

# Cadena Alura Store

*Javiera Torres - Enero 2026*

## I. Resumen Ejecutivo

Este informe presenta un análisis comparativo del desempeño de las cuatro tiendas de la cadena **Alura Store**, con el objetivo de apoyar al Sr. Juan en la decisión de cuál de ellas debería vender para iniciar un nuevo emprendimiento. Para ello, se utilizaron datos históricos de ventas, rendimiento operativo y reseñas de clientes, contenidos en archivos CSV.

El análisis se realizó utilizando la biblioteca **Pandas** para la carga, limpieza y manipulación de los datos, y **Matplotlib** para la creación de visualizaciones. Se evaluaron métricas clave como los ingresos totales, las categorías de productos más vendidas, la cantidad y calificación de reseñas, los productos con mayor volumen de ventas y el costo promedio de envío por tienda.

A partir de estos indicadores, se identificaron diferencias significativas en el rendimiento entre las tiendas. Los resultados muestran que una de las sucursales presenta consistentemente menores ingresos, menor volumen de ventas y un desempeño inferior en métricas de satisfacción del cliente en comparación con las demás.

Con base en estos hallazgos, se emite una recomendación final sobre cuál tienda debería ser vendida, sustentada en evidencia cuantitativa y visualizaciones claras. Este análisis busca demostrar cómo el uso de datos puede apoyar la toma de decisiones estratégicas en contextos empresariales reales.

## II. Objetivo del Análisis

El objetivo de este análisis es evaluar y comparar el desempeño de las cuatro tiendas de la cadena **Alura Store**, utilizando datos de ventas, rendimiento operativo y reseñas de clientes, con el fin de identificar cuál de ellas presenta la menor eficiencia global.

A través del análisis de métricas clave como los ingresos totales, las categorías de productos más vendidas, el volumen de ventas, la satisfacción de los clientes y el costo promedio de envío, se busca generar evidencia objetiva que apoye la toma de decisiones estratégicas del Sr. Juan respecto a qué tienda debería vender para iniciar un nuevo emprendimiento.

### **III. Descripción de Dataset**

El conjunto de datos utilizado en este proyecto corresponde a información de ventas y reseñas de clientes de una de las tiendas de la cadena **Alura Store**, almacenada en un archivo en formato CSV.

El dataset contiene un total de **2.358 registros**, donde cada fila representa una transacción individual realizada por un cliente. Cada registro incluye información tanto comercial como geográfica y de satisfacción del cliente.

Las columnas incluidas en el conjunto de datos son las siguientes:

- **Producto:** nombre del producto vendido.
- **Categoría del Producto:** categoría a la que pertenece el producto (por ejemplo, electrónicos, muebles, libros, entre otros).
- **Precio:** precio de venta del producto.
- **Costo de envío:** valor del envío asociado a la compra.
- **Fecha de compra:** fecha en que se realizó la transacción.
- **Vendedor:** identificador o nombre del vendedor.
- **Lugar de compra:** ubicación donde se realizó la compra.  
**Calificación:** calificación otorgada por el cliente al producto o a la experiencia de compra.
- **Método de pago:** medio de pago utilizado por el cliente (tarjeta, transferencia, etc.).
- **Cantidad de cuotas:** número de cuotas en que se realizó el pago.
- **lat:** latitud del lugar de compra.
- **lon:** longitud del lugar de compra.

Este conjunto de datos permitió analizar el comportamiento de ventas, la distribución de productos por categoría, los ingresos generados, los costos de envío, la satisfacción de los clientes y la distribución geográfica de las compras, aportando una base sólida para la comparación del desempeño entre tiendas.

### **IV. Metodología**

El análisis se desarrolló utilizando el lenguaje de programación **Python** y las bibliotecas **Pandas** y **Matplotlib** para la manipulación, análisis y visualización de los datos.

#### **a) Extracción y carga de datos**

En primer lugar, se cargaron los archivos CSV correspondientes a las tiendas de la cadena Alura Store mediante la biblioteca Pandas. Los conjuntos de datos fueron almacenados en estructuras tipo DataFrame para facilitar su exploración y posterior procesamiento.

#### b) Exploración y preparación de los datos

Se realizó una exploración inicial para comprender la estructura del dataset, revisar los tipos de datos de cada columna e identificar posibles valores faltantes o inconsistentes.

Asimismo, se verificó el formato de variables clave como *Precio*, *Costo de envío*, *Calificación* y *Fecha de compra*, realizando las conversiones necesarias para su correcto uso en los cálculos posteriores.

#### c) Cálculo de métricas clave por tienda

A partir de los datos cargados, se calcularon las siguientes métricas principales:

- **Ingreso total de cada tienda**, sumando los valores de la columna *Precio* por transacción.
- **Ventas por categoría**, mediante la agrupación de los registros por *Categoría del Producto* y el conteo de transacciones.
- **Productos más vendidos y menos vendidos**, identificados a partir del número de veces que aparece cada producto en el dataset.
- **Costo de envío promedio**, calculado como el promedio de la columna *Costo de envío* por tienda.
- **Valoración media por tienda**, obtenida como el promedio de la columna *Calificación*.

Estas métricas fueron calculadas utilizando funciones de agregación y agrupación de Pandas, tales como `groupby()`, `count()`, `sum()` y `mean()`.

#### d) Análisis del desempeño geográfico

Como parte del componente adicional del desafío, se realizó un **análisis del desempeño geográfico** utilizando las columnas *lat* y *lon*.

Este análisis permitió visualizar la distribución espacial de las compras y detectar zonas con mayor concentración de ventas, aportando una perspectiva territorial al comportamiento de los clientes.

#### e) Visualización de resultados

Para comunicar los hallazgos de manera clara y visual, se generaron distintas visualizaciones utilizando la biblioteca Matplotlib, incluyendo:

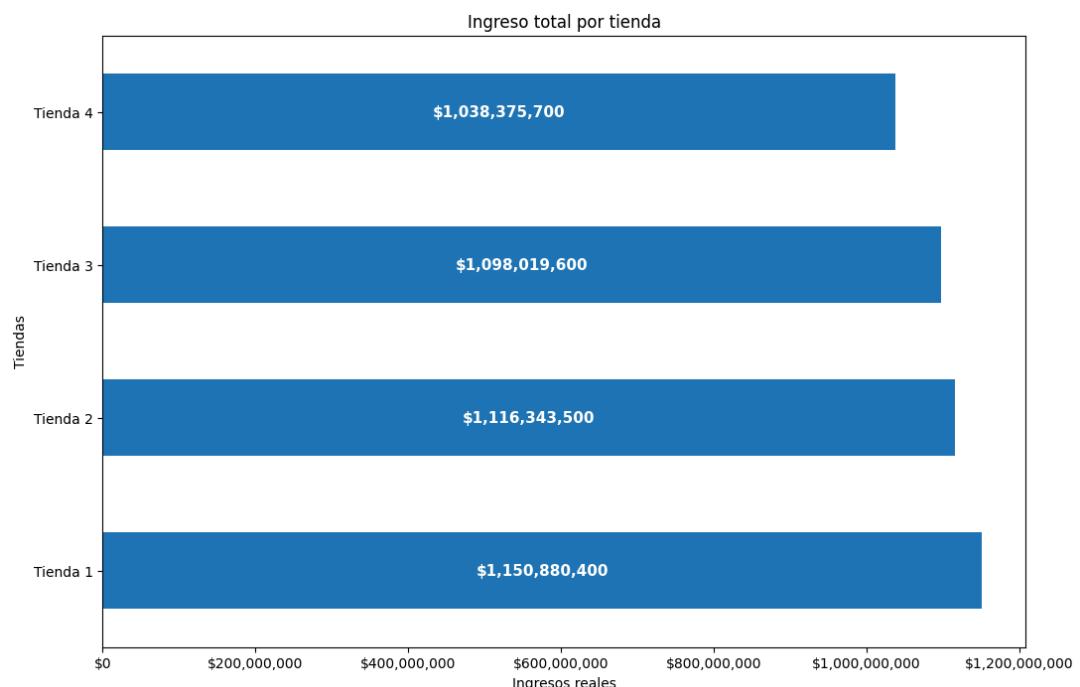
- Gráficos de barras para comparar el ingreso total por tienda.

- Gráficos de barras para representar las ventas por categoría.
- Gráficos para mostrar los productos más y menos vendidos.
- Visualizaciones descriptivas del costo de envío promedio y la valoración media por tienda.
- Representaciones gráficas del desempeño geográfico basadas en coordenadas de latitud y longitud.

