

# PRUEBA TÉCNICA SUMMAN

ISABELA TOLEDO BALCÁZAR

MARZO 2024

# AGENDA

- CONTEXTO DE LA PRUEBA
- PARTE 1: ANÁLISIS EXPLORATORIO DE DATOS (EDA)
- PARTE 2: APLICACIÓN DE MODELO DE SEGMENTACIÓN DE CLIENTES
- ANEXOS
- BIBLIOGRAFÍA

# 1. CONTEXTO DE LA PRUEBA

- Se proporciona un conjunto de datos de ventas de un negocio.
- **Periodo de ventas:** 2020-2023
- **Total registros:** 11.744
- **Tablas entregadas:**
  - **Sales:** incluye registros de ventas
  - **Product:** incluye información de los productos del negocio
  - **Customer:** incluye información de los clientes
  - **Regional\_manager:** incluye información de los vendedores por región
  - **Returns:** incluye información de los pedidos correspondientes a devoluciones del cliente

## **2. PARTE 1: ANÁLISIS EXPLORATORIO DE DATOS (EDA)**

### **INSTRUCCIÓN:**

**“Realizar análisis exploratorio de los datos entregados. Los resultados deben ser presentados a través de una herramienta de visualización”**

## 2. PARTE 1: ANÁLISIS EXPLORATORIO DE DATOS (EDA)

- **EDA:** es una técnica que permite obtener una **comprensión más profunda** de los conjuntos de datos e implica explorar los datos para descubrir **patrones, relaciones y anomalías**.

El objetivo del EDA no es confirmar hipótesis sino que se centra **en generar preguntas** y sus posibles direcciones para las investigaciones futuras. [1]

- **Herramientas utilizadas:** Python y Power BI

# 2. PARTE 1: ANÁLISIS EXPLORATORIO DE DATOS (EDA)

## PROCESO GENERAL:

- **IMPORTACIÓN DE DATOS**

### **2. LIMPIEZA**

- Exploración inicial
- Remover duplicados
- Gestionar los datos faltantes
- Corregir datos erróneos
- Dar formato adecuado a las columnas
- Manejo de outliers
- Análisis inicial con estadística descriptiva

