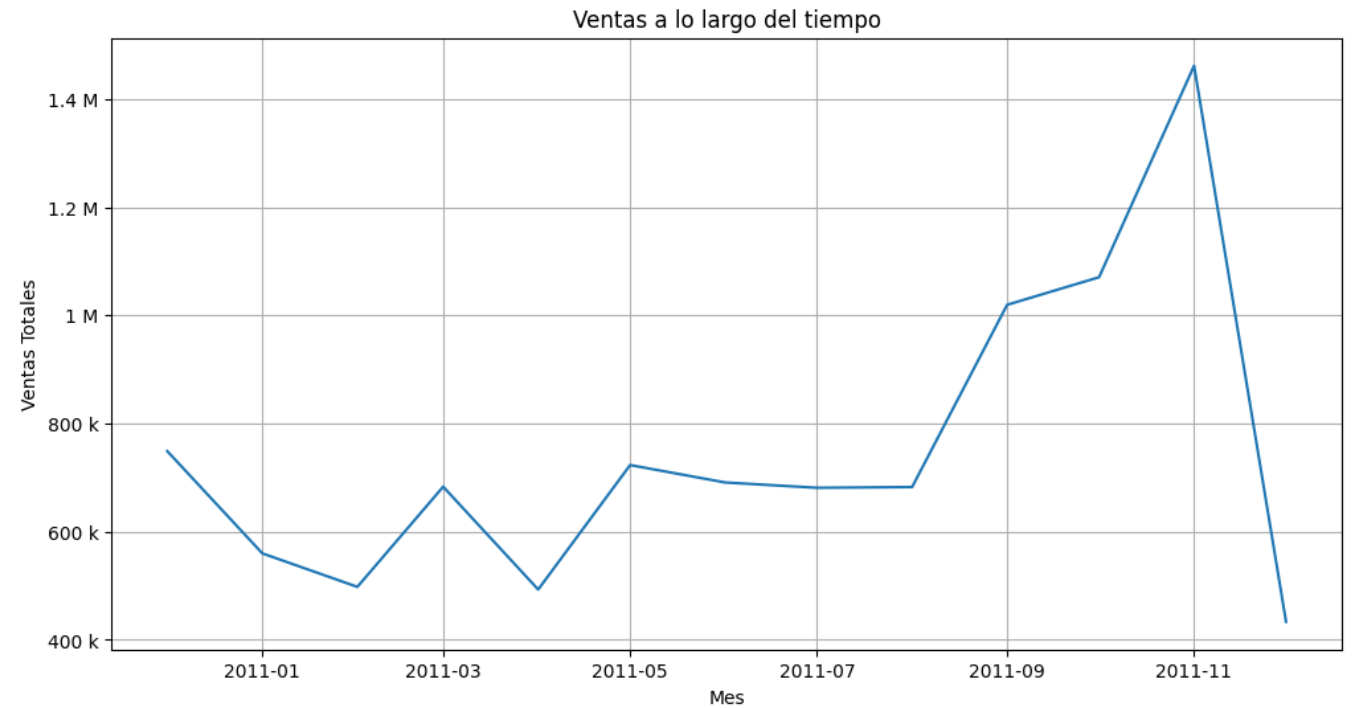


Segmentación de clientes y KPIs

Carlos Isaac



Objetivo del Análisis

Descripción de la Base de Datos:

- La base de datos contiene transacciones de clientes, registrando detalles específicos de cada compra. Cada entrada en la base de datos representa un ítem comprado, incluyendo información como el número de factura (InvoiceNo), el código del producto (StockCode), la descripción del producto (Description), la cantidad comprada (Quantity), la fecha de la factura (InvoiceDate), el precio unitario del producto (UnitPrice), el ID del cliente (CustomerID), el país del cliente (Country) y las ventas totales para esa entrada (TotalSales).
- Estos datos proporcionan una visión detallada de las compras realizadas por los clientes a lo largo del tiempo, permitiendo analizar patrones de compra, segmentación de clientes, y métricas clave de rendimiento.

Propósito del Análisis:

- El objetivo principal del análisis es comprender mejor el comportamiento de compra de los clientes y obtener insights que puedan ayudar a la empresa a tomar decisiones informadas. Mediante la segmentación de clientes y el cálculo de KPIs, se busca identificar oportunidades de mejora, áreas de crecimiento y estrategias específicas que puedan ser implementadas para mejorar la relación con el cliente y aumentar las ventas. En particular, el análisis se centra en identificar segmentos de alto valor, entender las tendencias de ventas y evaluar el rendimiento general basado en métricas clave.