## V. Análisis y resultados

### 1. Análisis de facturación

El gráfico presenta el ingreso total generado por cada una de las cuatro tiendas de Alura Store, lo que permite comparar directamente el desempeño financiero de cada sucursal.



Los resultados muestran que:

- **Tienda 1** es la tienda con **mayor facturación**, alcanzando un ingreso total aproximado de **\$1.150.880.400**, lo que la posiciona como la sucursal con mejor desempeño económico dentro de la cadena.
- **Tienda 2** ocupa el segundo lugar, con ingresos cercanos a **\$1.116.343.500**, manteniéndose también en un nivel elevado de facturación.
- **Tienda 3** registra un ingreso total de aproximadamente **\$1.098.019.600**, ligeramente inferior al de las dos primeras tiendas.

- **Tienda 4** presenta la menor facturación, con ingresos alrededor de **\$1.038.375.700**, ubicándose como la tienda con el desempeño financiero más bajo del conjunto.

En términos comparativos, se observa una diferencia de aproximadamente **\$112.504.700** entre la tienda con mayor facturación (Tienda 1) y la de menor facturación (Tienda 4). Esta brecha es significativa y sugiere que la Tienda 4 genera considerablemente menos ingresos que el resto de las sucursales.

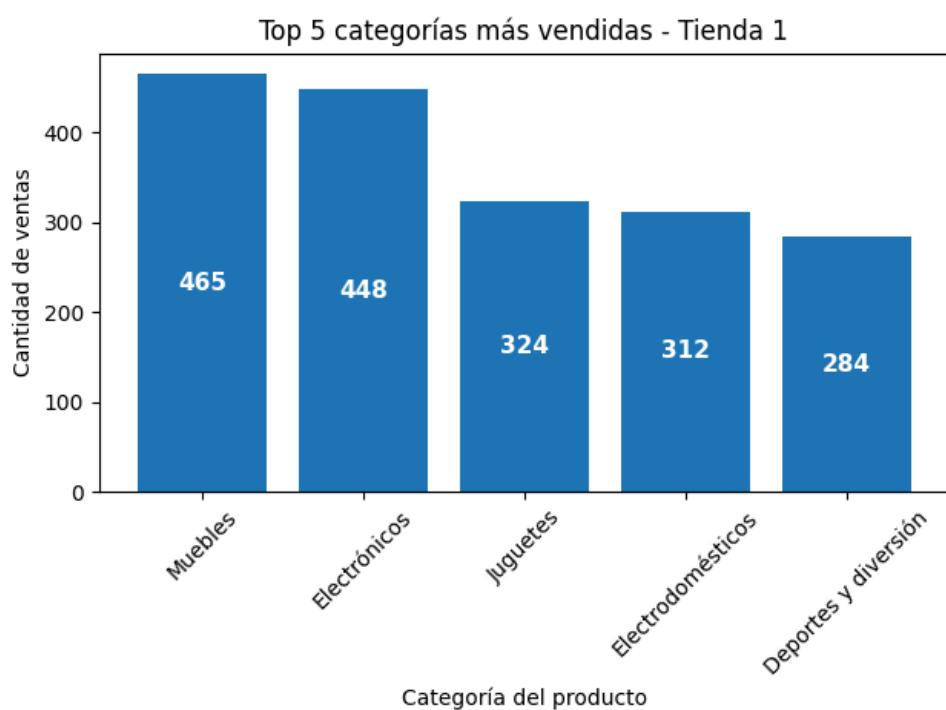
Este resultado indica que **Tienda 4** presenta un desempeño económico inferior en relación con las demás tiendas, lo que puede estar asociado a un menor volumen de ventas, menor valor promedio por transacción o una combinación de ambos factores.

## 2. Ventas por categorías

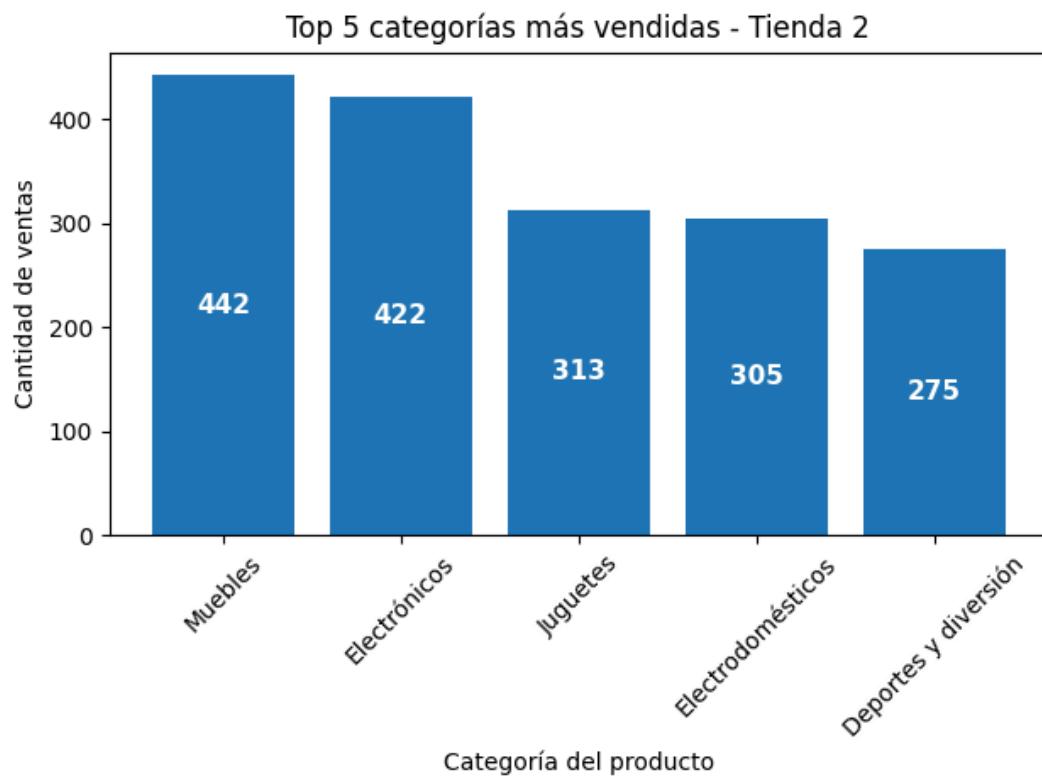
Para analizar las categorías más populares de cada tienda, se agruparon los datos por la variable *Categoría del Producto* y se contó el número de transacciones asociadas a cada categoría.

Posteriormente, se ordenaron los resultados de forma descendente para identificar las categorías con mayor volumen de ventas en cada sucursal.