### **3. EXPORTACIÓN DE DATOS LIMPIOS**

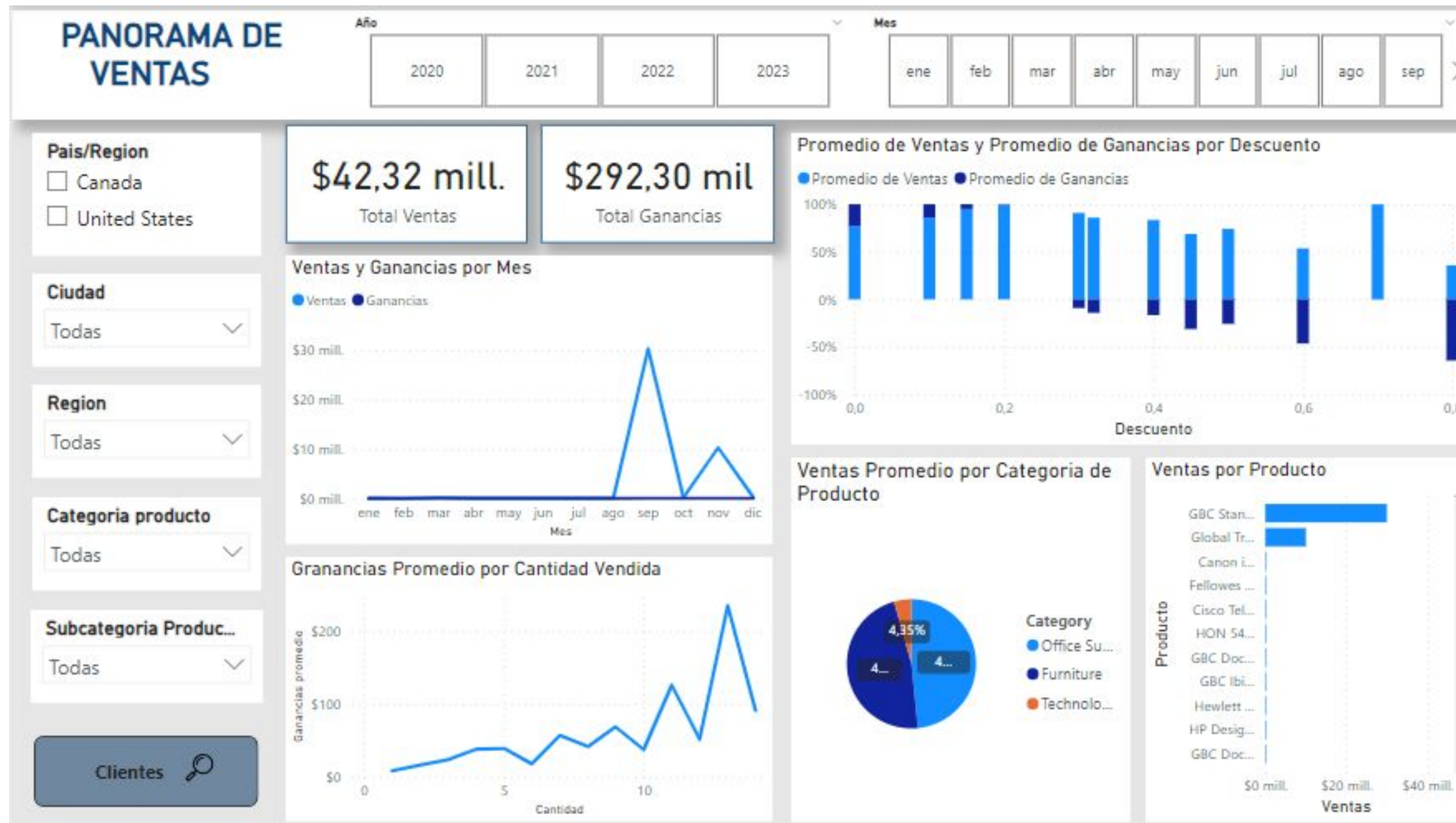
### **4. GENERACIÓN DE PREGUNTAS**

### **5. GENERACIÓN DE VISUALIZACIONES**

## 2. PARTE 1: ANÁLISIS EXPLORATORIO DE DATOS (EDA)

- ¿Cómo varían las ventas y las ganancias a lo largo de los meses y años?
- ¿Los descuentos tienden a aumentar las ventas o las ganancias, o tienen un impacto negativo?
- ¿Qué productos o categorías de productos generan más ventas?
- ¿Cómo afecta la cantidad de productos vendidos a la ganancia?

## 2. PARTE 1: ANÁLISIS EXPLORATORIO DE DATOS (EDA)





## 2. PARTE 1: ANÁLISIS EXPLORATORIO DE DATOS (EDA)

¿Cómo varían las ventas y las ganancias a lo largo de los meses y años?

- Se observan **picos significativos** en las ventas en ciertos meses del año (como septiembre y octubre), lo que podría indicar **temporadas de alta demanda** o **posiblemente grandes ventas individuales**. Estos picos no siempre coinciden **con aumentos proporcionales en las ganancias**, lo que sugiere que algunas de estas ventas de alto volumen podrían haber involucrado descuentos significativos o productos con márgenes de ganancia más bajos.
- Las **ganancias**, aunque generalmente siguen la tendencia de las ventas, **muestran una variabilidad considerable** de un mes a otro. Esto podría deberse a la mezcla de productos vendidos, los costos variables asociados con ellos, o la estrategia de precios y descuentos.

## 2. PARTE 1: ANÁLISIS EXPLORATORIO DE DATOS (EDA)

¿Los descuentos tienden a aumentar las ventas o las ganancias, o tienen un impacto negativo?