Verificación de la Calidad de Datos

- Limpieza: Eliminación de entradas con datos faltantes y manejo de valores atípicos.
- Transformación: Estandarización del formato de fecha y creación de la columna InvoiceMonth.
- Enriquecimiento: Cálculo de TotalSales por transacción.
- Deduplicación: Verificación y eliminación de entradas duplicadas.
- Normalización: Preparación de datos para técnicas de análisis.

Calidad de Datos:

- Consistencia: Formato uniforme en columnas.
- Validez: Reglas para valores válidos (ej. Quantity positiva).
- Integridad: Coherencia entre columnas (ej. $\text{TotalSales} = \text{Quantity} \times \text{UnitPrice}$).
- Completeness: No hay valores faltantes en columnas clave.



Principales Indicadores de Rendimiento



Ventas Totales: \$9,747,747



Número Total de Clientes: 4,372



Venta Promedio por Transacción: \$398.75

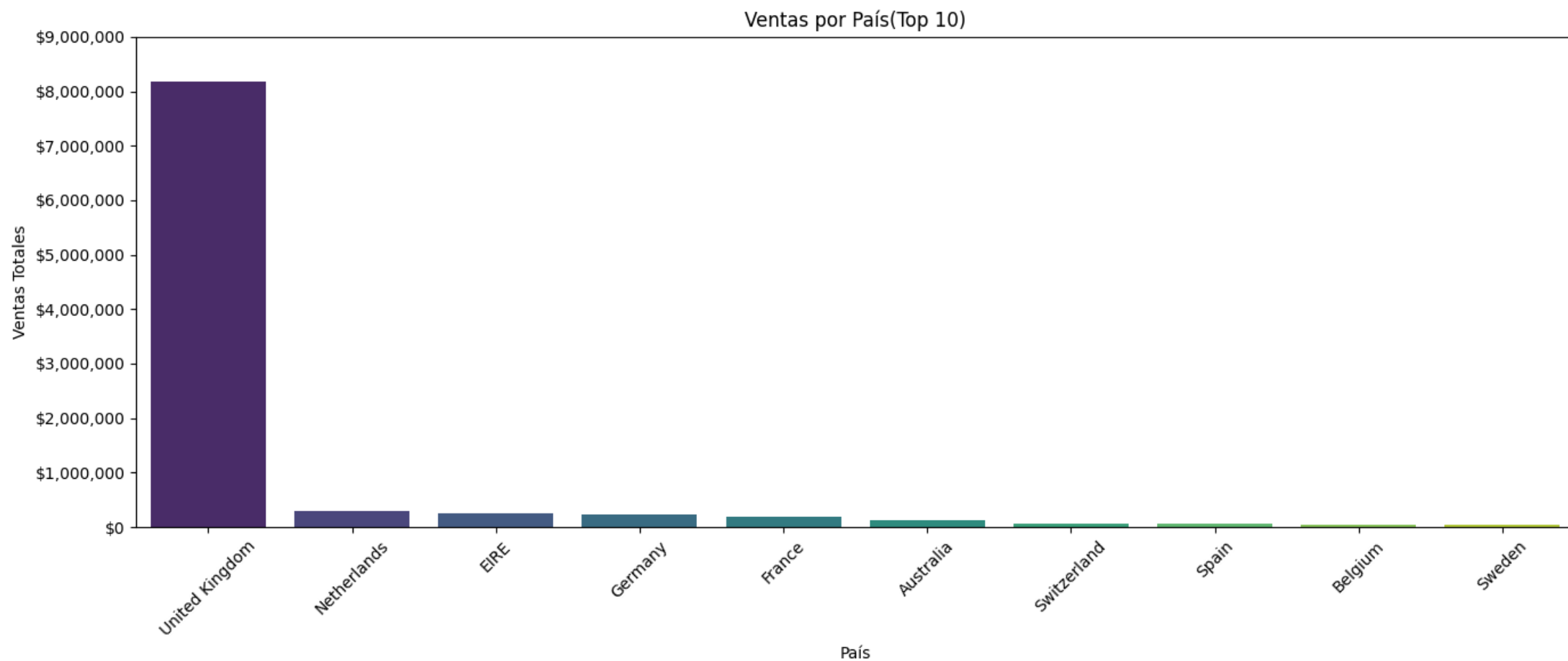


Precio Promedio de Venta por Producto: \$4.62

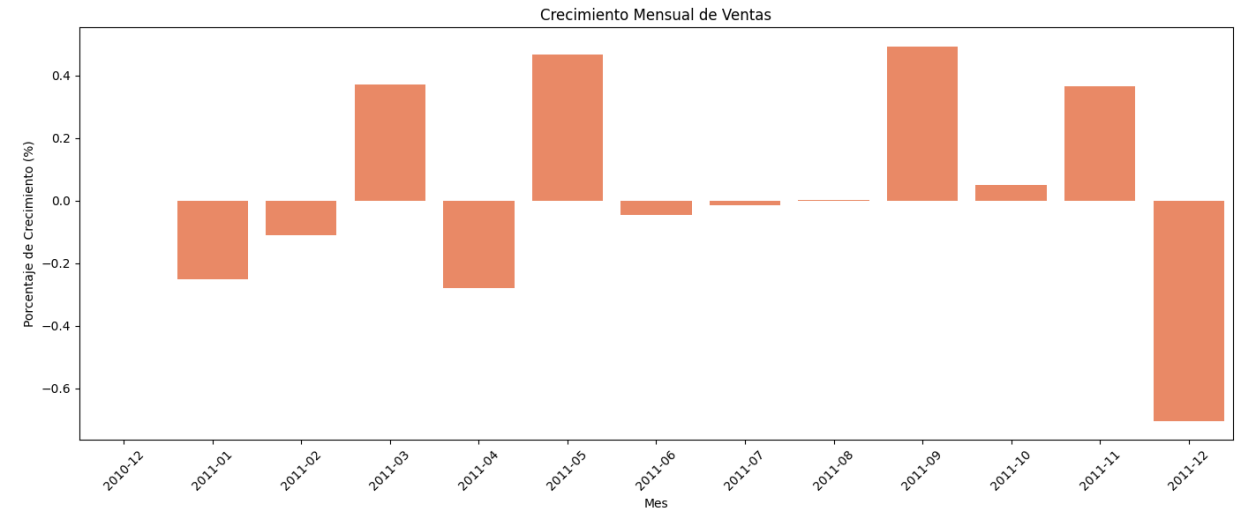
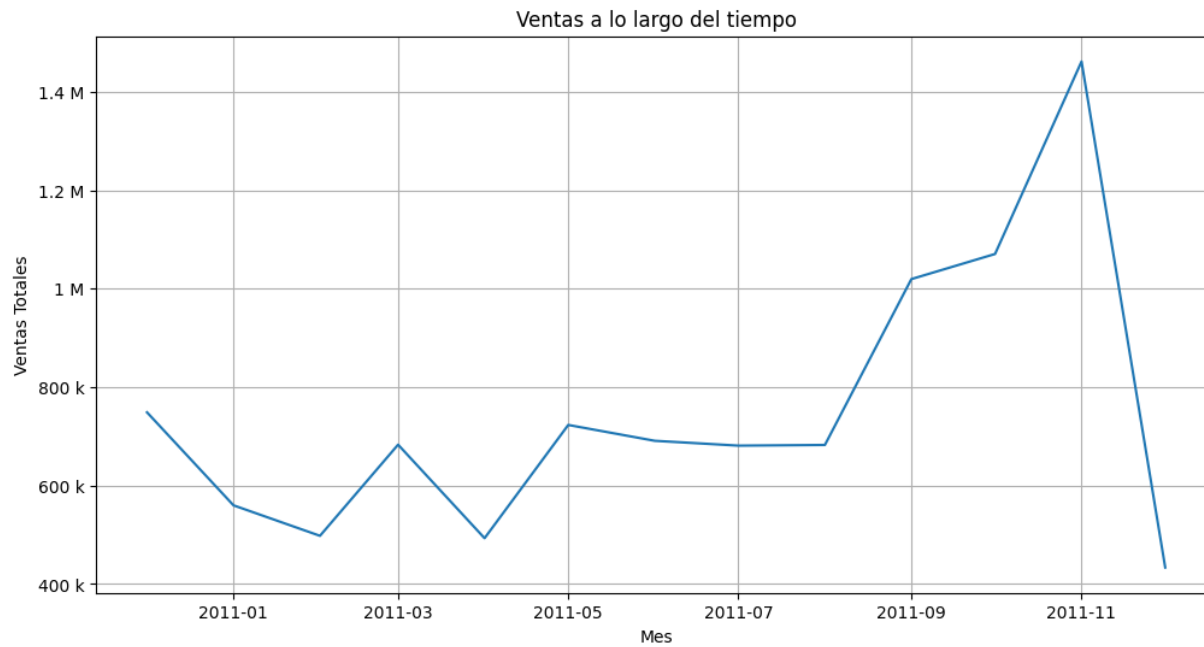


