



ESCUELA POLITÉCNICA NACIONAL

ESCUELA DE FORMACIÓN DE TECNÓLOGOS



DISEÑO DE INTERFACES

TDSD

ASIGNATURA:

Análisis de Datos

PROFESOR:

Ing. Juan Pablo

PERÍODO ACADÉMICO:

2023-B

Prueba de Análisis de Datos

Nombre del estudiante:

Mateo Garzon

TÍTULO:

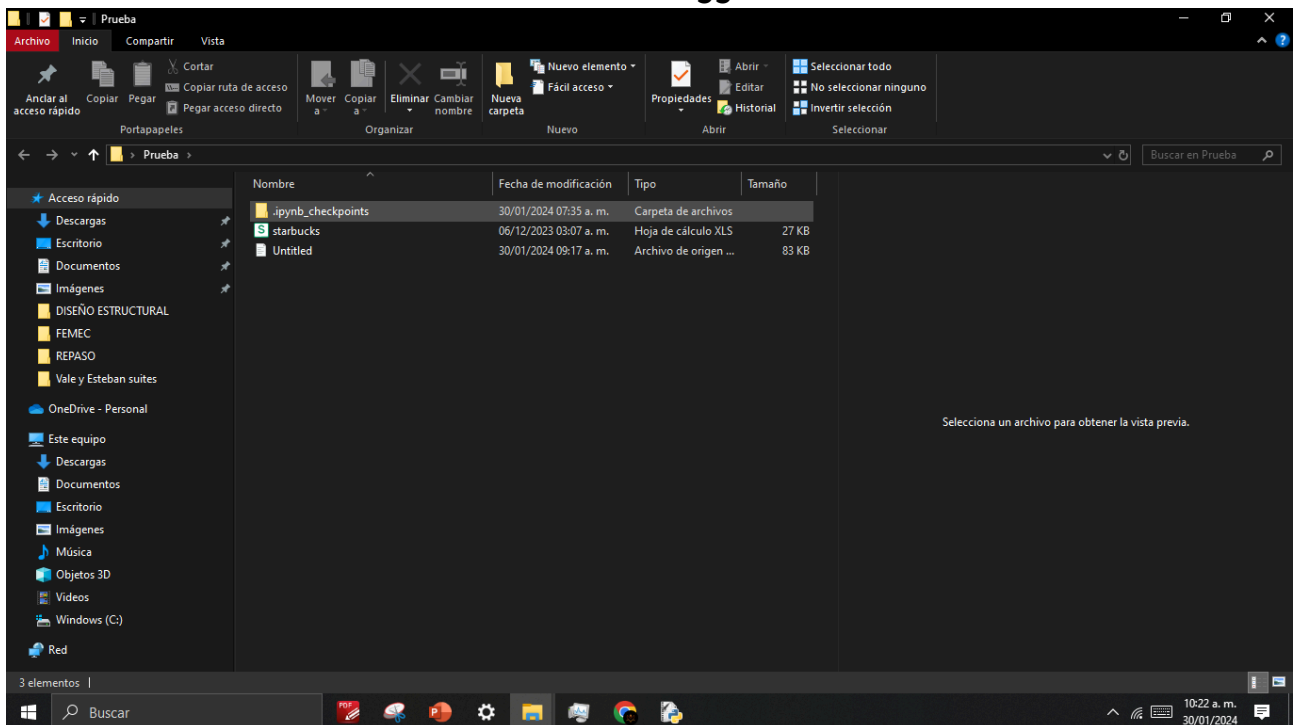
Prueba 1 Segundo Bimestre

2023-B

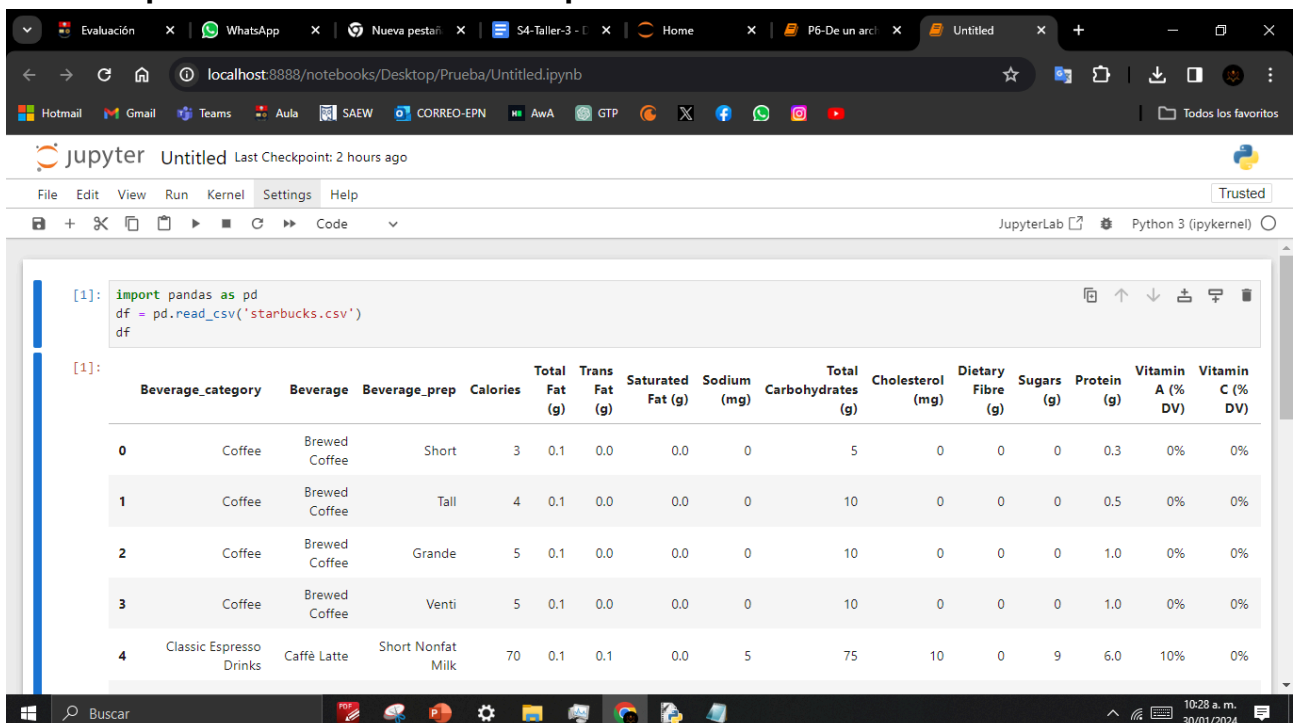
Caso 5

<https://github.com/WesitosFsa/PRUEBA-ANALISIS>

Primero Traemos el archivo starbucks csv de Kaggle



Paso 2 Importamos el dataframe csv a pandas



```
[1]: import pandas as pd
df = pd.read_csv('starbucks.csv')
df
```

	Beverage_category	Beverage	Beverage_prep	Calories	Total Fat (g)	Trans Fat (g)	Saturated Fat (g)	Sodium (mg)	Total Carbohydrates (g)	Cholesterol (mg)	Dietary Fibre (g)	Sugars (g)	Protein (g)	Vitamin A (% DV)	Vitamin C (% DV)
0	Coffee	Brewed Coffee	Short	3	0.1	0.0	0.0	0	5	0	0	0	0.3	0%	0%
1	Coffee	Brewed Coffee	Tall	4	0.1	0.0	0.0	0	10	0	0	0	0.5	0%	0%
2	Coffee	Brewed Coffee	Grande	5	0.1	0.0	0.0	0	10	0	0	0	1.0	0%	0%
3	Coffee	Brewed Coffee	Venti	5	0.1	0.0	0.0	0	10	0	0	0	1.0	0%	0%
4	Classic Espresso Drinks	Caffè Latte	Short Nonfat Milk	70	0.1	0.1	0.0	5	75	10	0	9	6.0	10%	0%

Paso 3 Importamos la librería matplotlib y numpy

```
<
:
import matplotlib.pyplot as plt
```

Paso 4 Establecemos que queremos encontrar y que queremos graficar

- 1 Gráfico de Líneas que nos saque un índice de cuantas grasas trans hay en cada bebida y cuántas grasas saturadas hay tambien ahi para establecer una diferencia y llegar a una conclusión
- 2 Gráfico de frecuencia de consumo de calorías establecer mediante un histograma cuantas calorías se consume mas y con que frecuencia dentro del dataframe
- 3 Gráfico de Barras Mirar cual es la bebida más consumida mediante un conteo de palabras en el dataframe
- 4 Gráfico de Dispersión mirar cual se consume más EL SODIO Y el colesterol
- 5 Gráfico de Pastel Mirar la categoría de bebidas más pedidas