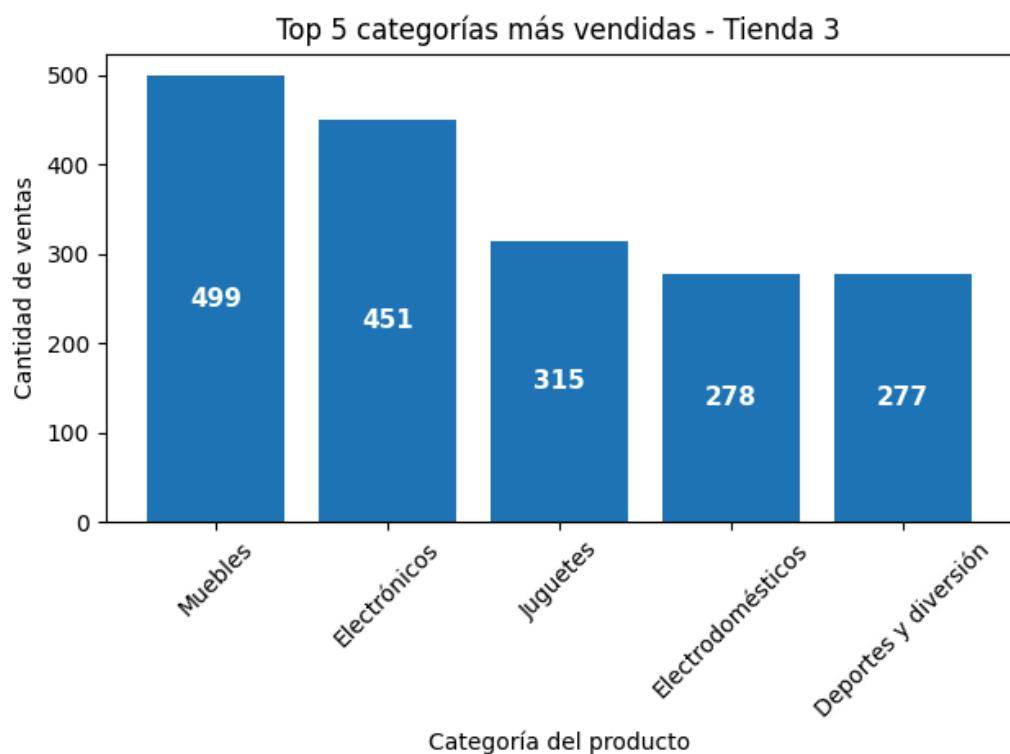
### Tienda 1



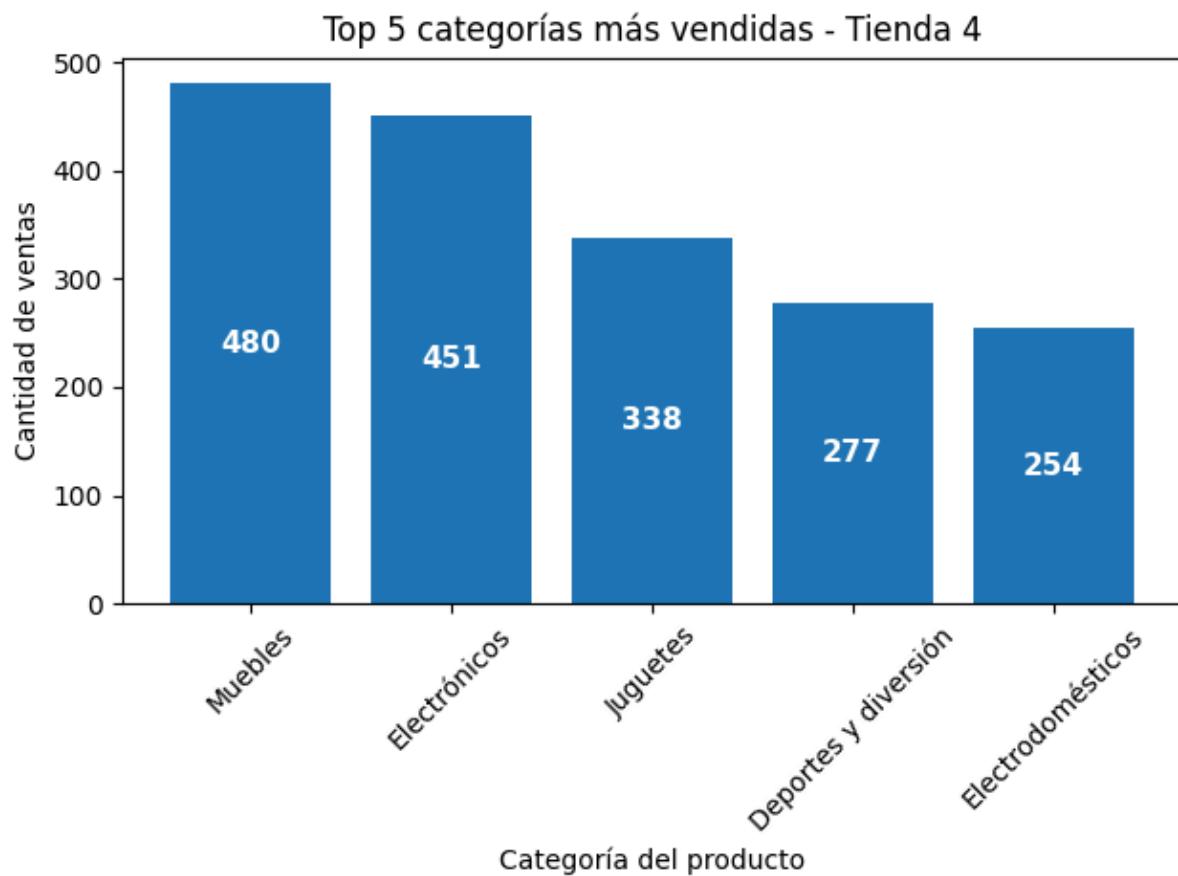
## Tienda 2



## Tienda 3



Tienda 4:

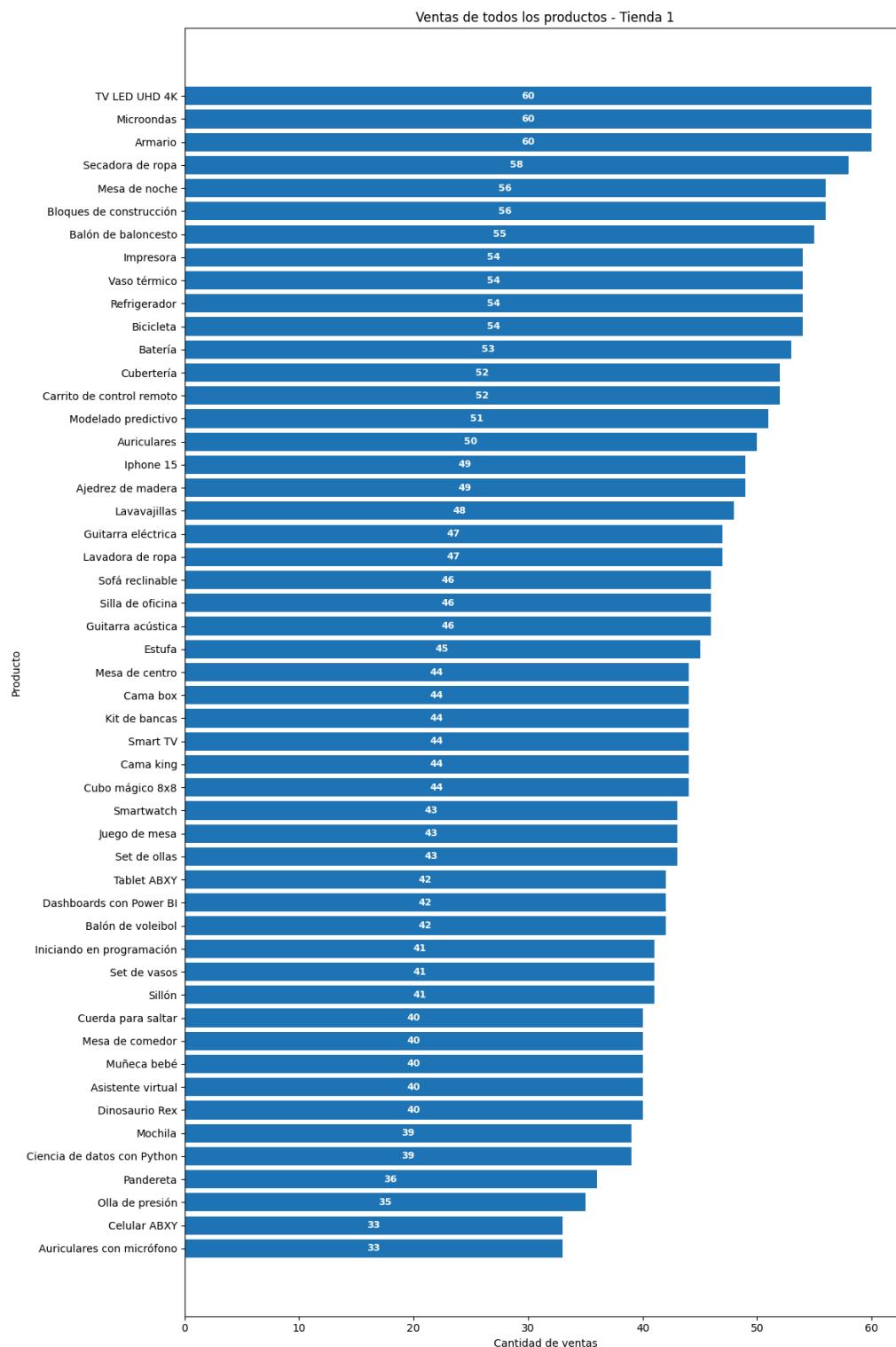


### 3. Productos más vendidos y menos vendidos

Para analizar el desempeño individual de cada producto, se agruparon los datos por la variable *Producto* y se contó el número de transacciones asociadas a cada uno en cada tienda.