- **Ventas promedio por nivel de descuento:** Las ventas promedio **tienden a ser más altas para ciertos niveles de descuento**, lo que sugiere que los descuentos pueden incentivar efectivamente un aumento en las ventas. Sin embargo, **no hay una correlación directa** entre el nivel de descuento y las ventas promedio.
- **Ganancia promedio por nivel de descuento:** La relación entre el nivel de descuento y la ganancia promedio **parece ser negativa en general**, lo cual puede esperarse, ya que **mayores descuentos** generalmente **reducen el margen de ganancia** por venta. También hay niveles de descuento específicos donde las **ganancias promedio aumentan** ligeramente, lo que puede indicar una **gestión efectiva** de los descuentos para ciertos productos o categorías.
- Aunque los descuentos pueden ser una estrategia efectiva para impulsar las ventas, su impacto en las ganancias es complejo. Los descuentos deben ser estratégicamente planeados y posiblemente combinados con otras tácticas de venta para asegurar que no solo aumenten las ventas sino también las ganancias.

## 2. PARTE 1: ANÁLISIS EXPLORATORIO DE DATOS (EDA)

¿Qué productos o categorías de productos generan más ventas?

- En general, los **5 productos más vendidos** son: GBC Binding System Combs, Global Troy Low-Back Titler, Canon Advanced Copier, Fellowes Electric Punch Comb, Cisco videoconferencing Unit.
- En general, la **categoría de productos más vendida** es la de **artículos de oficina** (48,45%), seguido de **Muebles** (47,2%) y la menos vendida es la de tecnología.
- Cabe resaltar que aunque un producto sea el más vendido, no necesariamente quiere decir que sea el que más ganancias produzca.
- En analizar las ventas para cada uno de los años, se observa que tanto los productos más vendido como las categorías más vendidas, varían a lo largo del tiempo.

## 2. PARTE 1: ANÁLISIS EXPLORATORIO DE DATOS (EDA)

¿Cómo afecta la cantidad de productos vendidos a la ganancia?

- Hay una **variabilidad notable** en la ganancia promedio a medida que cambia la cantidad vendida. Esto sugiere que **la relación entre cantidad y ganancia no es lineal** y puede estar influenciada por **varios factores** como el tipo de producto, el precio unitario, los costos y las estrategias de descuento.
- Existe un **pico prominente** en ganancias promedio cuando la **cantidad vendida es alta** (alrededor de 13 unidades). Esto podría indicar que las ventas en mayores cantidades podrían estar sujetas a mejores márgenes de venta que aumentan la rentabilidad.

# **3. PARTE 2: MODELO DE SEGMENTACIÓN DE CLIENTES**

**INSTRUCCIÓN:**

**“implementar un modelo de segmentación de clientes considerando la mayor cantidad de variables posibles”**

# 3. PARTE 2: MODELO DE SEGMENTACIÓN DE CLIENTES

## MODELO RFM: RECENCIA, FRECUENCIA Y VALOR MONETARIO

Es una **técnica de segmentación de clientes** basada en el **comportamiento de compra pasado**, se utiliza para ayudar a las empresas a clasificar a los clientes en función de la **recencia, la frecuencia y el valor monetario** de una transacción.

- **Recencia:** mide el tiempo desde la última compra de un cliente. Menos tiempo indica una mayor interacción y compromiso con la marca.
- **Frecuencia:** Número de veces que cada cliente realizó una compra.
- **Valor Monetario:** mide el total de dinero gastado por un cliente.

# 3. PARTE 2: MODELO DE SEGMENTACIÓN DE CLIENTES

## MODELO RFM: RECENCIA, FRECUENCIA Y VALOR MONETARIO

### JUSTIFICACIÓN:

- Es ampliamente utilizado por su **fácil compresión y aplicabilidad**.
- Disponibilidad de todos los **datos de transacción** necesarios para realizar el modelo.
- Las métricas ofrecen una visión profunda del **compromiso del cliente** y su **valor para el negocio**.
- Ayuda a **identificar a los clientes más valiosos**.
- Ayuda a **personalizar la comunicación, ofertas y promociones** de diferentes grupos de clientes.
- Ayuda **mejorar la experiencia del cliente** y utilizar los **recursos de marketing de forma eficaz**.