- 6 Gráfico un gráfico de Líneas que nos dice el valor de cuantas proteínas y fibra dietética es consumida mediante un índice
- 7 Graficar un histograma que nos dice cuanto es el máximo colesterol que se consume y con qué frecuencia
- 8 Graficar un histograma Mirar cuánta cafeína y en qué frecuencia suelen los consumidores consumir esto mg de cafeína
- 9 Gráfico de Dispersión mirar qué tanta diferencia hay entre el azúcar y los carbohidratos
- 10 Gráfico de Pastel con que preparaciones hay más y menos sacando un porcentaje de estos

Paso 5 Empezamos con la codificación

1 GRÁFICO

```
plt.figure(figsize=(10, 6))
plt.plot(df['Grasas Trans (g)'], label='Trans Fat(g)')
plt.plot(df['Grasas Saturadas (g)'], label='Saturated Fat (g)')
plt.legend()
plt.title('Grasas trans y Saturada')
plt.xlabel('indice')
plt.ylabel('Valor')
plt.show()
```

Se creó un gráfico de líneas utilizando la función plt.plot.
Se utilizó la columna 'Grasas Trans' para el eje x y la columna 'Grasas Saturadas' para el eje y.
Se añadió una etiqueta y leyenda para facilitar la interpretación.
plt show muestra el gráfico

2 GRÁFICO

```
plt.figure(figsize=(10, 6))
plt.hist(df['Calories'], bins=20, color='blue', alpha=0.7)
plt.title('Frecuencia de Consumo de Calorias')
plt.xlabel('Caloriass')
plt.ylabel('Frecuencia')
plt.show()
```

Se creó un histograma utilizando la función plt.hist.
Se utilizó la columna 'Calories' para representar la distribución de las calorías.
Se ajustaron los bins para tener una representación más clara de la frecuencia.

3 GRÁFICO

```
plt.figure(figsize=(12, 6))
df['Beverage_category'].value_counts().plot(kind='bar', color='green')
plt.title('Bebida mas consumida')
plt.xlabel('Categoría de Bebidas')
plt.ylabel('Conteo')
plt.show()
```

Se creó un gráfico de barras utilizando la función value_counts en la columna 'Beverage_category'.
Se representa la frecuencia de cada categoría de bebida.