Los resultados se ordenaron de forma ascendente y se visualizaron mediante gráficos de barras horizontales, lo que permitió identificar tanto los productos con mayor volumen de ventas como aquellos con baja rotación en cada sucursal.

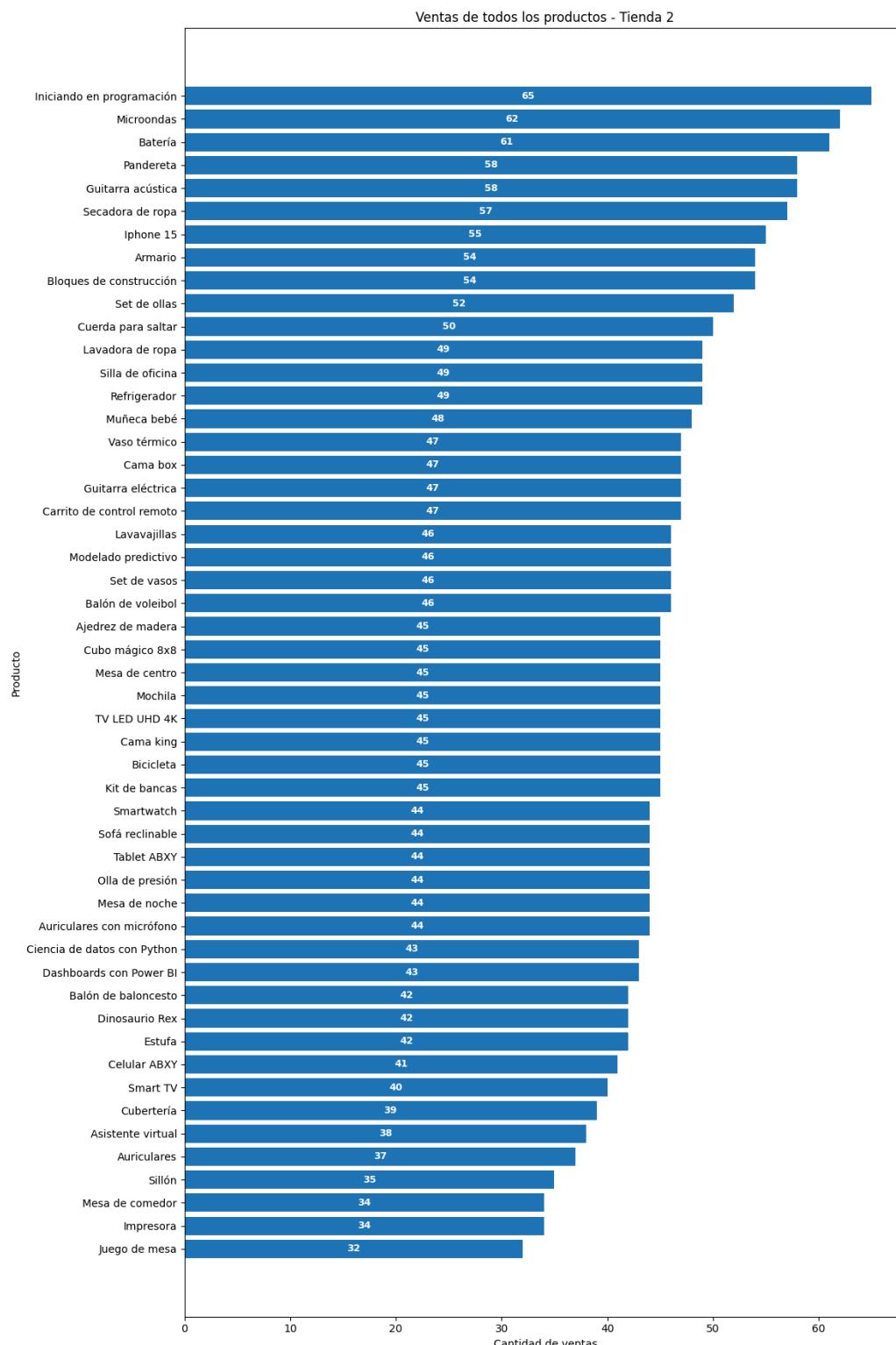
## Tienda 1:



Los productos más vendidos de la tienda 1 son: TV LED UHD 4K con 60 productos vendidos, le sigue Microondas con la misma cantidad y le iguala Armarios. Por el contrario los productos menos vendidos con Olla de presión con 35 productos vendidos, le sigue celular

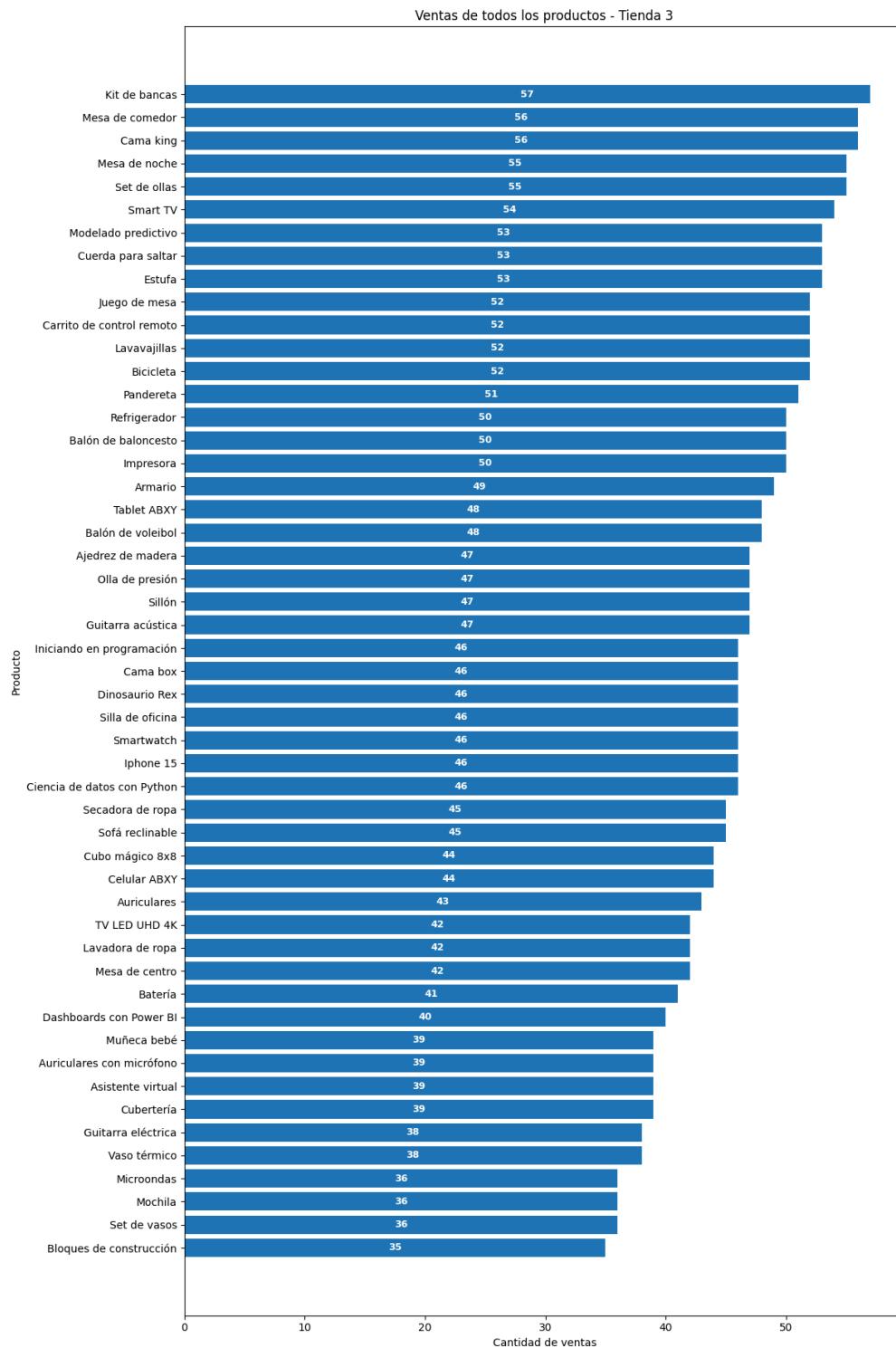
ABXY con 33 productos vendidos y el artículo más vendido es Auriculares con micrófono que se vendió 33 veces.