Número Total de Productos Vendidos: 5,190,059

Ventas por País



Crecimiento Mensual en Ventas



Segmentación de Clientes con RFM

- **RFM** es un método utilizado para analizar el valor del cliente. Se basa en tres criterios:
- **Recency (R)**: ¿Cuándo fue la última vez que un cliente hizo una compra? Los clientes que han comprado recientemente son más propensos a comprar nuevamente en comparación con aquellos que lo hicieron hace tiempo.
- **Frequency (F)**: ¿Con qué frecuencia un cliente realiza una compra? Si un cliente compra con frecuencia, es probable que siga haciéndolo en el futuro. Estos clientes suelen ser más valiosos que aquellos que compran con menos frecuencia.
- **Monetary Value (M)**: ¿Cuánto dinero gasta un cliente? Los clientes que gastan más son más propensos a continuar comprando en comparación con aquellos que gastan menos.



Conclusión análisis RFM

Segmento Oro (1386 clientes):

- Recency: 63 (compra reciente).
- Frequency: 12.73 (compras frecuentes).
- Monetary Value: Promedio de \$5,845, rango de -\$71.46 a \$1,448,167.36.
- Observación: Clientes valiosos que compran con frecuencia y gastan mucho.

Segmento Plata (2563 clientes):

- Recency: 119.5 (última compra hace tiempo).
- Frequency: 2.38 (frecuencia moderada).
- Monetary Value: Promedio de \$601, rango de -\$4,287.63 a \$12,393.70.
- Observación: Clientes con frecuencia moderada y gasto moderado.

Segmento Bronce (423 clientes):

- Recency: 20 (compra bastante reciente).
- Frequency: 1.64 (pocas compras).
- Monetary Value: Promedio de \$250.
- Observación: Clientes que compran menos frecuentemente y gastan menos.

Resumen:

El Segmento Oro contiene a los clientes más valiosos con compras frecuentes y alto gasto. El Segmento Bronce se caracteriza por clientes que compran menos y gastan menos. El Segmento Plata ocupa un punto intermedio en términos de frecuencia y gasto.