4 GRÁFICO

```
# 4. Gráfico de dispersión
plt.figure(figsize=(10, 6))
plt.scatter(df['Cholesterol (mg)'], df['Sodium (mg)'], alpha=0.5)
plt.title('Colesterol y Sodio')
plt.xlabel('Cholesterol mg')
plt.ylabel('Sodio mg')
plt.show()
```

Se creó un gráfico de dispersión utilizando la función plt.scatter.

Se utilizó 'Cholesterol' en el eje x y 'sodio' en el eje y para explorar la relación entre estas variables.

5 GRÁFICO

```
9]: # 5. Gráfico de pastel
plt.figure(figsize=(8, 8))
df['Beverage_category'].value_counts().plot(kind='pie', autopct='%1.1f%%')
plt.title('Categoria de bebidas mas pedidas')
plt.ylabel('')
plt.show()
```

Se creó un gráfico de pastel utilizando la función value_counts en la columna 'Beverage_category'.

Se representa la proporción de cada categoría de bebida en el conjunto de datos.

Segundo Gráfico de Líneas:

6 GRÁFICO

```
plt.figure(figsize=(10, 6))
plt.plot(df['Protein (g)'], label='Protein (g)', color='orange')
plt.plot(df['Dietary Fibre (g)'], label='Dietary Fibre (g)', color='purple')
plt.legend()
plt.title('Segundo Gráfico de Líneas')
plt.xlabel('Índice')
plt.ylabel('Valor')
plt.show()
```

Se creó otro gráfico de líneas con diferentes columnas ('Protein (g)' y 'Dietary Fibre (g)') para comparar tendencias.

7 GRÁFICO

```
# 7. Segundo histograma
plt.figure(figsize=(10, 6))
plt.hist(df['Cholesterol (mg)'], bins=15, color='red', alpha=0.7)
plt.title('Segundo Histograma - Cholesterol')
plt.xlabel('Cholesterol (mg)')
plt.ylabel('Frecuencia')
plt.show()
```

Se creó otro histograma para la columna 'Cholesterol (mg)' para explorar la distribución de colesterol en las bebidas.

8 GRÁFICO

```
plt.figure(figsize=(12, 6))
df['Caffeine (mg)'].value_counts().plot(kind='bar', color='pink')
plt.title('Frecuencia de bebidas ')
plt.xlabel('Cafeina')
plt.ylabel('Conteo ')
plt.show()
```

Se creó otro gráfico de barras utilizando la función value_counts en la columna 'Cafeina ' para explorar la frecuencia de consumo de cafeína en las bebidas.

9 GRÁFICO

```
plt.figure(figsize=(10, 6))
plt.scatter(df[' Sugars (g)'], df[' Total Carbohydrates (g)'], alpha=0.5, color='brown')
plt.title('Azucares vs Carbohidratos')
plt.xlabel('Azucares(g)')
plt.ylabel('Carbohidratos(g)')
plt.show()
```

Se creó otro gráfico de dispersión con 'carbohidratos' en el eje x y 'azúcares' en el eje y para explorar la relación entre estos nutrientes.

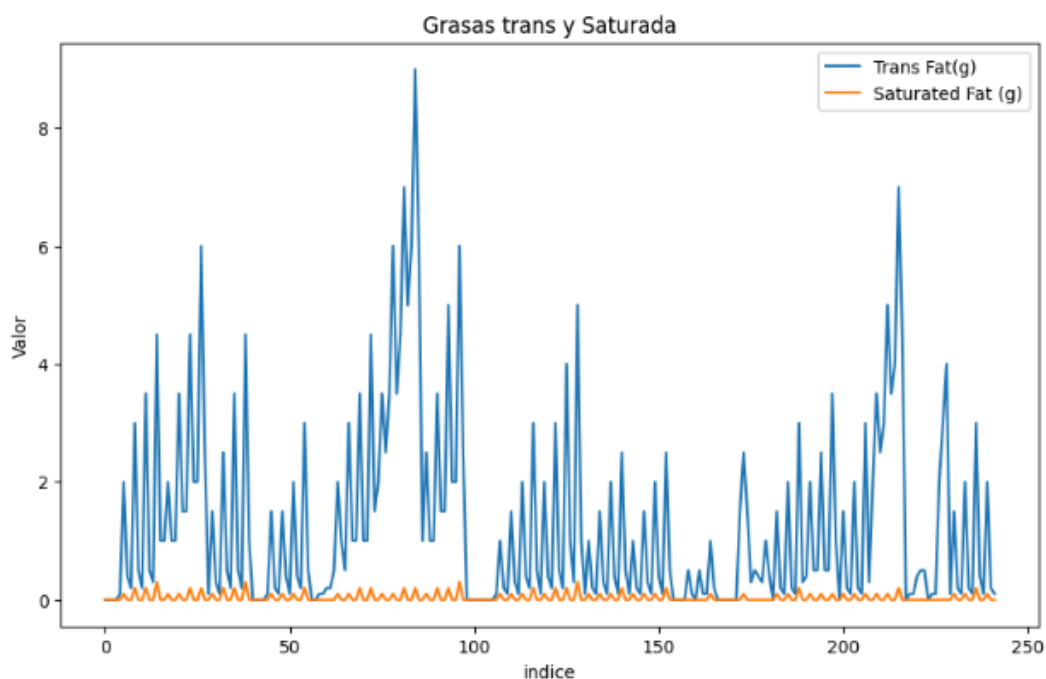
10 GRAFICO

```
plt.figure(figsize=(8, 8))
df['Beverage_prep'].value_counts().plot(kind='pie', autopct='%1.1f%%')
plt.title('Mayores preparaciones')
plt.ylabel('')
plt.show()
```

otro gráfico de pastel utilizando la función value_counts en la columna 'Beverage_prep' para representar la proporción de preparaciones de bebidas

