



Evento: Prueba Tecnica - Ruta análisis de datos 2023-1

Autores: Juliana Lalinde Velasquez (jlalinde@eafit.edu.co),

Santiago Puerta Florez (spuertaf@eafit.edu.co)

Fecha: 07/03/2023 (DD-MM-AA)

El presente documento contiene los detalles en específico sobre la prueba tecnica para la ruta de datos realizada a los recién ingresados a la misma.

JUSTIFICACION:

Al analizar las carencias de los integrantes Trainee y Jr. de la ruta de datos se halló un claro vacío en el concepto "Exploratory Data Analysis" (véase Tabla 1) conocido como "EDA", concepto básico a la hora de analizar un conjunto de datos dado.

Rol	Calificación obtenida EDA
Trainee Data Head	3
Jr. Data Head	4
Jr. Data Head	1
Jr. Data Head	1

Tabla 1. Calificaciones obtenidas para el concepto EDA de Trainee y Jr's.

Debido al vacío notado y teniendo en cuenta la necesidad del conocimiento del concepto se decidió hacer la prueba tecnica para así explicar de mejor manera el concepto.

LA METODOLOGÍA:

1.1 Modificando el conjunto de datos:

Se escogió el conjunto de datos "[Supermarket Sales](#)" de la página Kaggle relacionado con ventas de un almacén, se les hace un cambio a los registros de la columna "Branch" para que sean los almacenes: "Éxito San Antonio", "Éxito Wow Poblado", "Éxito Laureles" del grupo éxito ubicados en la ciudad de Medellín.



A su vez se reemplazan un 10% de los datos de la columna "Total" por NAN's para así poder evaluar que hacen los participantes ante un escenario de datos faltantes.

1.2 Las preguntas orientadoras

Debido a que la mayoría de los participantes afirmaban no tener experiencia llevando a cabo un EDA se decide elaborar una serie de preguntas orientadoras que ayuden al análisis de los datos.

La siguiente tabla contiene las preguntas elaboradas, así como la justificación de estas:

ID	Pregunta	Justificación
1	Los gerentes de la cadena de supermercados quieren ampliar el catálogo de productos que ofrecen en sus supermercados. ¿De qué productos deberían ampliar el catálogo?	En esta pregunta lo interesante puede ser ver si hacen el análisis separado por cada rama/ciudad, para revisar si se deberían hacer estrategias distintas o una global en términos de crecer la oferta de productos de la mejor línea. Lo otro es ver que concepto usan para crecer la línea, tienen distintas variables que pueden analizar, la que más vende en cantidades, en valores, la que más de gross income, la que tenga mejor calificación de la gente y cosas así, creo que es amplia a interpretación pero que si tiene un fin específico para todos. También pueden revisar por ejemplo qué género compra más y proponer un aumento en la línea específico de los productos que ese género consume más.
2	Desde la gerencia de la cadena de supermercados quieren realizar una campaña para incentivar la participación de las personas en sus programas de fidelidad (membresías). ¿Qué descuentos o programas recomendaría realizar?	Esta me parece bacana porque creo que es muy amplia, pero si plantea algo importante y es diferenciar entre clientes normales y con membresía. Los descuentos y programas también pueden tener demasiados enfoques: por ejemplo, revisar en que horarios se venden más las cosas y dar descuentos en esas franjas, revisar



		<p>en que días se venden más ciertas líneas de productos y dar descuentos (tipo los supermercados que hacen descuentos de carnes los martes, de vinos los sábados y así). También pueden revisar que es lo que más compran los miembros y dar descuentos en eso, o por el contrario que es lo que más compran los que no son miembros y dar descuento en eso para que se afilien o algo así. Descuentos según el método de pago.</p>
3	<p>La cadena de supermercados está pasando por una mala situación económica y necesita decidir cuál de sus ramas cerrar, ¿cuál recomendaría usted cerrar?</p>	<p>Esta última también es amplia y que puede verse de distintas maneras: ¿En cuál tienen más miembros? ¿En cuál se venden más cantidades? ¿En cuál se vende más en cantidad de dinero?</p>
4	<p>La cadena de supermercados quiere implementar una estrategia cashless, es decir, que no se reciban pagos en efectivo. ¿Cree usted que es viable y traería un beneficio?</p>	<p>Esta los obliga a mirar los datos desde otra separación: ya miraron género, membresías y ahora desde métodos de pago. Pueden revisar cantidades y total de dinero que se consume desde los distintos tipos de pago (efectivo y los dos que son cashless) y ver en general quienes los usan, porque tipo los miembros son menos probables que se vayan de supermercado que la gente que no tienen fidelidad, entonces de pronto es más importante con qué cosas pagan los miembros.</p>

1.3 El análisis del contexto

Una vez analizados los datos y respondidas las preguntas se les solicita a los participantes hacer un análisis del entorno o del contexto del cual los datos nos están hablando. ¿El barrio en el cual está ubicado cada uno de estos éxitos y los productos comprados tienen sentido? ¿Qué insights podemos hallar en los datos y que describen lo que está pasando en la ciudad?