## Tienda 2:



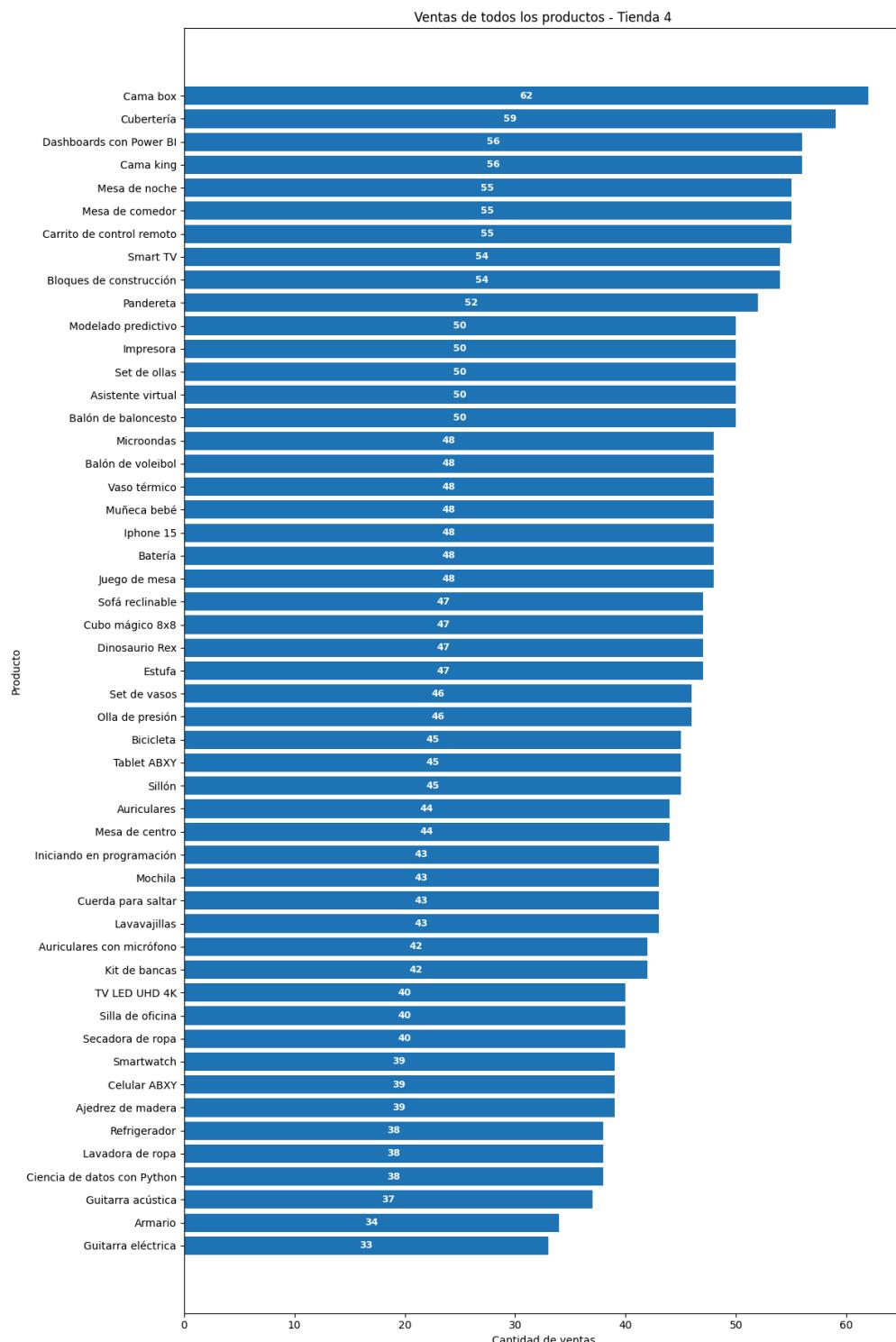
Los productos que más se vendieron fueron Iniciando en la programación vendido 65 veces, le sigue Microondas vendido 62 veces y Batería vendida 61 veces. Por otro lado, los productos menos vendidos son: Juego de mesa vendido 32 veces, Impresora vendida 34 veces y Mesa de Comer vendida 34 veces.

### Tienda 3:



Los productos más vendidos de la tienda 4 son: Kid de bancas que se vendió 57 veces, le sigue Mesa de comedor que se vendió 56 veces y Cama king que se vendió de igual manera 56 veces. Por el contrario los productos que menos se vendieron fueron: Bloques de construcción vendidas 35 veces, le sigue kit de vasos que se vendió 36 veces y el igualó mochila con la misma cantidad vendida.

#### Tienda 4:



Los productos más vendidos de la tienda 4 son: Cama box con 62 productos vendidos, le sigue Cubría con 62 productos vendidos y Dashboards con Power Bi que se vendió 59 veces. Sin embargo, los productos menos vendidos son: Guitarra Eléctrica vendida 33 veces, le sigue Armario vendido 34 veces y Guitarra acústica con 37 productos vendidos.

#### 4. Valor del envío promedio por tienda



El gráfico muestra el costo promedio de envío para cada una de las cuatro tiendas de Alura Store, permitiendo comparar el nivel de gasto logístico asociado a cada sucursal.

Los resultados indican que:

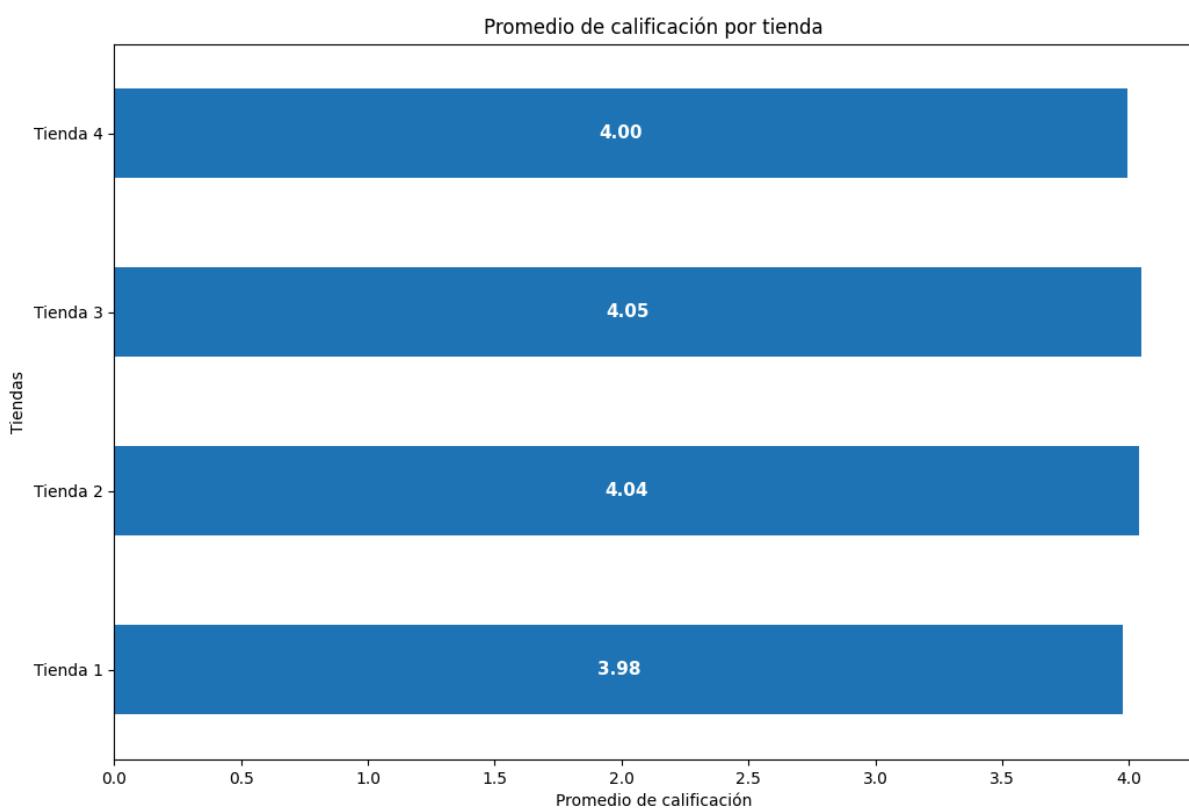
- **Tienda 1** presenta el **costo promedio de envío más alto**, con un valor aproximado de **\$26.018,61**, lo que sugiere que esta tienda incurre en mayores gastos logísticos por pedido en comparación con las demás.
- **Tienda 2** ocupa el segundo lugar, con un costo promedio cercano a **\$25.216,24**, manteniéndose también en un rango elevado de costos de envío.
- **Tienda 3** muestra un costo intermedio, con un valor promedio de **\$24.805,68**, ligeramente inferior al de las dos primeras tiendas.
- **Tienda 4** registra el **costo promedio de envío más bajo**, con aproximadamente **\$23.459,46**, lo que indica una operación logística relativamente más eficiente o una mejor ubicación respecto a sus clientes.