Paso 6 Imprimimos las gráficas y establecemos 2 conclusiones de cada uno

Gráfico 1

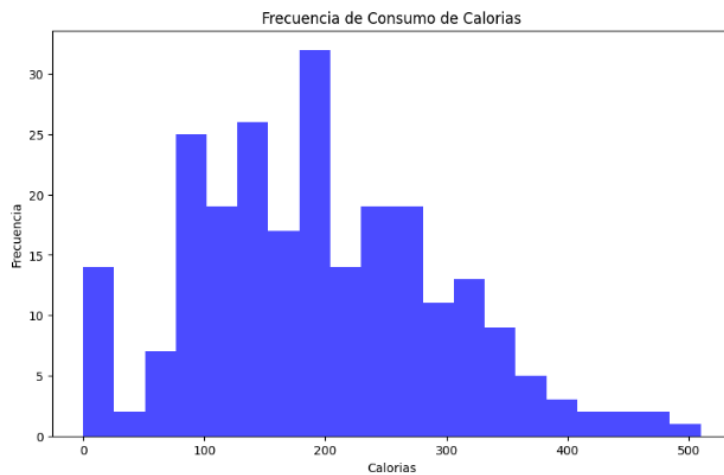


Conclusiones

-Preocupación por las grasas trans: Es importante estar alerta sobre las grasas trans porque están relacionadas con problemas de salud, como el aumento del riesgo cardíaco además de que están en mayor consumo que las grasas saturadas como se ve en el gráfico.

-Grasas saturadas : Aunque también debemos tener cuidado con las grasas saturadas, la comparación muestra que están presentes en menor cantidad que las grasas trans generando un mayor aporte en la salud

Gráfico 2

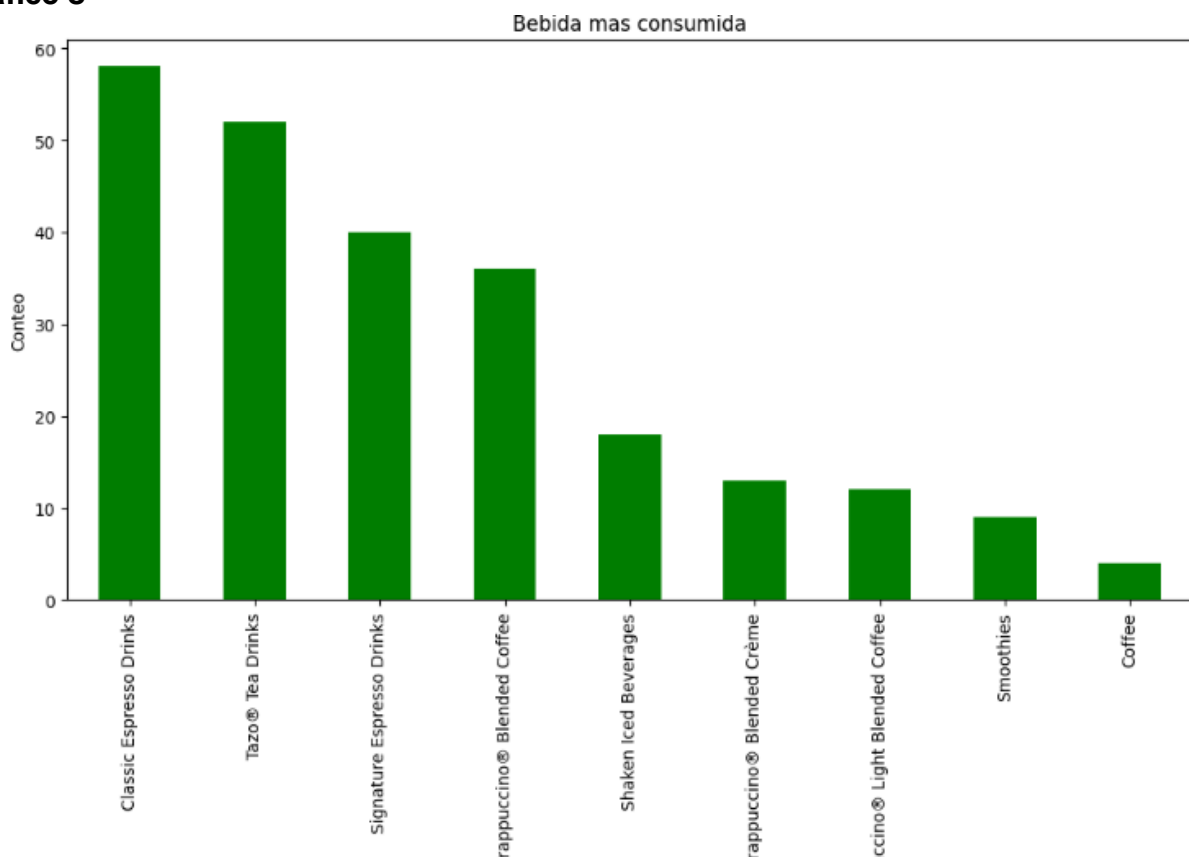


Conclusiones

-Concentración de calorías en un rango específico: Al observar que muchas bebidas tienen alrededor de 200 y 150 calorías, parece que hay un grupo de bebidas con valores similares de calorías. Esto podría indicar una tendencia en la oferta de bebidas dentro de un cierto rango calórico.

