

## Informe del proyecto de Introduccion a la Ciencia de Datos

Fuentes de datos utilizadas:

- 30 Mipymes → carpeta mipymes
- Mercatoria → carpeta mercatoria
- El toque → carpeta el toque

Los datos de Mercatoria fueron recopilados mediante scrapping. Se normalizaron los precios, que originalmente se encontraban en USD, mediante una funcion que recibe el precio del dolar según los cálculos de ElToque y realiza la aritmética correspondiente para convertirlos a moneda nacional.

Todos los datos se encuentran en sus respectivas carpetas dentro de una carpeta llamada json.

De Mercatoria se recopilaron datos sobre el arroz, frijoles, y yogurt, incluyendo el precio, país de origen, peso/gramaje, y marca. Del yogurt y los frijoles tambien se recopiló el tipo de frijol y sabor del yogurt.

En otra carpeta llamada notebook, se encuentra el jupyter notebook y una biblioteca llamada Tools, desde donde se importan las funciones necesarias para la realizacion del analisis.

Se utilizo la biblioteca plotly para la graficacion.

En una primera instancia se analiza el promedio de los siguientes productos (arroz, frijoles, pollo, aceite, sal y el azucar) productos basicos para cocinar, lo que nos permite notar la diferencia entre los precios y promediar su costo. La funcion utilizada para el analisis fue promedio\_canasta\_basica.

Una segunda grafica muestra cuatro graficos de donut que permite notar la diferencia entre mercatoria y las mipymes, subdivido entre arroz y frijoles, los paises que mas importan estos. Las funciones utilizadas fueron pais\_moda\_frijoles, pais\_moda\_arroz y mercatoria\_moda. La lógica del programa es buscar primero por la marca del producto, se guarda en una lista y luego se comprueba si está o no y. Esto nos permite contar el pais y asegurarnos de no contarlos mas de una vez.

La siguiente grafica tambien es una grafica de barras pero es parte de este analisis, pero permite ilustrar mejor los paises que tienen en conjunto, diviendo una parte frijoles y la otra arroz. En el .py la biblioteca llamada graph, en la función llamada barr\_com la primera parte del código realiza este proceso de escoger solo los paises en comun.

Se realizo una grafica heatmap que muestra el precio del kilogramo de frijol por tipo de frijol y país exportador, la función correspondiente es heatmap\_tabla en Tools. Se utilizan unicamente los datos de frijoles de Mercatoria.

Las dos últimas gráficas muestran datos referentes al yogurt en Mercatoria. La primera es una gráfica de frecuencias absolutas de los sabores de yogurt encontrados, la segunda un gráfico de pastel de los yogurts según el tipo (regular, probiótico, etc.). Los datos utilizados para crear esta gráfica se agregaron con la función yogurt\_data de la biblioteca Tools, la cual itera sobre toda la lista de yogurts recopilados de Mercatoria y cuenta en un diccionario la cantidad de apariciones de cada sabor y tipo de yogurt.