# 3. PARTE 2: MODELO DE SEGMENTACIÓN DE CLIENTES

## MODELO RFM: RECENCIA, FRECUENCIA Y VALOR MONETARIO

### IMPLEMENTACIÓN:

- Calcular las métricas RFM.
- Asignar una puntuación RFM a cada una de las métricas.
- Crear segmentos de clientes combinando las puntuaciones RFM.
- Añadir variables demográficas para un análisis más profundo de los segmentos de los clientes.



# 3. PARTE 2: MODELO DE SEGMENTACIÓN DE CLIENTES

## MODELO RFM: RECENCIA, FRECUENCIA Y VALOR MONETARIO

	Recency	Frequency	Monetary	R_Score	F_Score	M_Score	RFM_Score	Segment
Customer_Name								
Aaron Bergman	415	6	886.156	1	1	1	111	No Recurrente
Aaron Hawkins	12	11	1744.700	4	2	2	422	Potencial
Aaron Smayling	88	10	3050.692	2	2	3	223	No Recurrente
Adam Bellavance	54	21	7957.180	3	4	4	344	Leal
Adam Hart	34	20	3250.337	3	4	3	343	Leal
Adam Shillingsburg	28	25	3255.310	4	4	3	443	VIP

**Cientes VIP (Valiosos):** RFM Score de "444" o "443", indica que son clientes recientes, frecuentes y con alto gasto.

**Cientes Leales:** Alta puntuación en Frecuencia, como "X4X".

**Cientes Potenciales:** Clientes con buenos valores de Recencia y Frecuencia, pero no necesariamente altos en Valor monetario.

