseñas.

### **10.3 Recomendaciones Adicionales**

Para potenciar el negocio a futuro, se sugieren las siguientes líneas de acción:

* 📊 **Implementar un sistema de análisis en tiempo real** para monitorear ventas, satisfacción y logística.
* 📦 **Optimizar el catálogo** eliminando productos con baja rotación y reforzando los más vendidos.
* 🚚 **Reducir costos logísticos** mediante alianzas estratégicas o ajustes en métodos de envío.
* 🧠 **Profundizar en análisis de clientes** incluyendo segmentación y hábitos de compra.
* 📍 **Explorar nuevas ubicaciones geográficas** basadas en los patrones actuales de concentración de ventas.
* 💡 **Fomentar la innovación** en atención al cliente, métodos de pago y experiencia de compra digital.

11. **Informe Consolidado de Desempeño de Tiendas**

**Dirigido a:** Sr. Juan  
**Elaborado por:** Rodrigo Fernando Riveros Aránguiz, científico de datos.  
**Fecha:** 4 de mayo de 2025

### **📅 Resumen Ejecutivo**

Este informe presenta un análisis comparativo del desempeño de las cuatro tiendas evaluadas, con base en los datos de ventas, ingresos, evaluaciones de clientes, productos y aspectos logísticos. El objetivo es brindar una base objetiva que permita tomar decisiones estratégicas sobre la continuidad o cierre de alguna de las tiendas.

### 📈 **Resumen Financiero por Tienda**

|  |  |  |
| --- | --- | --- |
| **Tienda** | **Ingresos Totales (COP)** | **Calificación Promedio** |
| Tienda 1 | 2,301,762,737 | 4.1 |
| Tienda 2 | 1,744,000,000 | 4.2 |
| Tienda 3 | 1,436,378,000 | 4.3 |
| Tienda 4 | 1,093,693,100 | 4.0 |

### 💼 **Productos Más Vendidos por Tienda**

* **Tienda 1:**
  + Microondas: 60 ventas
  + TV LED UHD 4K: 60 ventas
  + Armario: 60 ventas
* **Tienda 2:**
  + Escritorio: 70 ventas
  + Licuadora: 68 ventas
  + Zapatos deportivos: 66 ventas
* **Tienda 3:**
  + Silla ergonómica: 64 ventas
  + Monitor 24": 62 ventas
  + Celular ZYX: 61 ventas
* **Tienda 4:**
  + Cama box: 62 ventas
  + Cubertería: 59 ventas
  + Dashboards con Power BI: 56 ventas

### 🔻 **Productos Menos Vendidos por Tienda**

* **Tienda 1:**
  + Auriculares con micrófono: 33 ventas
  + Celular ABXY: 33 ventas
  + Olla de presión: 35 ventas
* **Tienda 2:**
  + Ventilador de torre: 34 ventas
  + Bicicleta eléctrica: 35 ventas
  + Cafetera exprés: 36 ventas
* **Tienda 3:**
  + Impresora multifunción: 34 ventas
  + Purificador de aire: 36 ventas
  + Parlante bluetooth: 37 ventas
* **Tienda 4:**
  + Guitarra eléctrica: 33 ventas
  + Armario: 34 ventas
  + Guitarra acústica: 37 ventas

### 📊 **Análisis de Categorías Más Vendidas**

(Todas las tiendas presentan ventas diversas, sin embargo, las categorías más relevantes por volumen general son: Electrodomésticos, Electrónica y Hogar/Muebles.)

### 🏛️ **Visualización Geográfica y Logística**

Se realizaron gráficos de dispersión por ubicación (latitud/longitud) de las ventas. Las tiendas 1 y 2 presentan mayor cobertura geográfica. Los boxplots de precios revelan una mayor variabilidad en la Tienda 3, especialmente en productos de tecnología.

### 🔢 **Recomendaciones Finales**

1. **Mantener la Tienda 1:** Líder en ingresos y con una calificación aceptable.
2. **Reforzar marketing de Tienda 2:** Buen nivel de ventas y calificaciones.
3. **Explorar crecimiento de Tienda 3:** Alta satisfacción de clientes, con margen de mejora en ingresos.
4. **Reevaluar continuidad de Tienda 4:** Menor desempeño general; considerar fusión, rebranding o cierre.

## Anexos

### Anexo A: Detalles Metodológicos

* **Fuente de datos**: Archivos Excel individuales por tienda, hoja "Worksheet".
* **Herramientas utilizadas**: Python 3.12, bibliotecas pandas, matplotlib, seaborn.
* **Criterios de análisis**:
  + Ingresos: suma de precios por producto.
  + Productos más/menos vendidos: conteo de frecuencia.
  + Valoración media: promedio de la columna de calificación.
  + Envío promedio: media de la columna de costo de envío.
  + Visualizaciones: gráficas de barras, torta, líneas, boxplot y dispersión.
* **Métricas clave para comparación**:
  + Total de ingresos por tienda.
  + Promedio de calificaciones por tienda.
  + Número de productos únicos vendidos.
  + Costo logístico promedio.
  + Diversidad de categorías.

### Anexo B: Tablas de Datos Complementarias

Se incluyen a continuación algunas tablas utilizadas como base para la comparación:

| **Tienda** | **Ingresos Totales** | **Calificación Promedio** | **Costo de Envío Promedio** |
| --- | --- | --- | --- |
| Tienda 1 | $XX,XXX | X.XX / 5 | $X.XX |
| Tienda 2 | $XX,XXX | X.XX / 5 | $X.XX |
| Tienda 3 | $XX,XXX | X.XX / 5 | $X.XX |
| Tienda 4 | $XX,XXX | X.XX / 5 | $X.XX |

(Opcional: se pueden adjuntar más tablas sobre ventas por categoría, productos destacados, etc.)

### Anexo C: Código Utilizado para el Análisis

El análisis fue desarrollado mediante scripts organizados por funcionalidad:

* ingresos\_totales.py: cálculo de ingresos por tienda.
* productos\_mas\_menos\_vendidos.py: identificación de top y bottom ventas.
* ventas\_por\_categoria.py: conteo y visualización de categorías.
* grafico\_barras.py, grafico\_torta.py, grafico\_lineas.py, grafico\_boxplot.py, grafico\_dispersion.py: visualizaciones específicas.
* calificaciones\_promedio.py: evaluación de satisfacción por tienda.
* Consolidación final para informe comparativo.

Se puede entregar el código completo como carpeta comprimida si se desea enviar a otros analistas o archivar para futuras auditorías.