

Bootcamp Data Science - Ejercicio S23 - Gerardo Rodríguez

Objetivo: Realizar un análisis RFM (Recency, Frequency, Monetary) para segmentar a los clientes del conjunto de datos "Online Retail.xlsx", identificar diferentes grupos de clientes basados en su comportamiento de compra, y proponer estrategias de marketing personalizadas para cada segmento.

Descripción del Conjunto de Datos: El archivo "Online Retail.xlsx" contiene transacciones de un periodo de tiempo específico para una tienda online. Las columnas relevantes incluyen:

- InvoiceNo (número de factura)
- InvoiceDate (fecha de la factura)
- CustomerID (identificador único del cliente)
- Quantity (cantidad de productos comprados)
- UnitPrice (precio por unidad del producto)

Tareas:

1. Preparación y Limpieza de los Datos:

- Cargar los datos desde el archivo "Online Retail.xlsx".
- Realizar una limpieza inicial de los datos, asegurándose de eliminar transacciones sin CustomerID, filtrar cualquier dato que pueda ser erróneo o no representativo para el análisis RFM (por ejemplo, cantidades negativas o precios cero).

2. Cálculo de Métricas RFM:

- Calcular las métricas RFM para cada cliente:
- Recency (R): El número de días desde la última compra de cada cliente
- Frequency (F): El número total de compras realizadas por cada cliente
- Monetary (M): El gasto total de cada cliente
- Asegúrate de utilizar la fecha más reciente del conjunto de datos como punto de referencia para calcular la métrica Recency (R).

3. Segmentación de Clientes:

- Clasificar a los clientes en segmentos basándose en cuartiles para cada métrica RFM.
- Crear un puntaje RFM combinando los cuartiles de cada métrica.
- Identificar segmentos de interés, como los mejores clientes (por ejemplo, aquellos con puntaje RFM '444').

4. Análisis de Segmentos y Estrategias de Marketing:

- Realizar un análisis de los segmentos de clientes identificados, destacando las características clave de cada grupo.
- Proponer estrategias de marketing personalizadas para al menos tres segmentos diferentes de clientes, basándose en sus comportamientos de compra identificados a través del análisis RFM.

Entregables:

Un informe que incluya:

- La metodología utilizada para la limpieza, preparación y análisis de los datos

- Los resultados del análisis RFM, incluyendo una descripción de los segmentos de clientes identificados
- Estrategias de marketing recomendadas para cada segmento de clientes identificado, con justificaciones basadas en el análisis RFM
- Anexo con el código en Python utilizado para realizar el análisis, incluyendo comentarios que expliquen las operaciones realizadas en cada paso

1. Metodología utilizada para la limpieza, preparación y análisis de datos.

Filtrado de Datos:

Se filtran las filas donde Quantity y UnitPrice sean mayores que 0 para asegurarse de que los datos sean válidos y positivos.

También se filtran las filas donde CustomerID sea mayor que 0 para asegurar que se tengan datos válidos de clientes.

