

Ejercicio – Data Science

Contexto

Nuestra empresa se encarga de las ventas de bienes de consumo masivos en un segmento B2B y cuenta con diferentes variedades de clientes que se conocen como puntos de ventas, la misma tiene una gran red de vendedores que se encargan de visitar a estos clientes de manera presencial.

El equipo de Nivel De Servicio esta a cargo de todo lo relacionado a la experiencia del cliente en la venta Offline, es decir, a través de una visita del vendedor presencial, dentro de esto mismo puede ver diferentes aspectos como mejorar la calidad y optimizar los tiempos de las visitas de los vendedores, entender que productos son mas recomendable vender en los diferentes clientes, entender que tan eficientes son las visitas que se hacen, entre otras cosas.

El equipo de Nivel De Servicios contacto al de Ciencia de Datos para entender cuál es la mejor manera de optimizar los tiempos de visita en los clientes, para lograr este objetivo se te solicita que generes un producto que permita entender cuál es la recomendación de tiempo de visita para cada uno de los clientes en base a distintas variables comerciales.

Esta herramienta debe calcular la duración total de visitas (en minutos) al mes más reciente disponible del dataset. Además, se requiere generar un score normalizado de esta recomendación de tiempo y, con ese score, elaborar un ranking regional que ordene a los clientes de cada región (es decir, que vas a tener que agrupar la información para tener un ranking 1 para cada región en el dataset).

Datos disponibles

Se suministra la siguiente información:

- dataset.csv
 - o cliente_id: Código correspondiente a cada cliente
 - o anio_mes: Período correspondiente al registro año mes en formato yyyyMM
 - o canal: Canal al que pertenece el cliente (Autoservicio, Kiosco MaxiQuiosco)
 - o region: Región de la cual forma parte el cliente
 - o minutos_visitados_mes: Minutos totales en el mes que se visitó el cliente en cada período
 - o cantidad_heladeras: Cantidad de heladeras que tenía el cliente en el período
 - o venta_total_negocios_mes: Venta del total de los negocios que tuvo el cliente unidad HI vendidos, esta venta puede no corresponder a la suma de los negocios por separado ya que incluye algunos productos que no se encuentran en los negocios (1,2,3,4)
 - o cantidad_productos_total_negocios_vendidos_mes: Cantidad de productos distintos vendidos en la venta total de todos los negocios en cada período
 - o cantidad_compras_total_negocios_mes: Cantidad de compras totales que el cliente hizo en el período
 - o dias_entre_compras_total_negocios_mes: días pasados entre cada compra realizada dentro del período

- venta_negocio1_mes: Venta correspondiente al negocio 1 en cada periodo, unidad HI vendidos
- venta_negocio2_mes: Venta correspondiente al negocio 2 en cada periodo, unidad HI vendidos
- venta_negocio3_mes: Venta correspondiente al negocio 3 en cada periodo, unidad HI vendidos
- venta_negocio4_mes: Venta correspondiente al negocio 4 en cada periodo, unidad HI vendidos

Consigna

En base a esto se solicita calcular la cantidad de minutos a visitar para cada uno de los clientes, se debe tener en cuenta lo siguiente:

- Se debe realizar un análisis (EDA) para entender lo siguiente:
 - Cuantos clientes tengo por región y canal
 - Cual es el ranking que tienen estos clientes por venta total
 - ¿Cuál negocio es el que más ventas tiene dentro de la compañía?
 - ¿Existen valores Duplicados?
 - ¿Existen valores Outliers?
- Se debera realizar un modelo de ML para calcular la cantidad de minutos a visitar para cada uno de los clientes
- Se debe buscar una manera de Normalizar esta recomendación para generar un score agrupado para cada región
- Se debe rankear a los clientes en base a este Score

El entregable debe incluir:

- Planteamiento del problema y solución
- Notebook de Python con la resolución.
- Tabla con la predicción en minutos, el score y el ranking para cada uno de los clientes.
- Un informe para el área de Nivel De Servicio con una explicación no técnica de los resultados. Este informe hay que presentarlo en una de las entrevistas del proceso de selección. (utilizar la herramienta que sea de su agrado para presentar puede ser una PPT, un Dashboard, etc)
- Aclaración: en caso de que algún requerimiento sea ambiguo, tomar una asunción y aclararla en el entregable