1.- ELECCIÓN QUE SUPERMERCADOS ESTÁN EL TOP (Organización con WORD, TRELLO, GITHUB)

2.- SCRAPEO DE LOS DATOS

3.- TABLAS EN COMÚN

| Supermercado | Categoría | Subcategoría | Producto | Precio | Fecha | Descripción | Cantidad | id\_producto |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |

4.- BASE DE DATOS

5.- ANÁLISIS DE DATOS

*si queréis cambiar o añadir cualquier cosa, adelante*

VAMOS A IR INCLUYENDO RESPUESTA A CADA PREGUNTA PARA SABER QUÉ NECESITAMOS Y CÓMO HACERLO.

QUÉ PIDE MERCADONA:

1. **¿Está su API Actualizada? ¿Cuántos productos faltan?**

Respuesta: una vez montado el dataset de Mercadona, comprobar la fecha de descarga de cada artículo (es más que previsible que no esté actualizada, porque si lo estuviese no harían la segunda pregunta). Para ello necesitamos hacer un web scraping de la página de Mercadona para compararlo con el dataset que ya nos facilitaron desde el inicio (donde están todos los competidores + Mercadona).

A continuación, para saber cuántos productos faltan/sobran, debemos buscar un factor de comparación como por ejemplo el número de artículo.

1. **¿Cuál es la diferencia media de sus precios vs los precios de Carrefour?**

Aquí deberemos tener descargado el dataset de Carrefour con una estructura de categorías igual que la de Mercadona. Y comparar los precios por categorías para tener una media acotada por cada una de ellas (si vamos uno por uno en precios, será inviable. Debemos hacerlo por el precio medio de todos los artículos de esa categoría.

1. **¿Cómo me posiciono según precio y variedad de productos en las principales categorías con respecto a todas las marcas?**

Para analizar la variedad, haríamos un grupby por categorías y enumerar las marcas dentro de cada una.

Para comparar el precio, deberíamos agrupar por categoría y comparar precio medio de cada categoría igual que en la pregunta nº2.

1. **¿Cuál es el producto en el que somos más caros? ¿y el más barato?**

Punto sencillo.

1. **Actualmente queremos destinar dinero para igualar los mejores precios de la competencia. ¿Cuál es la cantidad total a invertir? PREGUNTAR A JULIÁN**

A partir de los resultados de la nº2, detectar dónde están las mayores diferencias en precios medios por categoría y ordenar por precio descendente para detectar los artículos con mayor diferencia.

Hacer un group by por categoría y un min/max por cada categoría. De esta forma tendremos los artículos con mayores diferencias.

Tendríamos que sacar el precio por kg, por litro o por unidad (columna de precio / cantidad una vez limpiado el dato). Este dato compararlo respecto a la competencia para detectar los productos con mayor diferencia de precio y definir el importe a reducir el precio de mercadona para igualar.

Calcular esa diferencia en base al nº de artículos vendidos en el periodo anterior y “simular” a igualdad de cantidades vendidas en el periodo actual.

1. **¿Hay algún precio de nuestros productos que no siga la tendencia de la competencia? Por qué podría ser?**

Confeccionar un boxplot donde tengamos los datos estadísticos de una serie de parámetros.

Agrupamos por categoría y hacemos el boxplot por cada una para ver cómo están organizados e incluir una regresión lineal.

1. **¿Cuántas marcas blancas tiene mi competencia? ¿estamos mejor o peor posicionados según precio y variedad?**

Aquí deberíamos extraer las marcas blancas de cada categoría y hacer un barplot para ver la que tiene cada marca.

Para el posicionamiento en precio, hacer un boxplot con los precios medios por cada categoría. Para la variedad podría limitarse a las categorias en que está presente la marca blanca (revisar ordenación de categorias de carrefour). Mirar si la categoría incluye a la marca o viceversa.

1. **¿Hay alguna categoría que quitarías o añadirías? ¿por qué?**

Hacemos el left-join de cada categoría para detectar qué categorías tenemos diferentes a la competencia (ya sea por presencia o ausencia). Aparte, debemos sacar el número de artículos que vendemos de cada una para poder comparar si está mejor posicionada que la competencia en base a ventas.

1. **Estamos pensando en abrir una nueva vertical llamada Mercadona eat&go. Su objetivo es que la gente pueda comprar y comer los productos en una sala acondicionada dentro de nuestro establecimiento, Te compartimos un csv con las compras de algunos usuarios en verano. ¿Qué tipo de alimentos recomendarías y por qué?**

Haríamos un barplot por numero de articulo vendido