-Oportunidad para opciones bajas en calorías: La presencia de muchas bebidas en el rango de 150 a 200 calorías podría sugerir una oportunidad para ofrecer más opciones de bebidas con pocas calorías o implementar estrategias para reducir las calorías en algunas bebidas existentes. Esto podría satisfacer la demanda observada en este intervalo específico de calorías.

Gráfico 3

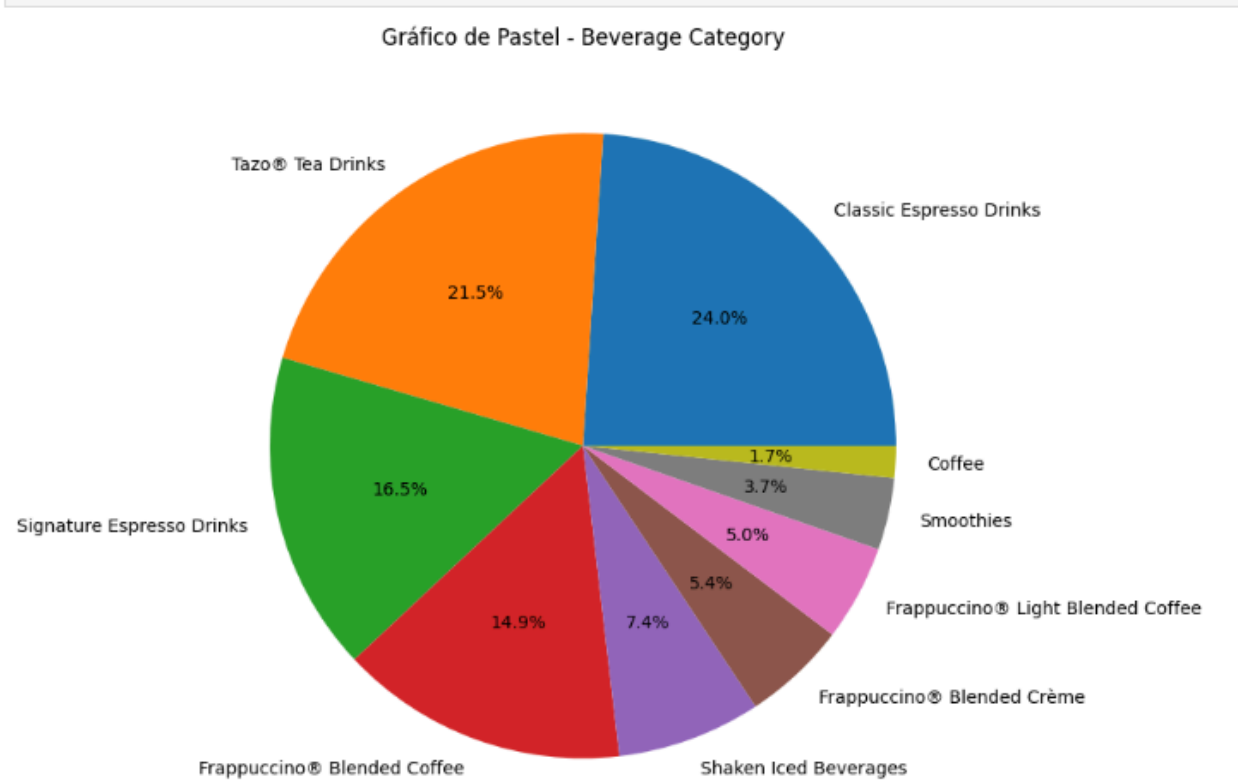


Conclusiones

-Popularidad del café clásico: El hecho de que el café clásico sea consumido con alta frecuencia indica que es la bebida más popular entre las opciones disponibles. Esta información puede ser útil para gestionar el inventario y tomar decisiones sobre qué ofrecer.

-Diversidad en las preferencias de bebidas calientes: Al ver que tanto el té como el café expreso tienen una presencia significativa, se sugiere una diversidad en las preferencias de bebidas calientes. Esto puede ser útil al planificar menús y estrategias de marketing, destacando la variedad de opciones disponibles para satisfacer diferentes gustos.