Número de clientes en cada segmento RFM

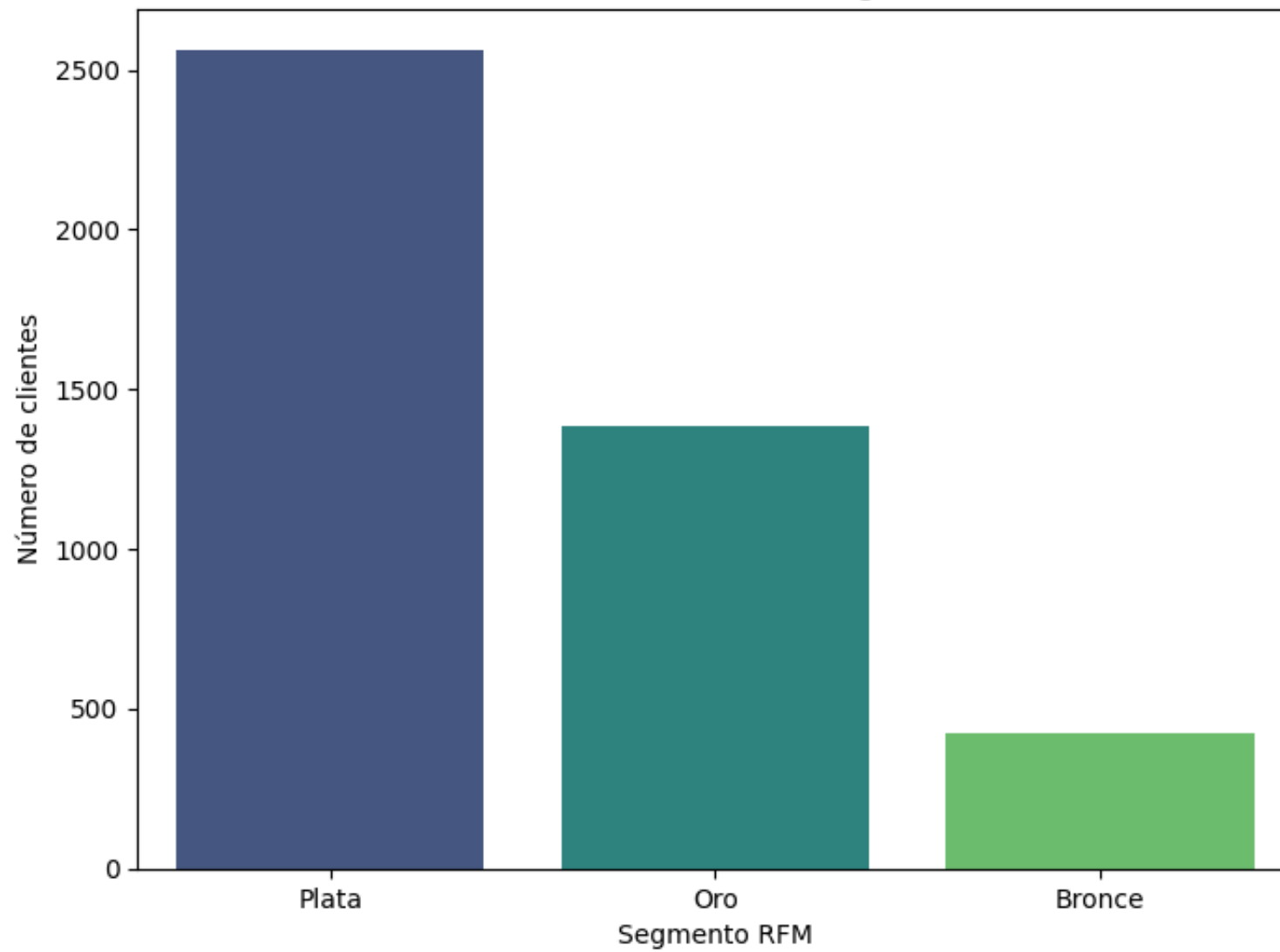
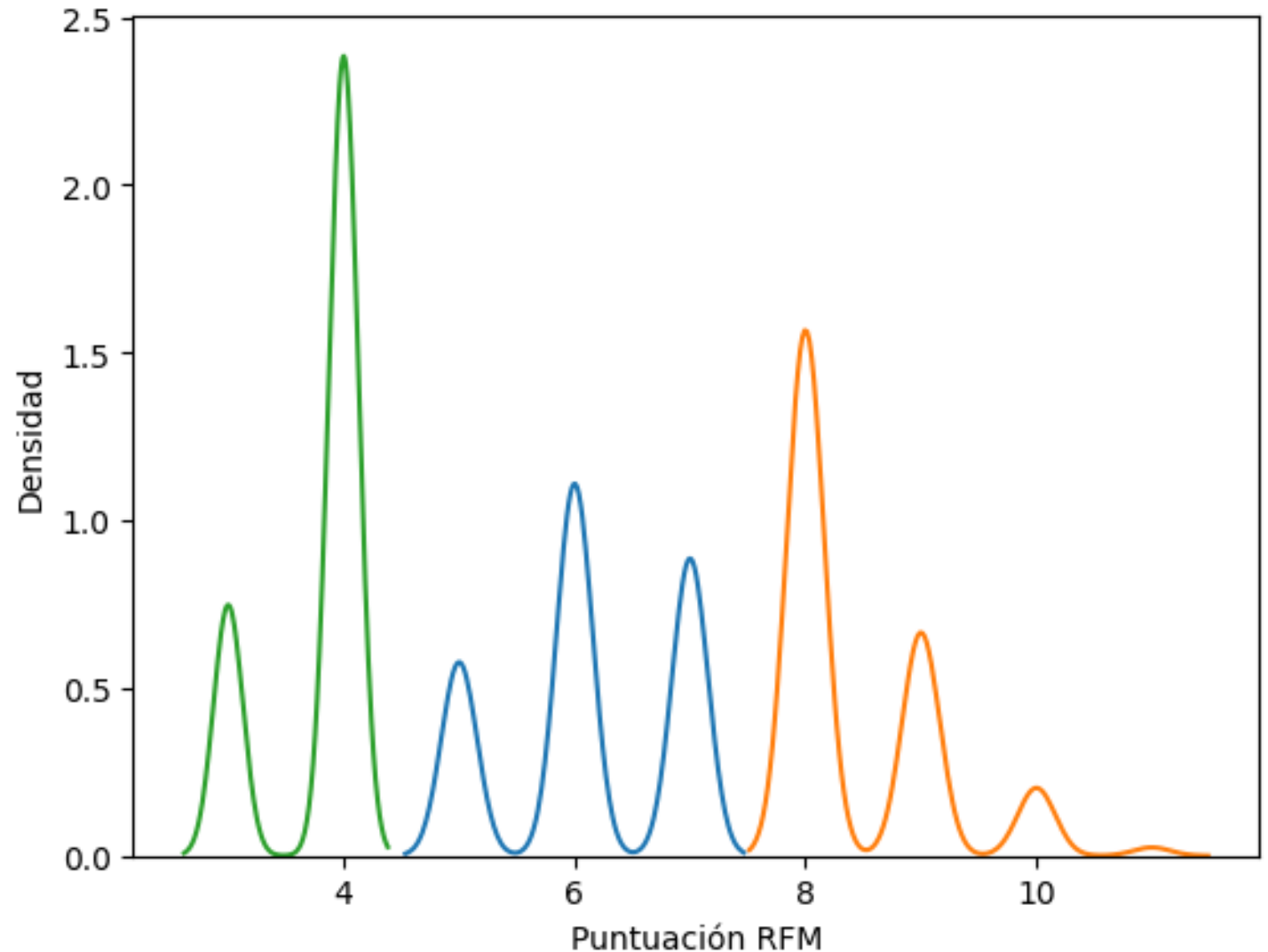