En términos comparativos, se observa una diferencia de alrededor de **\$2.559** entre la tienda con el mayor costo promedio de envío (Tienda 1) y la más eficiente en este aspecto (Tienda 4). Aunque esta brecha no es extremadamente amplia, puede tener un impacto significativo en la rentabilidad total cuando se considera el volumen acumulado de ventas.

Este análisis sugiere que **Tienda 4** cuenta con una ventaja logística, posiblemente debido a una mejor distribución geográfica de sus clientes, menores distancias de envío o acuerdos más favorables con proveedores de transporte. Por el contrario, **Tienda 1** podría estar enfrentando desventajas operativas que elevan sus costos de despacho.

En conjunto, el costo promedio de envío es un indicador clave para evaluar la eficiencia operativa de cada tienda. Estos resultados pueden ser utilizados para apoyar decisiones estratégicas, como la optimización de rutas logísticas, la renegociación de tarifas de envío o, en el contexto del desafío, como un factor adicional para identificar la tienda menos eficiente.

## 5. Valoración media por tienda



El gráfico presenta el promedio de calificación otorgado por los clientes a cada una de las cuatro tiendas de Alura Store, lo que permite evaluar el nivel de satisfacción general asociado a cada sucursal.

Los resultados muestran que:

- **Tienda 3** obtiene la **mejor valoración promedio**, con una calificación aproximada de **4,05**, lo que indica un alto nivel de satisfacción por parte de sus clientes.
- **Tienda 2** se posiciona muy cerca, con un promedio de **4,04**, evidenciando también una percepción positiva y consistente del servicio y los productos ofrecidos.
- **Tienda 4** alcanza una calificación media de **4,00**, manteniéndose dentro de un rango favorable y alineado con el desempeño general del resto de las tiendas.
- **Tienda 1** presenta la **valoración promedio más baja**, con aproximadamente **3,98**, aunque la diferencia con las demás tiendas es mínima.

En términos generales, se observa que **todas las tiendas cuentan con calificaciones cercanas a 4 puntos**, lo que refleja un nivel de satisfacción alto y relativamente homogéneo entre los clientes. La diferencia máxima entre la tienda mejor valorada (Tienda 3) y la menos valorada (Tienda 1) es de apenas **0,07 puntos**, lo que sugiere que la experiencia del cliente es consistente en toda la cadena.

Este resultado indica que, desde la perspectiva de satisfacción del cliente, **no existen brechas significativas entre las tiendas**, por lo que la calificación promedio, por sí sola, no constituye un criterio determinante para decidir qué tienda debería venderse. Sin embargo, combinada con otras métricas como ingresos, volumen de ventas y costos de envío, esta variable aporta un contexto importante para evaluar el desempeño global de cada sucursal.

## 6. Desempeño geográfico

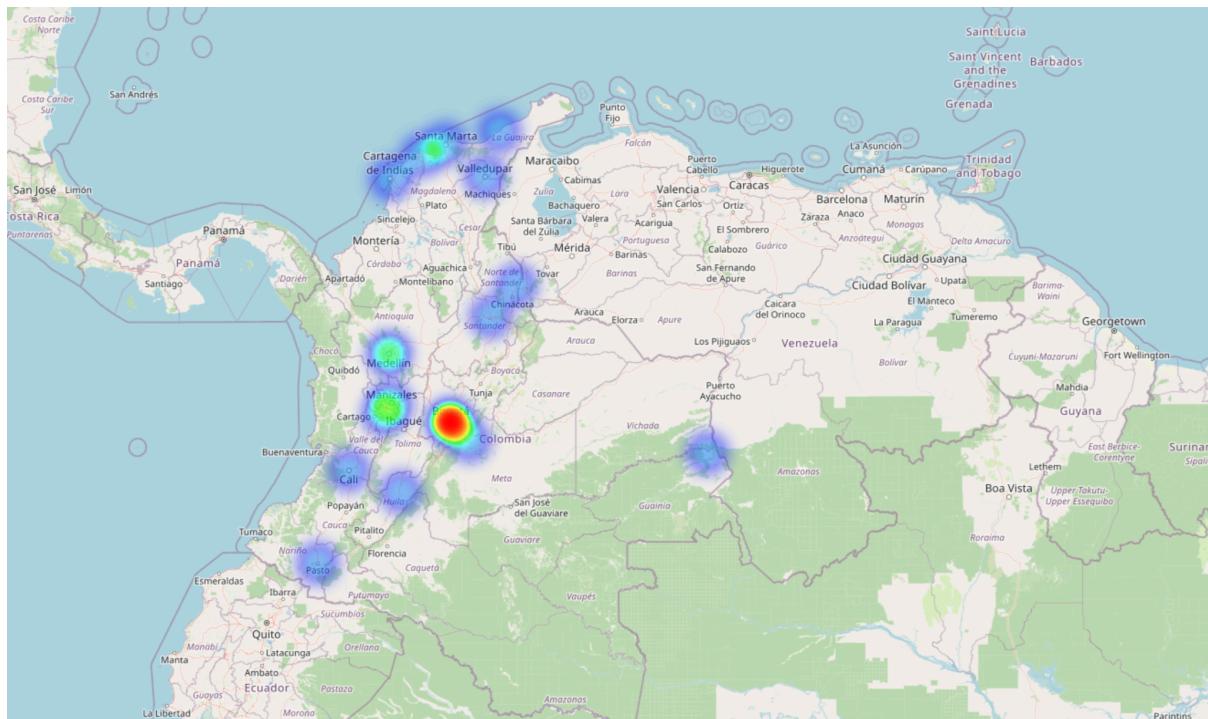
Como parte del extra del desafío, se analizó la distribución geográfica de las ventas utilizando las coordenadas de latitud y longitud incluidas en el dataset.

A través de visualizaciones de dispersión, se identificaron zonas con una mayor concentración de compras, lo que evidencia que las ventas no se distribuyen de manera uniforme en el territorio.

Asimismo, se exploró la relación entre la ubicación geográfica, los ingresos y las calificaciones de los clientes, observándose que algunas regiones presentan un mejor desempeño comercial y mayores niveles de satisfacción.

Este análisis aporta una visión territorial del comportamiento de los clientes y permite detectar patrones geográficos que pueden influir en el rendimiento de las tiendas.

## Tienda 1:



El mapa de calor muestra la distribución geográfica de las ventas realizadas por la Tienda 1, utilizando las coordenadas de latitud y longitud de cada transacción. La intensidad del color representa la concentración de ventas, donde los tonos más cálidos (rojo y amarillo) indican una mayor densidad de pedidos, mientras que los tonos fríos (azul) reflejan una menor actividad comercial.

### Principales hallazgos

Se observa una **alta concentración de ventas en la región central de Colombia**, particularmente en el área correspondiente a **Bogotá y sus alrededores**, donde aparece el punto de mayor intensidad en el mapa. Esto sugiere que la Tienda 1 tiene una fuerte presencia comercial en esta zona, probablemente asociada a una mayor densidad poblacional, mayor poder adquisitivo y mejor conectividad logística.

Además, se identifican **focos secundarios de ventas** en otras ciudades importantes del país, como: Medellín, Cali, Barranquilla y Cartagena.

Estas áreas presentan una intensidad media, lo que indica una participación relevante en el volumen total de ventas, aunque menor en comparación con la región central.

Por otro lado, se detecta una **baja densidad de ventas en las regiones del sur y sureste del país**, así como en zonas rurales y de menor densidad poblacional. Esto podría estar relacionado con:

- Menor acceso a servicios logísticos
- Mayores costos o tiempos de envío
- Menor penetración del comercio electrónico
- Menor visibilidad de la marca en estas regiones

### **Implicancias para el rendimiento de ventas**

El patrón observado sugiere que el desempeño de la Tienda 1 está **altamente concentrado geográficamente**, dependiendo en gran medida de un núcleo urbano principal para generar la mayor parte de sus ingresos.