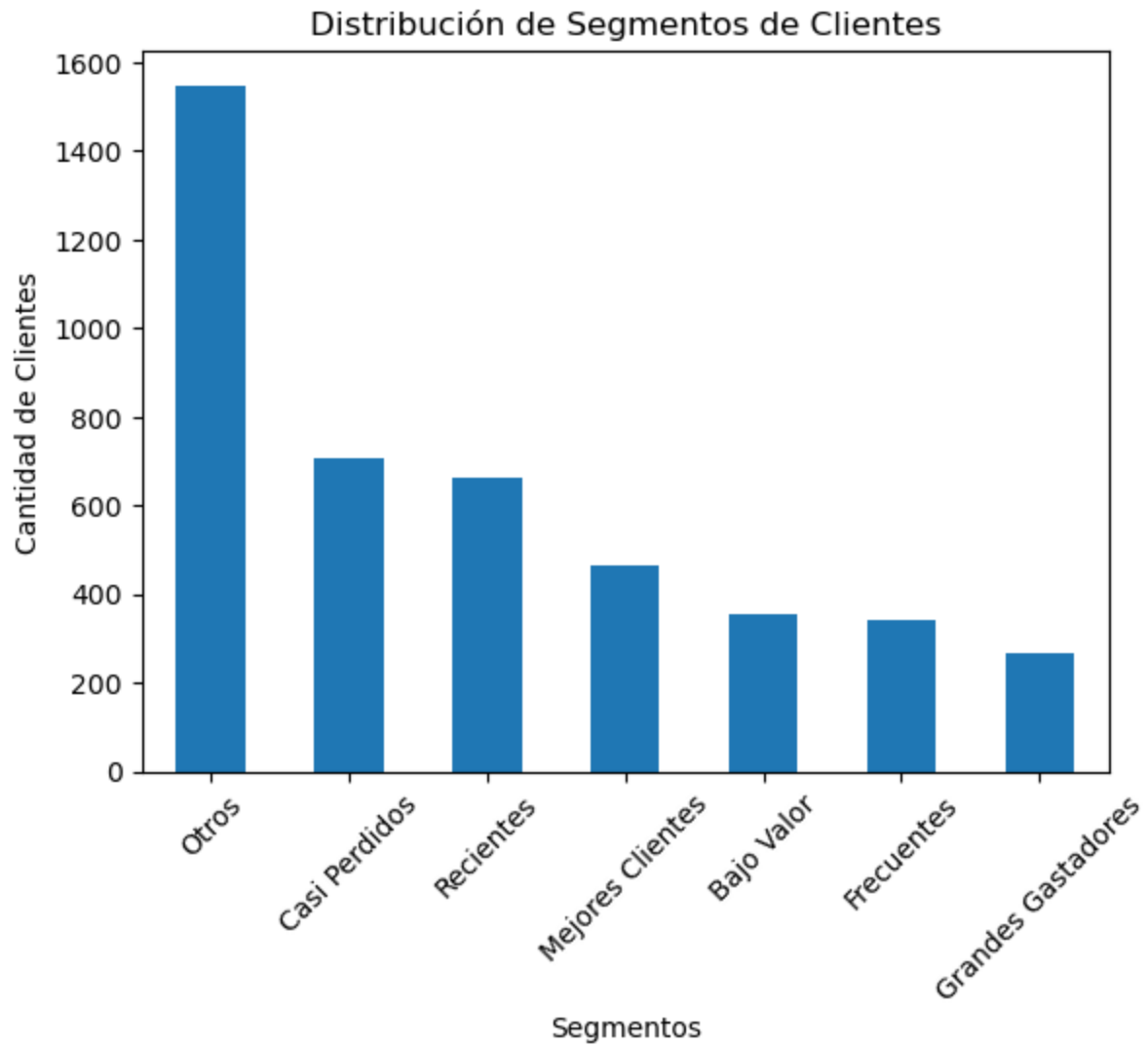
Conversión de Fechas:

Las fechas de las facturas (InvoiceDate) se convierten al formato datetime para facilitar el análisis temporal.

2. Los resultados del análisis RFM, incluyendo una descripción de los segmentos de clientes identificados

El análisis RFM (Recency, Frequency, Monetary) clasifica a los clientes en diferentes segmentos basados en sus comportamientos de compra. Aquí se muestran los resultados del análisis RFM realizado con los datos proporcionados.

	Segmento	Cantidad de Clientes
0	Otros	1548
1	Casi Perdidos	704
2	Recientes	663
3	Mejores Clientes	462
4	Bajo Valor	353
5	Frecuentes	343
6	Grandes Gastadores	265