Gráfico 4

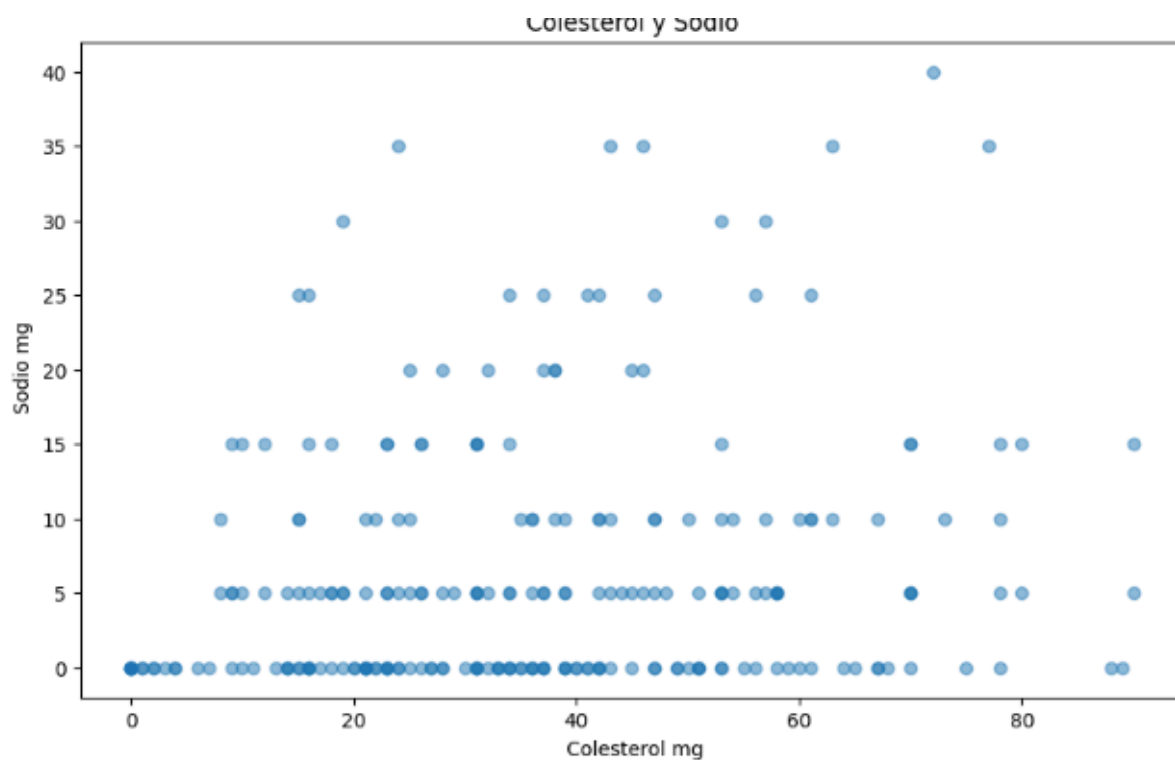


Conclusiones

-Bajas ventas en ciertas categorías: La posición de los cafés, batidos y Frappuccino en la región de menor consumo sugiere que estas categorías podrían tener una demanda más baja en comparación con otras. Esto podría ser importante para tomar decisiones sobre la promoción de estas categorías o considerar ajustes en la oferta.

-Oportunidad de promoción: Dado su menor presencia en el gráfico, existe una oportunidad para implementar estrategias de promoción específicas para estas categorías. Esto podría incluir ofertas especiales, marketing dirigido o la introducción de nuevas variaciones de productos para aumentar su atractivo y, en consecuencia, impulsar las ventas.

Grafico 5

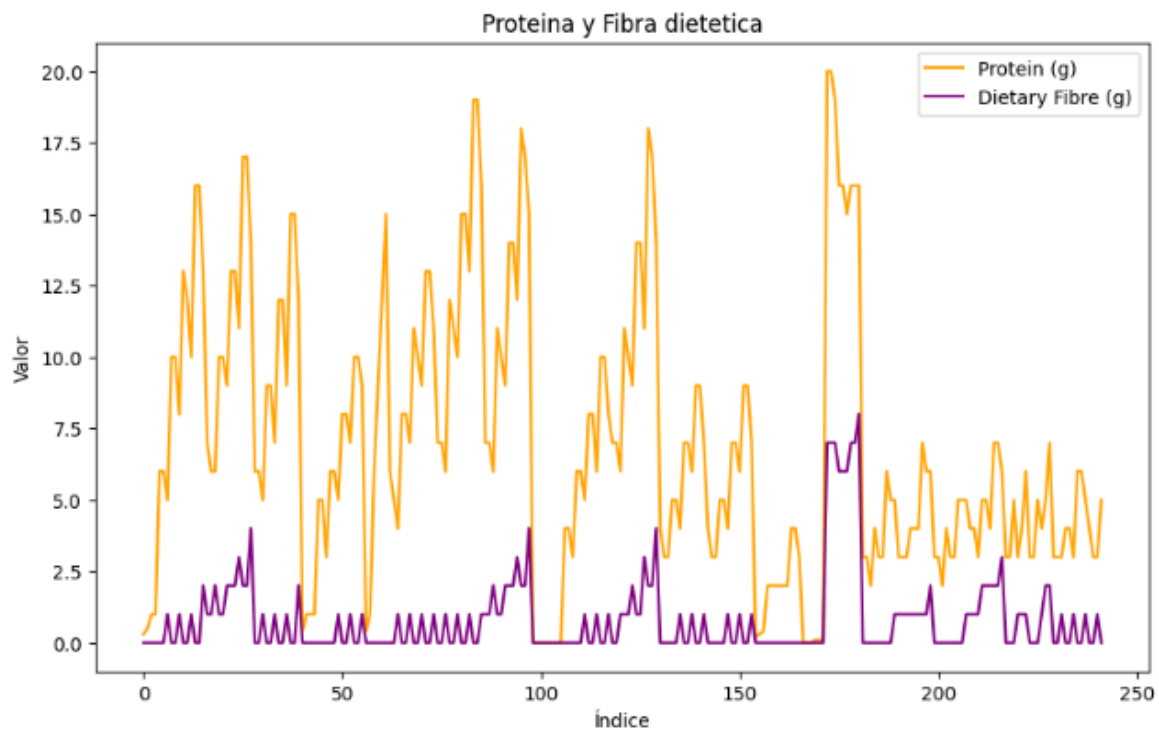


Conclusiones

-la atención al sodio: La concentración más alta en sodio puede indicar que las bebidas tienden a contribuir significativamente a la ingesta de sodio en comparación con el colesterol. Esto resalta la importancia de la atención a los niveles de sodio en las bebidas para aquellos que están monitoreando su consumo de sodio por razones de salud.