Esto implica que:

- La tienda es **muy fuerte en mercados urbanos consolidados**, pero
- Presenta una **penetración limitada en regiones periféricas**.

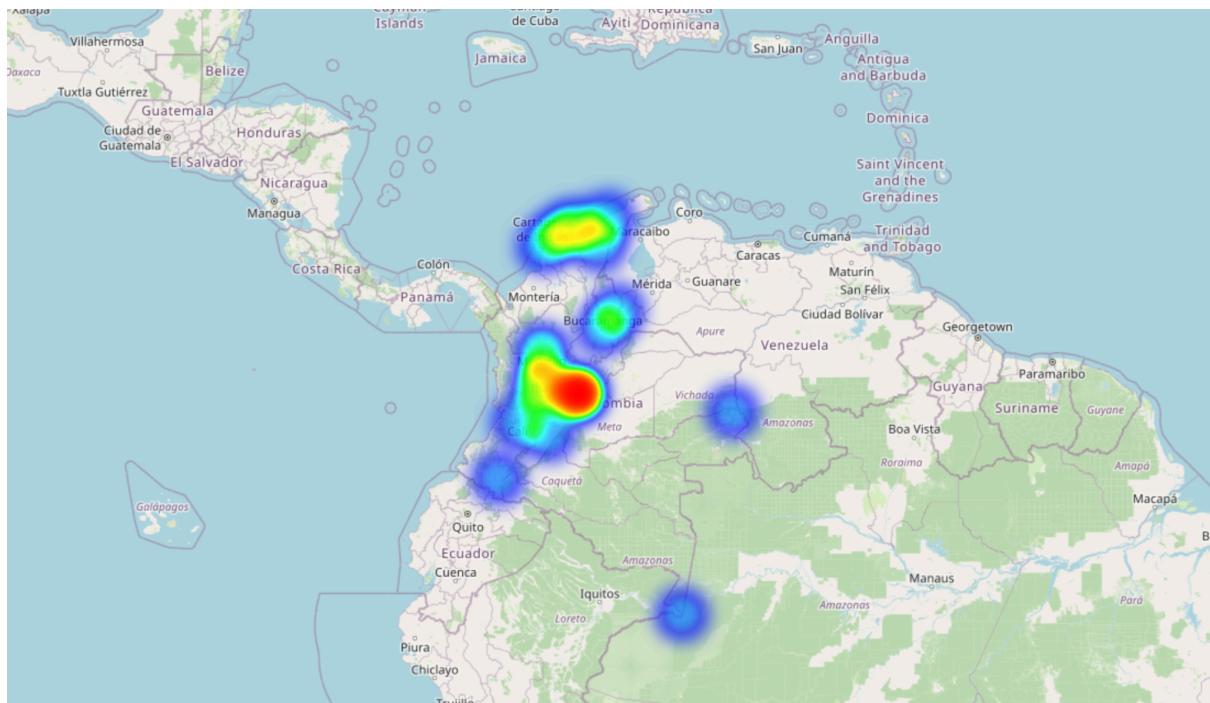
### **Interpretación**

El mapa de calor evidencia que las ventas de la Tienda 1 se concentran principalmente en la región central de Colombia, con un foco dominante en el área de Bogotá y presencia secundaria en otras ciudades principales como Medellín y Cali.

Este patrón sugiere que el rendimiento de la tienda está estrechamente vinculado a mercados urbanos de alta densidad poblacional, mientras que su alcance en regiones más alejadas o rurales es limitado.

La distribución geográfica de las ventas podría estar influyendo tanto en los ingresos totales como en los costos de envío, y representa una oportunidad para diseñar estrategias de expansión territorial o de optimización logística.

## Tienda 2:



El mapa de calor muestra la distribución espacial de las ventas realizadas por la Tienda 2, utilizando las coordenadas de latitud y longitud asociadas a cada transacción. La intensidad del color representa la concentración de ventas, donde los tonos más cálidos (rojo y amarillo) indican una mayor densidad de pedidos.

### Principales hallazgos

A diferencia de la Tienda 1, la **Tienda 2 presenta una distribución geográfica más dispersa**, con múltiples focos de alta intensidad repartidos en distintas regiones del país.

Se identifica un **núcleo principal de ventas en la región central de Colombia**, nuevamente en el área de **Bogotá y alrededores**, donde se observa la mayor concentración de transacciones.

Sin embargo, a diferencia del patrón altamente centralizado de la Tienda 1, la Tienda 2 muestra **focos secundarios de intensidad alta y media** en:

- La región Caribe (zona de Barranquilla y Cartagena)
- La región noroccidental (área de Medellín)
- El suroccidente del país (zona de Cali)

Además, se observan **puntos aislados de ventas en regiones más alejadas**, incluyendo zonas del sur y sureste (Amazonas y áreas limítrofes), lo que indica una mayor cobertura territorial en comparación con la Tienda 1.

### **Implicancias para el rendimiento de ventas**

El patrón geográfico de la Tienda 2 sugiere que:

- Cuenta con una **base de clientes más diversificada territorialmente**.
- No depende exclusivamente de un único núcleo urbano para sostener su volumen de ventas.
- Presenta una **mayor penetración en regiones periféricas** del país.

Esto puede explicar, al menos en parte, por qué la Tienda 2 logra mantener un nivel alto de facturación, aunque con un **costo promedio de envío mayor** que el de la Tienda 4, debido a:

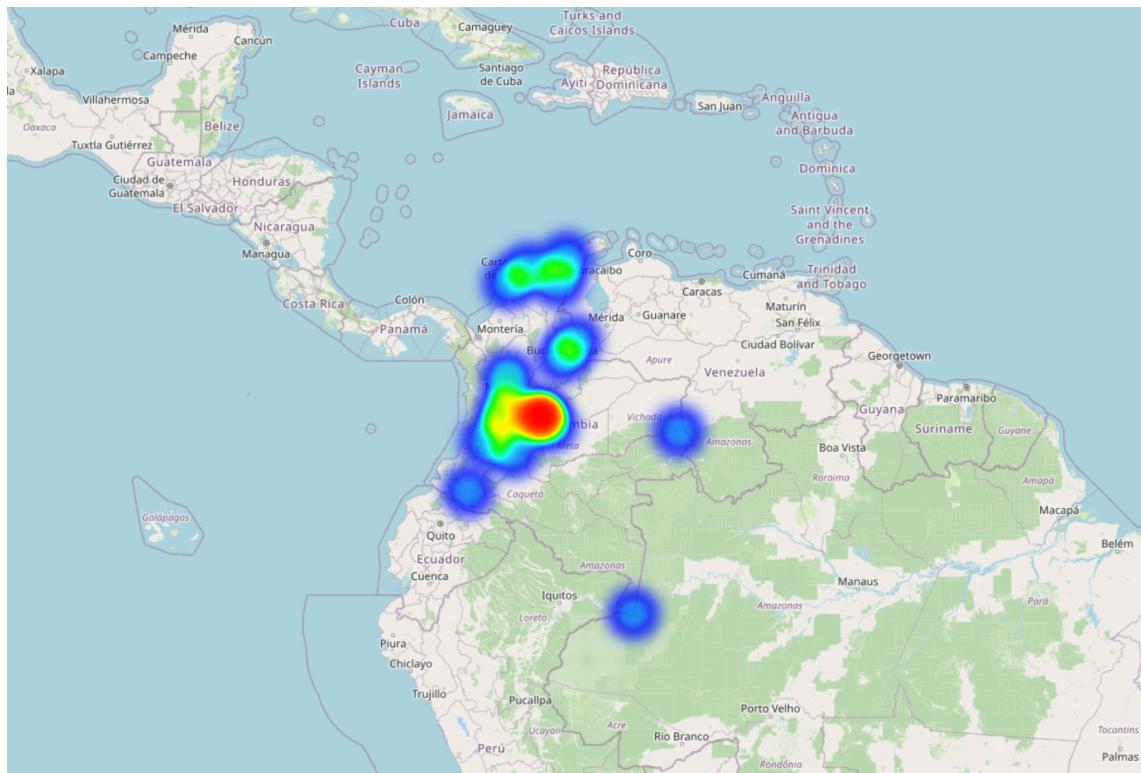
- Mayores distancias de despacho
- Mayor dispersión geográfica de los pedidos

### **Interpretación**

El mapa de calor evidencia que la Tienda 2 posee una distribución de ventas más amplia y menos concentrada que la Tienda 1, con múltiples focos relevantes en distintas regiones del país, incluyendo la región central, la costa Caribe y el suroccidente.