Gráfico de Distribución de RFM Score por Segmento:

- Este gráfico muestra cómo se distribuyen los valores de "RFM Score" para cada uno de los tres segmentos: Oro, Plata y Bronce.
- Eje X (Horizontal): Representa el "RFM Score", que es una combinación del recency, frequency y monetary value de cada cliente.
- Eje Y (Vertical): Indica la densidad o la cantidad de clientes en cada punto del "RFM Score".
- Cada línea (Oro, Plata, Bronce) nos muestra la tendencia de sus respectivos clientes. Por ejemplo, si la línea de Oro tiene un pico en un valor alto del eje X, significa que muchos clientes "Oro" tienen un "RFM Score" alto.



Segmentación mediante K-means

Especial (1 cliente):

- Última compra: Muy reciente.
- Frecuencia de compra: 2259 veces.
- Gasto total: \$1,448,167.36.
- Observación: Cliente extremadamente valioso con comportamiento de compra excepcional.

Oro (2920 clientes):

- Última compra: Hace 31 días.
- Frecuencia de compra: 6.54 veces.
- Gasto total: \$2,557.20.
- Observación: Clientes muy valiosos que compran con frecuencia y gastan significativamente.

Plata (817 clientes):

- Última compra: Hace 152 días.
- Frecuencia de compra: 2.55 veces.
- Gasto total: \$704.83.
- Observación: Clientes valiosos con una frecuencia y gasto moderados.

Bronce (634 clientes):

- Última compra: Hace 295 días.
- Frecuencia de compra: 1.56 veces.
- Gasto total: \$404.90.
- Observación: Clientes con la menor frecuencia y gasto.



Conclusión

- La segmentación RFM nos permitió clasificar rápidamente a nuestros clientes en categorías fácilmente comprensibles (Oro, Plata, Bronce) basadas en su comportamiento de compra reciente, su frecuencia de compra y su gasto total. Esto nos proporciona una visión general del valor de los diferentes segmentos de clientes para el negocio y puede informar estrategias de marketing y retención de clientes de alto nivel.
- Por otro lado, la segmentación K-means nos permitió detectar patrones más sutiles y específicos en los datos. Nos proporcionó un grupo adicional (Especial) para un cliente muy valioso que gastó una cantidad excepcionalmente alta. Aunque este grupo solo contiene un cliente, es crucial para el negocio debido a su alto valor. Este tipo de información detallada puede ser útil para estrategias de marketing y servicio al cliente más personalizadas y dirigidas.