Para poder identificar los segmentos que más nos dejan dinero los clasificamos por ingresos que nos dejan de mayor a menor.

```

# Crear la tabla de suma de gastos por segmento
monetary_by_segment = RFM.groupby('Segment')['Monetary'].sum().reset_index()
monetary_by_segment.columns = ['Segmento', 'Total Monetario']

# Ordenar la tabla por el total monetario en orden descendente
monetary_by_segment = monetary_by_segment.sort_values(by='Total Monetario', ascending=False)

# Mostrar la tabla
print(monetary_by_segment)

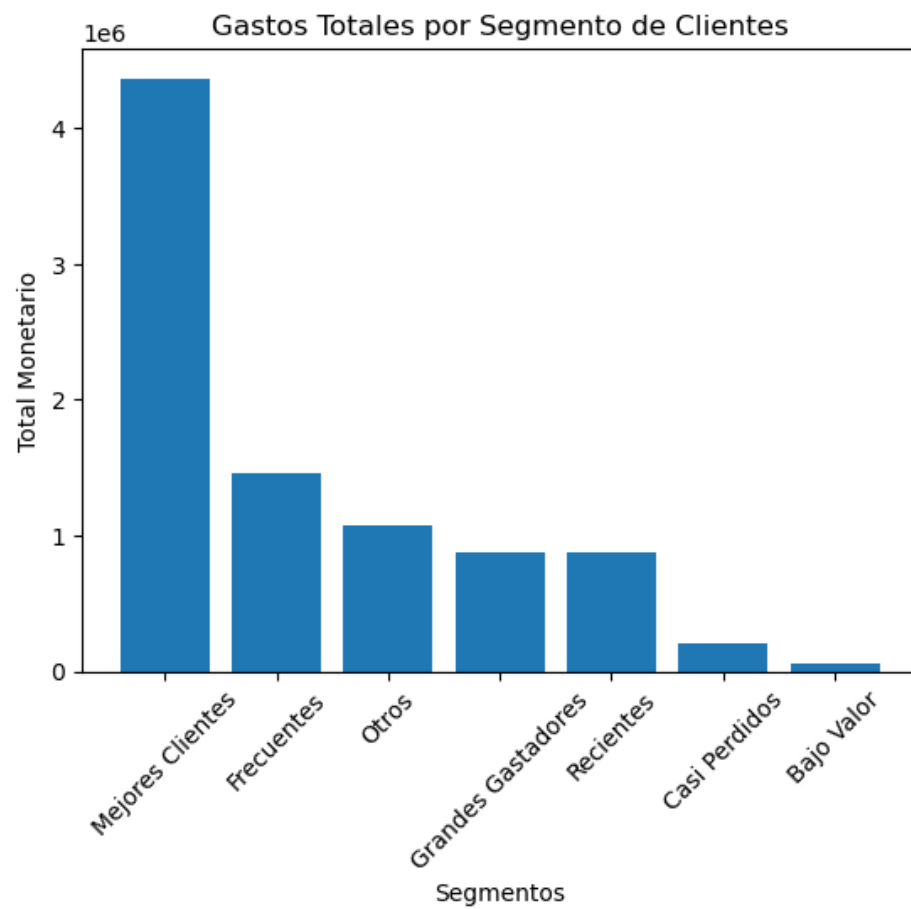
# Crear el gráfico de barras
fig, ax = plt.subplots()
ax.bar(monetary_by_segment['Segmento'], monetary_by_segment['Total Monetario'])
ax.set_xlabel('Segmentos')
ax.set_ylabel('Total Monetario')
ax.set_title('Gastos Totales por Segmento de Clientes')
plt.xticks(rotation=45)
plt.show()

```

Identificamos que nuestra estrategia debería enfocarse de manera constante en conservar y mejorar el segmento de Mejores Clientes y Clientes Frecuentes.

Sin embargo, rubros como Grandes gastados y Recientes pueden ser interesantes para convertirlos al rubro de Mejores clientes y/o Frecuentes.

Segmento	Total Monetario
Mejores Clientes	4369199.090
Frecuentes	1454127.700
Otros	1074028.303
Grandes Gastadores	878180.031
Recientes	870034.430
Casi Perdidos	204204.010
Bajo Valor	61634.340



3. Estrategias de marketing recomendadas para cada segmento de clientes identificado, con justificaciones basadas en el análisis RFM

Aunque brindamos ideas aplicables para cada segmento, recomendamos enfocarnos en los principales grupos: Mejores clientes,

a) Mejores Clientes (444):

Descripción: Clientes con las puntuaciones más altas en recencia, frecuencia y valor monetario.

Estrategia: Mantener estos clientes satisfechos con ofertas especiales, programas de fidelización y excelente servicio al cliente.

b) Recientes:

Descripción: Clientes que han comprado recientemente.

Estrategia: Aprovechar su reciente actividad con promociones adicionales para fomentar la repetición de compras.

c) Frecuentes:

Descripción: Clientes que compran con alta frecuencia.

Estrategia: Mantener su lealtad con descuentos por volumen y ofertas exclusivas.

d) Grandes Gastadores:

Descripción: Clientes que gastan grandes cantidades de dinero.

Estrategia: Ofrecer productos premium y servicios personalizados para mantener su interés.

e) En Riesgo:

Descripción: Clientes que solían ser activos pero no han comprado recientemente.

Estrategia: Enviar campañas de reactivación con descuentos o recordatorios de productos que les puedan interesar.

f) Casi Perdidos:

Descripción: Clientes que muestran señales de abandono.

Estrategia: Intentar reactivarlos con ofertas especiales o encuestas para entender mejor sus necesidades.

g) Bajo Valor:

Descripción: Clientes con baja frecuencia y bajo valor monetario.

Estrategia: Evaluar si vale la pena invertir en su retención o enfocarse en otros segmentos.

h) Otros:

Descripción: Clientes que no caen específicamente en las categorías anteriores.

Estrategia: Analizar individualmente para determinar posibles acciones.