

Alumno: Federico Gauna

**Título:** Análisis de la evolución de las ventas de un e-commerce

### **Objetivos e Hipótesis**

En este proyecto, se analizarán las ventas de un e-commerce en crecimiento. Se nos presentan datasets con información sobre los clientes, vendedores, las transacciones entre estos y un dataset provisto por un tercero con el poder de compra de cada cliente. También, se usará un dataset con el tipo de cambio a dólar para poder evaluar este crecimiento en términos reales debido al contexto inflacionario.

En términos generales, lo que se hará es realizar un análisis sobre estos datos para visualizar como es la evolución de las ventas, revisando los puntos altos y bajos con respecto a la cantidad de ventas y montos totales, clientes más frecuentes y vendedores más populares.

El periodo de tiempo sobre el cual se encuentran los datos es entre enero y octubre del 2023. Pero, se trabajarán los datos de enero hasta septiembre del 2023 *-explicado a más a detalle en el preprocesamiento-*.

Algunos puntos a tener en cuenta en el análisis son los siguientes:

- Cantidad de ventas de cada vendedor y cantidad de compras de cada cliente.
- Crecimiento o decrecimiento de las ventas en pesos y dólares.
- Cantidad de oferta de productos por categoría.
- Monto total por categoría en pesos y dólares.
- Cantidad de compras de los clientes según su poder de compra.

### **Alcance**

Del análisis propuesto se armará un dashboard a nivel operacional, el cual tiene como finalidad mostrar la evolución de las ventas del e-commerce en términos reales e identificar tanto los puntos destacados que deben ser preservados como las áreas de oportunidad que requerirán una atención especial. Estos últimos serán el enfoque primordial para implementar mejoras significativas. Por esto, este análisis será dirigido a los Managers del emprendimiento.

## **Dataset**

Contamos con cuatro datasets: customers, customers\_purchasing\_power, merchants y transactions.

El dataset de “customers” contiene información sobre los clientes que utilizan el e-commerce para comprar productos.

El dataset de “customers\_purchasing\_power” contiene información sobre el segmento socioeconómico de los clientes, indicando “low”, “mid” y “high” según tengan más o menos poder de compra.

El dataset de “merchants” contiene la información de los vendedores que utilizan el e-commerce para publicar y vender sus productos.

El dataset de “transactions” contiene toda la información sobre las transacciones que se realizan dentro del e-commerce entre los vendedores y los clientes.

El dataset de “datos-tipo-cambio-usd” contiene la información sobre el tipo de cambio a dólar desde el 2002 hasta el 2023.

## **Columnas**

Dataset “customers”:

- CustomerID (INT): Clave primaria.
- Name (STRING): Nombres de los clientes.
- Surname (STRING): Apellidos de los clientes.
- Gender (STRING): Género de los clientes.
- Birthdate (DATE): Fecha de nacimiento de los clientes.

Dataset “customers\_purchasing\_power”:

- CustomerID (INT): Clave primaria.
- Purchasing\_Power (STRING): Indicador de poder de compra.

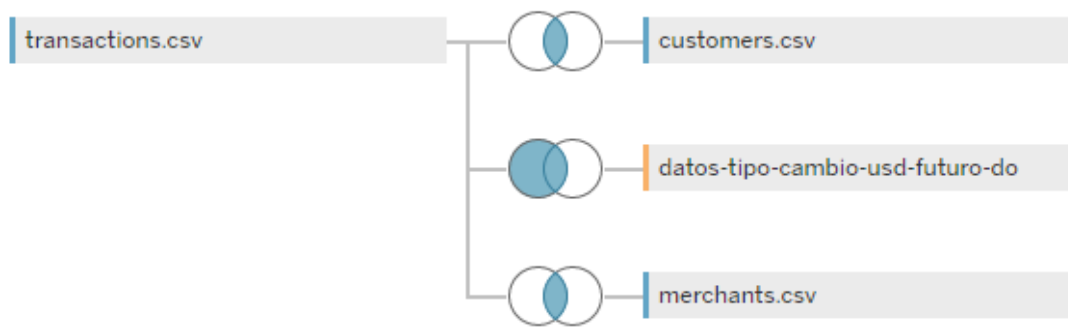
Dataset “merchants”:

- MerchantID (INT): Clave primaria.
- Merchant\_Name (STRING): Nombres de los vendedores.
- Category (STRING): Categoría de los productos que se venden.

Dataset “transactions”:

- TransactionID (INT): Clave primaria.
- Customer\_ID (INT): Clave foránea. ID de los clientes.
- Date (STRING): Fecha de la transacción.
- MerchantID (INT): Clave foránea. ID de los vendedores.
- Transaction\_Amount (INT): Monto de la venta en pesos.

## Modelo de Datos



Se hicieron 3 relaciones físicas. Por un lado, se relacionó, a través de inner join, la tabla de “transactions” con “customers” utilizando la clave foránea “Customer ID” y con “merchants” utilizando la clave foránea “MerchantID”. Por otro, se hizo un left join con “datos-tipo-cambio-usd” usando la columna “Date”.



Se hizo una relación lógica con “customers\_purchasing\_power”.