Este patrón sugiere una mayor cobertura territorial y una menor dependencia de un solo mercado urbano, lo que representa una fortaleza comercial. No obstante, esta dispersión geográfica también podría estar contribuyendo a un mayor costo promedio de envío.

### Tienda 3:



El mapa de calor de la **Tienda 3** muestra una fuerte concentración de ventas en la **zona central de Colombia**, especialmente en el eje que conecta **Bogotá, Medellín y Cali**. Esta franja aparece marcada con colores rojo y amarillo, lo que indica una **alta densidad de transacciones** en estas áreas metropolitanas.

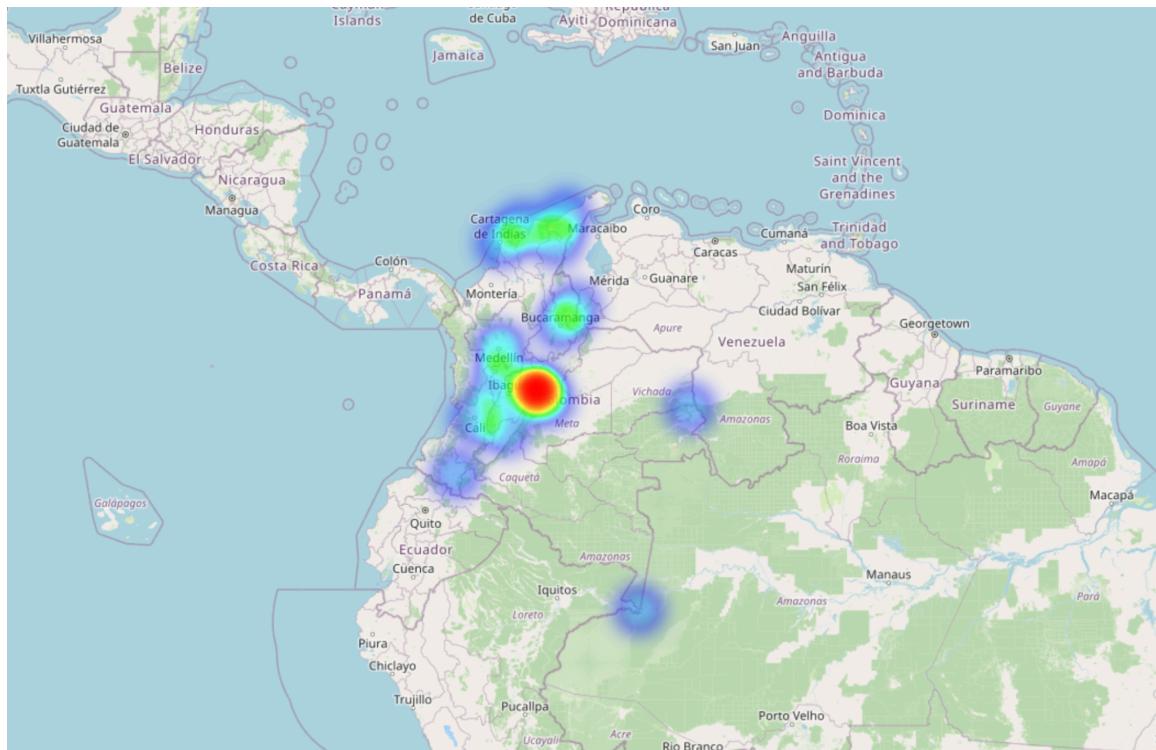
También se observan focos secundarios en:

- La región Caribe (zona de Cartagena y Barranquilla).
- El suroccidente del país (cercanías de Pasto y el eje cafetero).
- Algunas zonas del oriente y sur del país, aunque con menor intensidad.

### Interpretación

- La Tienda 3 tiene una **cobertura amplia y relativamente equilibrada**, con presencia tanto en grandes ciudades como en regiones periféricas.
- La alta concentración en el centro del país sugiere que esta tienda tiene una **fuerte penetración en los principales mercados urbanos**.
- La dispersión moderada hacia el sur y el oriente indica oportunidades de crecimiento en esas zonas.

#### Tienda 4:



En el caso de la **Tienda 4**, el mapa de calor muestra un patrón muy similar, pero con una **concentración aún más intensa en la zona central**, especialmente alrededor de **Bogotá y Medellín**, donde se observa el punto más caliente (rojo).

Además, se identifican:

- Focos relevantes en la región Caribe.
- Presencia en el suroccidente (Cali y alrededores).
- Algunos puntos aislados en el sur y la Amazonía, aunque con baja densidad.

#### Interpretación

- La Tienda 4 presenta una **estrategia claramente enfocada en los grandes centros urbanos**.
- Su rendimiento parece estar **altamente concentrado en pocas regiones clave**, lo que puede indicar una dependencia geográfica.
- La baja intensidad en regiones periféricas sugiere un **potencial de expansión territorial**

## **Conclusión global del análisis geográfico (Tiendas 1–4)**

El análisis geográfico de las ventas muestra un patrón consistente en las cuatro tiendas: la **mayor concentración de ventas se ubica en la zona central del país**, especialmente en grandes ciudades como Bogotá, Medellín y Cali. Estas áreas concentran la mayor actividad comercial debido a su alta densidad poblacional y mejor infraestructura.

En contraste, las regiones periféricas presentan una **menor intensidad de ventas**, lo que sugiere oportunidades de mejora en cobertura, logística o estrategias comerciales. En general, los resultados confirman que la **ubicación geográfica influye directamente en el rendimiento de las tiendas** y es un factor clave para la toma de decisiones estratégicas.

## **IV. Conclusiones**

A partir del análisis de los datos de ventas, costos logísticos y valoraciones de clientes de las cuatro tiendas de Alura Store, se puede obtener una visión integral del desempeño de cada sucursal.

En primer lugar, el **análisis de facturación** muestra que la **Tienda 4** es la sucursal con **menor ingreso total**, ubicándose por debajo de las demás tiendas con una diferencia significativa respecto a la de mayor facturación (Tienda 1). Este resultado indica que la Tienda 4 genera menos valor económico para la cadena, lo que la posiciona como la más débil desde una perspectiva financiera.

En segundo lugar, el **análisis del costo promedio de envío** revela que la **Tienda 4 presenta el menor costo logístico**, lo que sugiere una mayor eficiencia en términos de distribución. Sin embargo, esta ventaja operativa no logra compensar su bajo nivel de ingresos, por lo que su impacto positivo en la rentabilidad global es limitado.

Por otro lado, el **análisis de la valoración media por tienda** muestra que todas las sucursales presentan calificaciones muy similares, cercanas a 4 puntos. Esto indica que el nivel de satisfacción del cliente es alto y homogéneo en toda la cadena. En consecuencia, la experiencia del cliente no representa un factor diferenciador relevante entre tiendas, ni justifica por sí sola el bajo desempeño financiero de la Tienda 4.

## **IV. Recomendaciones**

Con base en el análisis de los ingresos, los costos promedio de envío y la valoración de los clientes, se recomienda que el Sr. Juan considere vender la **Tienda 4**, ya que presenta el desempeño financiero más bajo dentro de la cadena.

A pesar de contar con costos logísticos más eficientes y una buena calificación promedio, estos factores no compensan su menor nivel de facturación. La venta de esta tienda permitiría liberar recursos para invertir en un nuevo emprendimiento con mayor potencial de crecimiento y rentabilidad.

#### **III. Limitaciones**

El análisis presenta limitaciones relacionadas con el alcance temporal de los datos, la ausencia de información sobre costos operativos totales y el uso de métricas promedio que pueden ocultar variabilidad. Además, no se consideran factores externos ni se profundiza en el volumen de ventas o en el impacto cuantitativo de la ubicación geográfica. Estas limitaciones deben tenerse en cuenta al interpretar los resultados y la recomendación final.

#### **IX. Líneas Futuras de Análisis**

Como trabajo futuro, se propone incorporar una dimensión temporal, calcular la rentabilidad neta por tienda e integrar cuantitativamente el análisis geográfico. Además, se sugiere segmentar clientes y productos, analizar volumen de ventas y aplicar modelos predictivos para apoyar decisiones estratégicas.