

## Desafío Quilmes Data Science – Algo Promo

El equipo de Revenue de ABI – INBEV se encuentra testeando un algoritmo que optimiza promociones comerciales, personalizando los descuentos que se ofrecen a nuestros clientes (Autoservicios, Kioscos, Almacenes), con el objetivo de incrementar las ventas y minimizar la inversión que la compañía realiza en dichos descuentos.

Los productos se comercializan a través de una aplicación B2B desde la cual el cliente visualiza los productos disponibles, precios y descuentos.

El testeo se realiza a través de un AB testing en el cual los clientes se asignan a grupos de tratamiento (%50) y control (%50). Al grupo de tratamiento se les asignan las promociones personalizadas a través del algoritmo y al de control se le asignan promociones fijas (ejemplo: para todos un 10% sobre el monto de compra de los productos de la promo).

Las promociones por lo general se realizan para un conjunto de productos (SKU), bastando con que el cliente compre 1 de los productos de la promo, esta acción se reconoce como “redención”.

Como Data Scientist del equipo, tu objetivo es obtener insights para determinar si las promociones asignadas a través del algoritmo (grupo de tratamiento) obtienen mejor performance que las asignadas de forma tradicional (grupo de control) y encontrar oportunidades de mejora.

Algunas de las métricas que se observan son:

- La “Redemption rate” [clientes redimidos / clientes con promoción], es decir de los clientes a los que se le ofreció la promoción, cuantos de estos finalmente compraron durante el período de vigencia.
- Volumen promedio: promedio del volumen vendido en Hectolitros.

### El algoritmo:

El algoritmo personaliza la cantidad mínima que un cliente debe comprar para acceder a un % de descuento dado según una escala (a mayor cantidad mayor descuento). A la vez las recomendaciones del algoritmo se ajustan semanalmente según la respuesta de los clientes. Si la primera semana el cliente no aplica a la promo, probablemente a la siguiente semana le realice una mejor oferta (por ejemplo, reducir la cantidad mínima).

Ejemplo de recomendación a un cliente del grupo de tratamiento para uno de los SKU's de una promoción:

UID_PROMO_DESIGN_ID	POC_ID	PROMO_ID	PERIOD_START_DT	PERIOD_END_DT	SKU	MEASUREMENT_GROUP	DISCOUNT	MIN_QTY	MAX_QTY
6568a24a019dca56d52730cd	465841	f336cb88-e15b-42e1-9021-d06ad00cc322	1/12/2023	7/12/2023	19496	Group 1	4.00	7	999
6568a24a019dca56d52730cd	465841	0086d394-6fb8-4f20-882c-34394ce95c55	8/12/2023	14/12/2023	19496	Group 1	2.00	4	5
6568a24a019dca56d52730cd	465841	0086d394-6fb8-4f20-882c-34394ce95c55	8/12/2023	14/12/2023	19496	Group 1	4.00	6	999
6568a24a019dca56d52730cd	465841	1d73bc7b-e61d-45d0-ad47-e9fed0414942	15/12/2023	21/12/2023	19496	Group 1	4.00	8	999

En el ejemplo se observa que durante la primera semana le pide 7 unidades para acceder a un 4% de descuento en el SKU 19.496, en la segunda semana le exige 4 unidades para un 2% de dto y 6 unidades para un 4%; en la última 8 unidades para un 4%.

Ejemplo escala de descuento:

Grupo Control

Cantidad mínima	Descuento
3	2%
6	4%

Grupo de tratamiento

Cantidad mínima	Descuento
Personalizado x Algoritmo	2%
Personalizado x Algoritmo	4%

**Datos proporcionados:**

Se le proporcionan 2 datasets:

**PROMOCIONES:** recomendaciones para una promoción de 701 SKU's de cerveza durante el mes de Diciembre 2023.

UID_PROMO_DESIGN_ID	Id de promoción
POC_ID	Id de cliente
PROMO_ID	Id de recomendación
PERIOD_START_DT	Fecha inicio de la recomendación
PERIOD_END_DT	Fecha de fin de la recomendación
SKU	Código de producto
MEASUREMENT_GROUP	Grupo de testeo: Grupo 1 = Tratamiento, Grupo 0 = Control. Resto eliminar
DISCOUNT	% de descuento
MIN_QTY	Cantidad mínima para acceder al descuento
MAX_QTY	Cantidad máxima del rango de descuento

**VENTAS:** ordenes de pedidos realizados en los últimos 13 meses. El dataset de ventas contiene TODAS las ventas de cualquier producto (esté o no esté incluido en la promo).

POC_ID	Id de cliente
ORDER_DATE	Fecha de orden
SKU	Código de producto
TOTAL_VOLUME	Volumen en Hectolitros

UNIDADES	Cantidad de unidades
----------	----------------------

### Ejercicio:

- 1) Cálculo de resultados comparando Grupo Tratamiento vs Grupo de control y obtener conclusiones de performance según las siguientes métricas:
  - a. Integrar las promociones ofrecidas en Diciembre 2023 con la información de ventas. Tener en cuenta que las ventas deben coincidir con el período de vigencia de cada descuento.
  - b. Calcular la redemption rate SIMPLE: proporción de clientes que al menos compraron 1 producto de la promo (sin importar la cantidad).
  - c. Calcular la redemption rate REAL: % clientes que compraron una cantidad de unidades mayor o igual a la asignada por el algoritmo.
  - d. Comparar el volumen total y el volumen promedio. Previamente eliminar outliers.
- 2) Si analizamos el comportamiento histórico de compra de los clientes de cada grupo ¿En qué medida son alcanzables las cantidades mínimas exigidas por el algoritmo? ¿Qué conclusiones obtenemos? ¿El algoritmo es exigente? Aplicar medidas estadísticas para fundamentar las conclusiones. Al seleccionar los períodos de comparación, tener en cuenta que la venta de cerveza es un negocio de alta estacionalidad.
- 3) Repetir los análisis 1 y 2 pero generando clusters de clientes similares (posibles variables para clusterizar: volumen, cantidad de skus comprados por el cliente, frecuencia de compra, etc). ¿Existen clusters con diferencias en la performance del algoritmo?

### Entregables:

- Notebook
- Breve informe de resultados.