

Nombre: Andrea Díaz

Empresa: Drogería Nena

Prueba técnica Ciencia de Datos – Resumen Ejecutivo

Para la realización de esta prueba, se escogió el problema “Sistema de Alerta Temprana (Churn Rate)” y se realizaron los siguientes pasos para su realización:

Metodología

1. **Estudio de los datos y compresión del problema:** En esta primera etapa, se buscó comprender la estructura de los datos recibidos y cómo estos se pueden utilizar para resolver el caso de estudio. Además, en esta etapa se estudió a profundidad el caso de estudio y se plantearon los objetivos a cumplir.
2. **Limpieza y preparación de los datos:** Este es un paso crucial ya que lleva a los datos a un estado ideal que permita realizar un análisis de mayor calidad. Se realizaron tareas como tratamiento de valores nulos, agrupamientos, filtros, eliminación de columnas, entre otras.
3. **Segmentación de clientes:** Durante esta fase se realizó un Análisis RFM de los clientes por trimestre, es decir, se estudió sus hábitos de compra en términos de frecuencia, última compra realizada y monto gastado en cada trimestre desde 2023 hasta 2025; este paso se incluyó para estudiar cómo han ido cambiando estos hábitos e identificar posibles clientes en fuga. A cada cliente se le otorgó un valor de 1 a 5 en los campos: frecuencia, recencia y gasto, y su puntaje RFM se calculó a partir de la suma final de estos; para cada campo, la clasificación del 1 al 5 se basó en los cuantiles de estos campos respectivamente. Luego, usando el puntaje RFM por trimestre se dividieron a los clientes en “Cliente Leal”, “Cliente Regular” y “Cliente en Riesgo”. Para lograr esta segmentación de los clientes, se usó el modelo de clustering K-Means y el “Elbow Method” para encontrar el número ideal de clústeres; se decidió utilizar este modelo para la clasificación debido a que una clasificación manual podría no hacer justicia a la distribución de los puntajes RFM y a las tres variables que lo comprenden. El algoritmo K-Means para clustering agrupa los datos según características similares y se usó por ser un algoritmo rápido y de simple interpretación.
4. **Modelación de los datos:** Para predecir el comportamiento de los clientes en los próximos trimestres y anticipar pérdidas de clientes, se aplicó el modelo de predicción para series temporales conocido como ARIMA. Antes

de seleccionar el modelo a usar, visualizamos cómo ha variado el puntaje RFM promedio usando una serie de tiempo por trimestres; debido a que los datos históricos son confiables, no incluyen huecos o comportamientos estacionales y muestran una tendencia positiva, es posible usar un modelo de predicción como ARIMA.

5. **Evaluación de resultados:** Despues de aplicar el modelo de predicción se evaluaron los datos para garantizar que la predicción obtenida se acerque con los valores actuales.
6. **Creación de dashboard:** Al terminar con el proceso de limpieza, transformación y modelación de los datos, estos se exportaron y se cargaron a PowerBI para la creación de un dashboard que, a través de visualizaciones, muestre los insights descubiertos.

Insights:

- El lifetime value de la empresa calculado para 2025 es de 444752.56\$, un incremento del 208.13% del desde 2023.
- Para el último trimestre de 2025, la empresa tiene 63 clientes en fuga de 480 clientes activos.
- El puntaje rfm promedio del último trimestre de 2025 es de 11.36, un aumento del 6.25% en comparación con el último trimestre de 2024.
- La Farmacia Iker Alivio de la cadena Grupo Esperanza Estrella es el cliente cuyo puntaje RFM promedio ha disminuido más desde el 2024.
- Las señales tempranas de un cliente en fuga, o en riesgo, son una disminución significativa en la cantidad de productos comprados y en la cantidad de compras realizadas, por lo que el monto total gastado disminuye.

Recomendaciones:

En primer lugar, se recomienda conservar a los clientes clasificados como “Clientes Leales” y “Clientes Regulares” a través de descuentos y promociones para recompensar su fidelidad. Además, debido a la alta cantidad de clientes que entran en estas categorías se recomienda identificar aquellos clientes que aportan el mayor lifetime value y ofrecerles una atención personalizada para garantizar su lealtad.

Luego, en cuanto a aquellos clientes en la zona de riesgo, se recomienda la activación de campañas dirigidas directamente a estos clientes que contengan recordatorios de lo que ofrece Drogería Nena y descuentos exclusivos. Además, se recomendaría estudiar más a fondo las razones por las que los clientes puedan estar prefiriendo a la competencia.