## Preprocesamiento

- No se encontraron datos incorrectos o valores faltantes.
- En la columna de “Gender” hay valores nulos. Se utiliza un campo calculado para crear una nueva columna llamada “\_Gender”, en base a “Gender”, en la cual se reemplazan los NULL con “Sin especificar”.
- No hay filas duplicadas.
- Se eliminan las columnas “MerchantID” y “Customer ID1” producto de la relación entre las tablas, ya que están duplicadas.
- Se cambia el nombre de “Merchant Name1” por “Merchant Name”.
- En el dataset “datos-tipo-cambio-usd” se eliminaron todas las columnas, excepto por “Indice Tiempo” y “Tipo Cambio Implícito En Adrs”.
- En Excel, se reemplazó los registros de la columna “Tipo Cambio Implícito En Adrs” que tenían punto por una coma, para que Tableau los reconozca como valor numérico.
- En el dataset “transactions” se filtraron todos los registros cuya fecha sea anterior al “01/10/2023”, por dos motivos:
  - En primer lugar, las ventas de octubre solo llegan hasta el día 14, siendo este faltante de información el causante de un

análisis incorrecto de los datos, ya que se están trabajando TODAS las ventas de TODO el mes.

- En segundo lugar, no se dispone del cambio a dólar en las fechas de octubre, por lo cual, no es posible analizar estas ventas en un contexto real.

## **Métricas**

Las métricas creadas para este análisis son las siguientes:

- Cantidad de Ventas:

$\text{Cantidad de ventas} = \text{COUNT}([\text{Transaction ID}])$

- Promedio de Ventas:

$\text{Prom\_Ven\_Men} = \text{COUNT}([\text{Transaction ID}]) / 9$

- Promedio de Monto (ARS):

$\text{Prom\_Mnt\_Men} = \text{SUM}([\text{Transaction Amount}]) / 9$

- Ventas (USD):

$[\text{Transaction Amount}] / [\text{Tipo Cambio Implicito En Adrs}]$

- Promedio de monto (USD):

$\text{Prom\_Mnt(USD)\_Men} = \text{SUM}([\text{Ventas (USD)}]) / 9$

## **Visualización**

Con respecto a la visualización, lo primero que tenemos es una página principal con un menú para ir a los distintos tableros: vendedores – clientes – ventas. Esta división fue hecha para agrupar a los gráficos en correspondencia a la temática que explican.

En la parte de “Vendedores”, utilizan gráficos de barras para visualizar los mayores vendedores y una línea de tiempo que muestra el crecimiento y decrecimiento, junto a sus distintos máximos y mínimos. Se representa tanto en cantidad como en monto (ARS y USD).

En la parte de “Clientes”, se usan gráficos de barras para ver cuales son los clientes que más compras y cuales los que menos. También, se usa una tabla en la que se visualiza una matriz RFM, en la cual se puede ver la distribución de los clientes más valiosos. Y, por último, con un gráfico de torta se muestra el porcentaje de clientes según su poder de compra.

En la parte de “Ventas”, se muestran gráficos de barras que se representa los ingresos totales de cada categoría en pesos y en dólares, y la cantidad de ventas segmentado por poder de compra para cada categoría.

## **Conclusiones**

Los insights que se encontraron después de hecho el análisis son los siguientes:

- Existió un crecimiento en la cantidad de ventas y en las ganancias en pesos, pero, al analizarlo en dólares, se ve un decrecimiento.
- La mayor parte de las transacciones son hechas por clientes con poco poder de compra, mientras que los que tienen alto poder de compra son la menor proporción.
- Hay una gran cantidad de clientes, pero muchos de ellos tienen una frecuencia y recencia en la aplicación muy baja.
- La categoría “Travel” es de las que tiene mayor monto mientras que “Restaurant” es de las que menos.
- La cantidad de ventas por categoría esta distribuida de forma bastante uniforme siendo “Electronics” la que más vende y “Travel” la que menos.

Como conclusión, se puede decir que, durante el periodo de enero a septiembre, aunque hubo un crecimiento en lo que respecta a las ventas, esto solo es cierto si se habla de cantidad de ventas, ya que las ganancias en pesos crecen, pero si se ven en un término real como pueden ser los dólares, hay un decrecimiento bastante notable debido al contexto inflacionario.

Se podría adjudicar el crecimiento en ventas y la gran cantidad de clientes a la novedad del e-commerce, ya que gran parte de estos clientes lo utilizan muy poco o, directamente, no vuelven después del primer uso de la aplicación. Una recomendación sería aplicar una estrategia de marketing para recuperar los clientes a punto de perderse y mantener a los pocos que son más activos.

Conocer a los mayores vendedores y a los clientes que más compran de la aplicación es importante para saber quienes son los usuarios que hay que cuidar para mantener activo el e-commerce.

Por otro lado, saber quienes son los clientes con menos compras también es importante, ya que es una potencial oportunidad del negocio. Se tendrá que hacer un análisis más profundo sobre estos y decidir si vale la pena invertir recursos en recuperarlos o no.

Con respecto al poder de compra de los clientes, nos encontramos lo que probablemente sea la mayor oportunidad para el e-commerce, ya que los clientes con poder de compra "High" y "Mid" abarcan el 40% de total. Llevar a cabo las estrategias necesarias para incentivar el uso de la aplicación a este grupo sería fundamental.

Por último, hay una oferta de categorías bastante uniforme, pero lo que más resalta es lo que sería la otra oportunidad de negocio, la categoría de "Travel". Es la que acumula mayores ingresos por mucha diferencia, pero al mismo tiempo es la que menos se compra. Se recomienda comunicarse con los vendedores y llegar a un acuerdo para promocionar más esta categoría, especialmente en los clientes con alto poder de compra.