-Consideraciones nutricionales: La diferencia en la concentración entre sodio y colesterol puede ser útil para la toma de decisiones relacionadas con la formulación de productos, resaltando la necesidad de mantener un equilibrio saludable de nutrientes y ofrecer opciones con bajos niveles de sodio o colesterol, según las preferencias y necesidades nutricionales del consumidor.

Gráfico 6

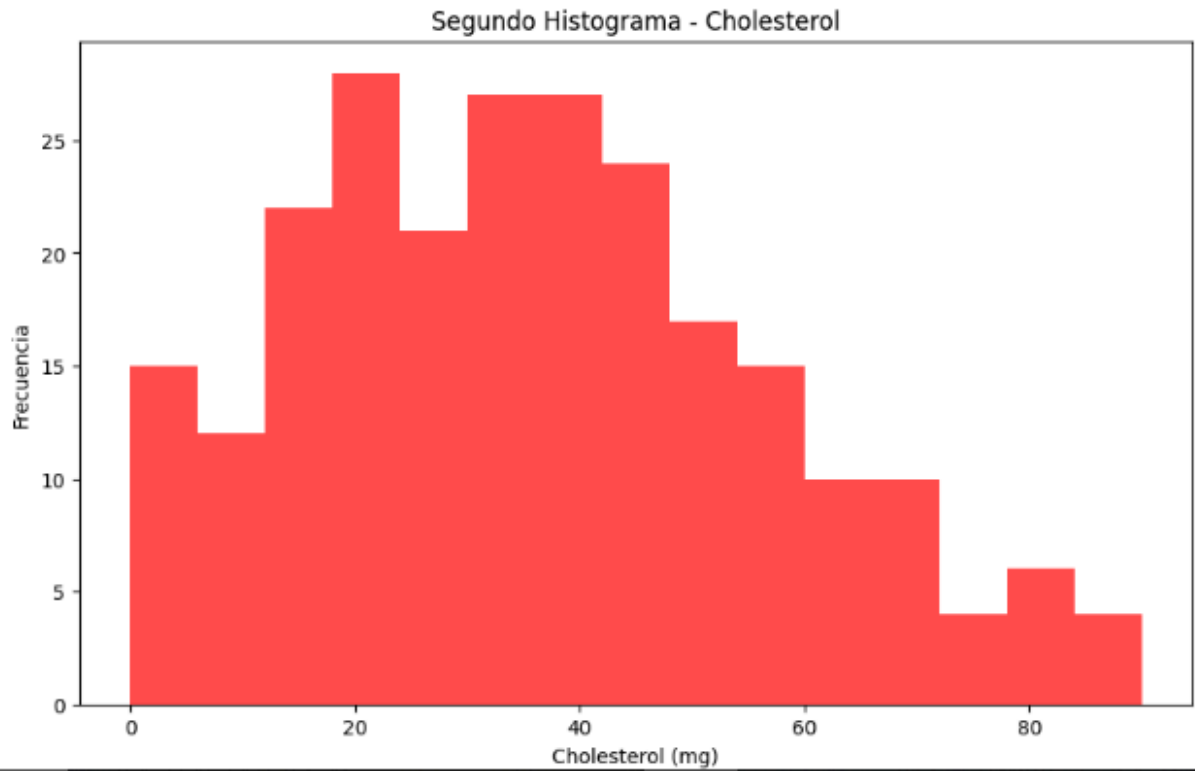


Conclusiones

-La mayor concentración de fibra en las bebidas resalta un aspecto positivo en términos de contenido nutricional. La fibra es esencial para la salud digestiva y puede ser un punto de venta positivo para aquellos que buscan opciones de bebidas más nutritivas.

-Oportunidad para la mejora de proteínas: La observación de que la concentración de proteínas es menor en comparación con la fibra podría señalar una oportunidad para mejorar la oferta de bebidas, especialmente para aquellos que buscan una mayor ingesta de proteínas. Introducir opciones con un mayor contenido proteico podría satisfacer las necesidades de un grupo demográfico que busca este componente nutricional específico.