**Cientes No Recurrentes:** Bajos en Frecuencia y posiblemente variados en Recencia y Valor Monetario.

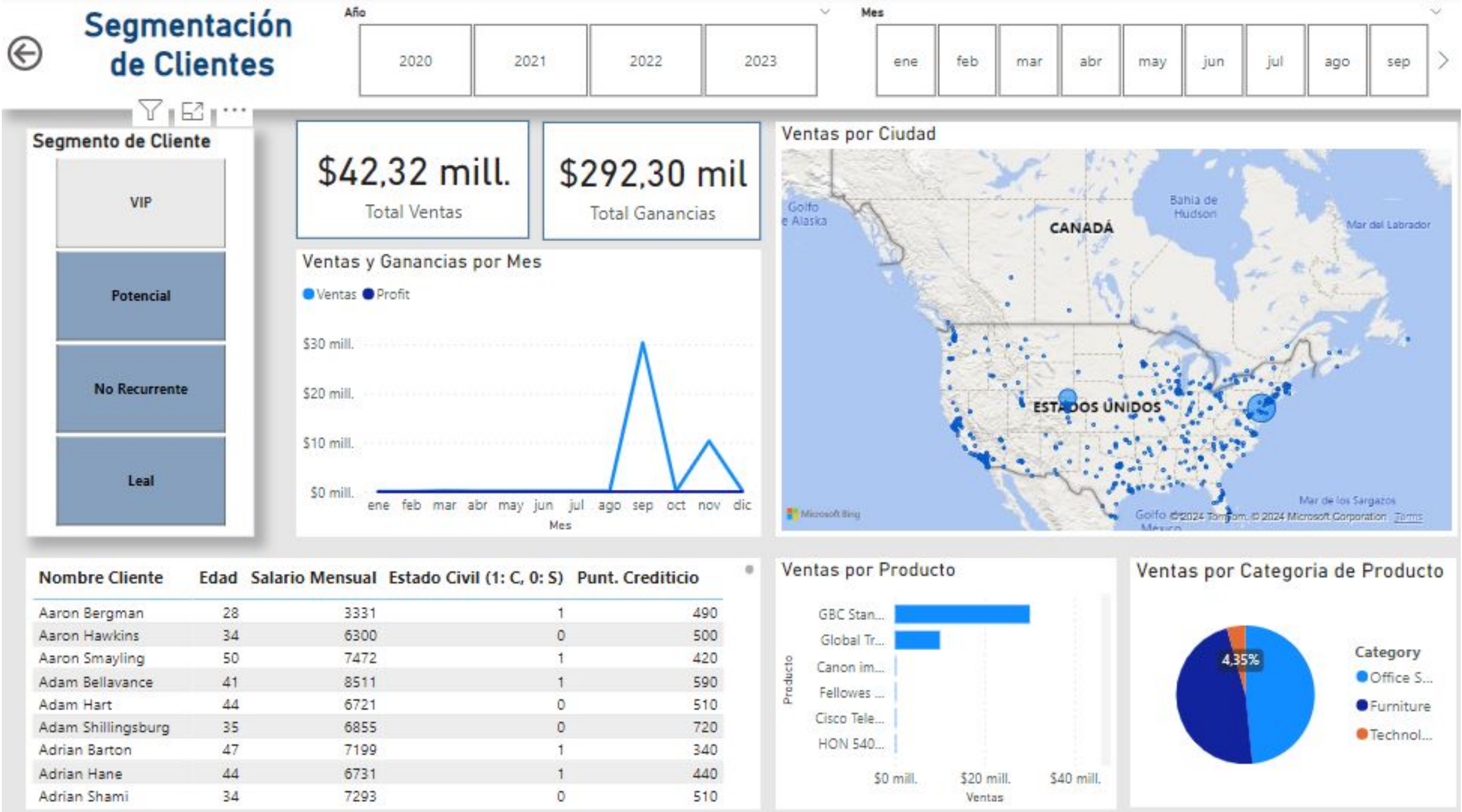
# 3. PARTE 2: MODELO DE SEGMENTACIÓN DE CLIENTES

## MODELO RFM: RECENCIA, FRECUENCIA Y VALOR MONETARIO

### INSIGHTS:

- La empresa tiene: 401 clientes "No recurrentes", 210 "clientes potenciales", 131 clientes "leales" y 53 clientes "VIP".
- **Edad:** La edad promedio en los segmentos es bastante uniforme alrededor de los 38 años.
- **Salario Mensual:** Los clientes leales y los VIP tienen los ingresos mensuales promedio más altos, lo que puede correlacionarse con su capacidad para realizar compras más frecuentes o más costosas.
- **Estado Civil:** La proporción de clientes casados en comparación con los solteros varía ligeramente, siendo más baja en el segmento VIP. Sugiriendo una posible libertad financiera mayor para los clientes solteros.
- **Puntaje de Crédito:** Los VIP tienen el puntaje de crédito promedio más alto, Lo que puede sugerir una mejor salud financiera, y los clientes No Recurrentes tienen un puntaje de crédito ligeramente inferior al de otros segmentos.

# 3. PARTE 2: MODELO DE SEGMENTACIÓN DE CLIENTES



## 4. ANEXOS:

LINK A REPOSITORIO EN GITHUB:

[https://github.com/Isabelatoledob/Prueba\\_Tecnica\\_summan](https://github.com/Isabelatoledob/Prueba_Tecnica_summan)

## 5. BIBLIOGRAFÍA

[1] <https://es.r4ds.hadley.nz/07-eda.html>

[2]

<https://mailchimp.com/es/resources/rfm-analysis/#:~:text=El%20modelo%20RFM%20considera%20los,mejor%20tus%20productos%20y%20servicios.>

[3]

<https://www.barilliance.com/es/guia-de-analisis-por-rfm-6-segmentos-clave-para-el-rfm-basado-en-marketing/#tab-con-18>

**¡GRACIAS!**