Gráfico 7



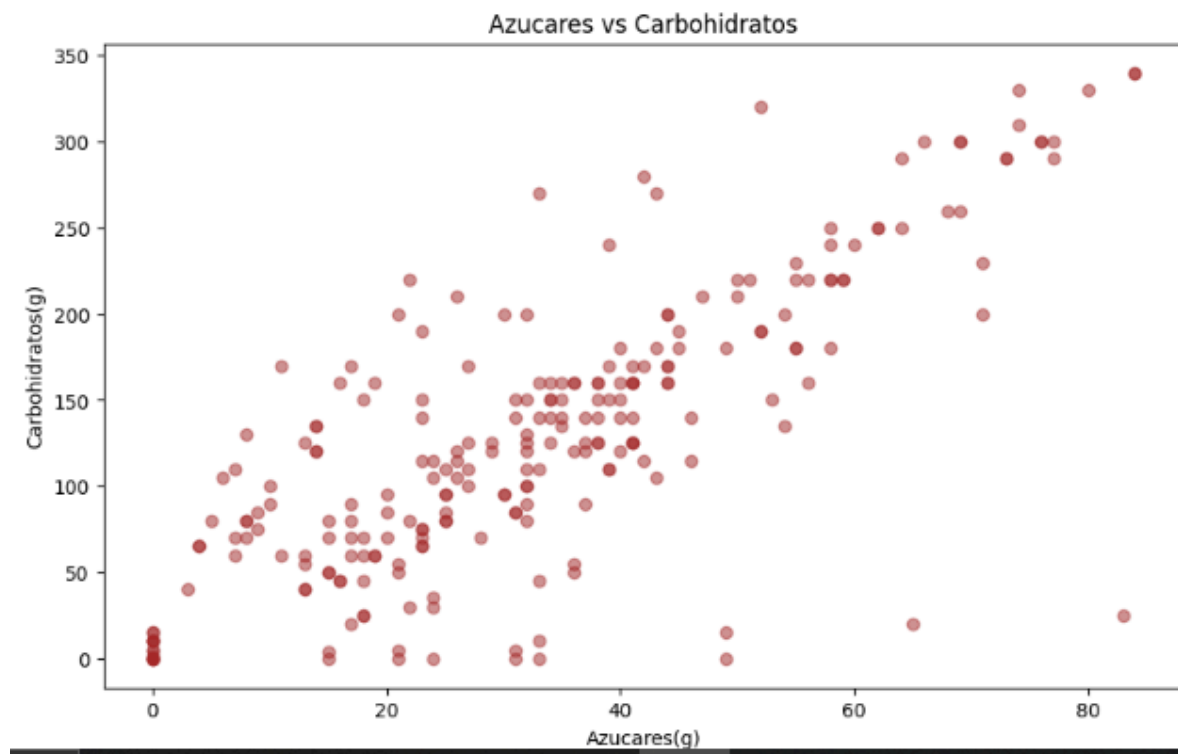
Conclusiones

Gráfico 8



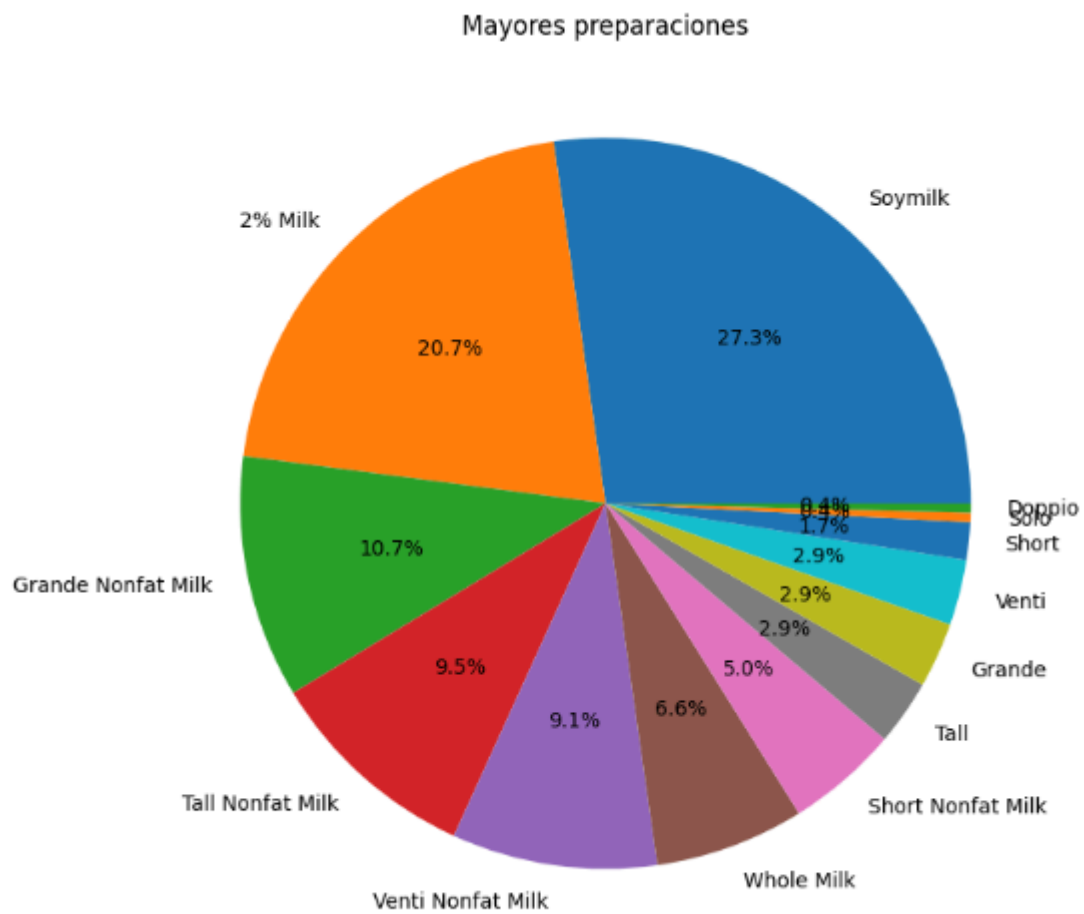
Conclusiones

Gráfico 9



Conclusiones

Gráfico 10



Conclusiones