Drink Menu Report

Name:
Doha Abdelaal Zakaria

Date

30/7/2022

1) Data analysis:

1.1) Data size:

```
(242, 18)
```

1.2) Data type of each column:

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 242 entries, 0 to 241
Data columns (total 18 columns):
 #   Column                    Non-Null Count  Dtype
---  ------                    --------------  -----
 0   Beverage_category         242 non-null    object
 1   Beverage                  242 non-null    object
 2   Beverage_prep             242 non-null    object
 3   Calories                  242 non-null    int64
 4    Total Fat (g)            242 non-null    object
 5   Trans Fat (g)             242 non-null    float64
 6   Saturated Fat (g)         242 non-null    float64
 7    Sodium (mg)              242 non-null    int64
 8    Total Carbohydrates (g)  242 non-null    int64
 9   Cholesterol (mg)          242 non-null    int64
 10   Dietary Fibre (g)        242 non-null    int64
 11   Sugars (g)               242 non-null    int64
 12   Protein (g)              242 non-null    float64
 13  Vitamin A (% DV)          242 non-null    object
 14  Vitamin C (% DV)          242 non-null    object
 15   Calcium (% DV)           242 non-null    object
 16  Iron (% DV)               242 non-null    object
 17  Caffeine (mg)             241 non-null    object
dtypes: float64(3), int64(6), object(9)
memory usage: 34.2+ KB
None
```

1.3) Number of nulls in each column:

```
Out[7]: Beverage_category            0
        Beverage                     0
        Beverage_prep                0
        Calories                     0
         Total Fat (g)               0
        Trans Fat (g)                0
        Saturated Fat (g)            0
         Sodium (mg)                 0
         Total Carbohydrates (g)     0
        Cholesterol (mg)             0
         Dietary Fibre (g)           0
         Sugars (g)                  0
         Protein (g)                 0
        Vitamin A (% DV)             0
        Vitamin C (% DV)             0
         Calcium (% DV)              0
        Iron (% DV)                  0
        Caffeine (mg)                1
        dtype: int64
```

2. Dataset Preparation:

2.1) Removing Duplicates:

It is made by considering the first duplication is unique and the rest are duplicates.

2.2) Replace Nulls:

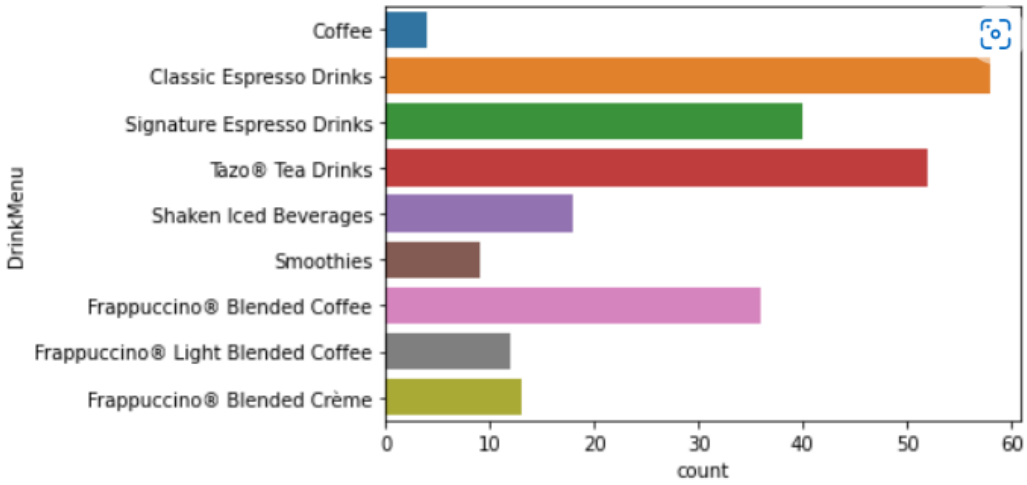
It is made by Replacing with the most repeated value in the column (Caffeine)

2.3) remove unnecessary columns:

```
The columns that was removed are: 'Beverage_prep','Trans Fat (g) ', 'Saturated Fat (
g)', ' Sodium (mg)',' Total Carbohydrates (g) ', ' Dietary Fibre (g)', ' Protein (g) ', 'Vita
min A (% DV) ', 'Vitamin C (% DV)',
       ' Calcium (% DV) ', 'Caffeine (mg)'
```

3. Dataset Visualization:

2.1) Plot Beverage_categry:



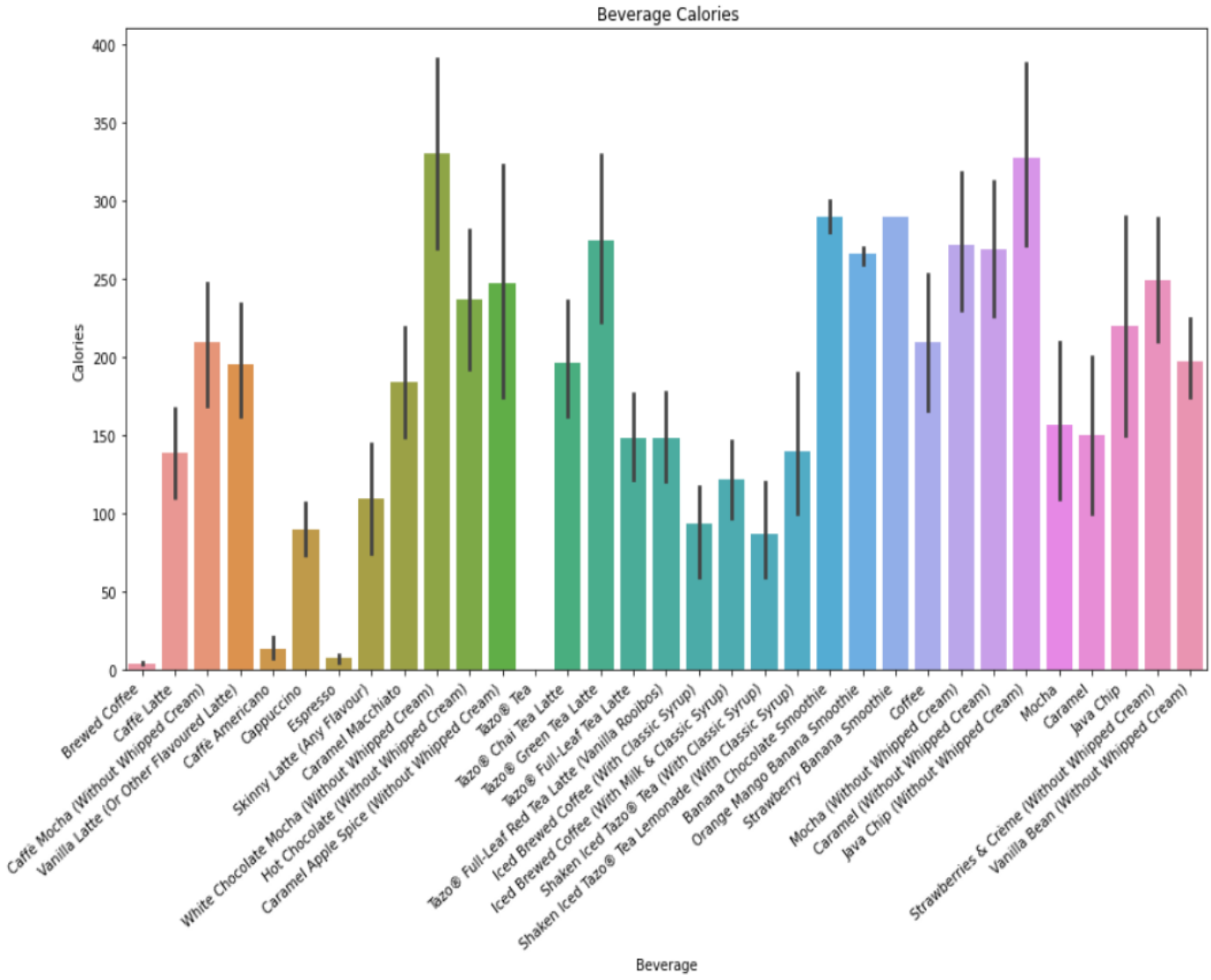
2.2) Plot of the drink that has the highest amount of calories

The following table shows the drinks in a descending order according to the sugar

Out[15]:

	DrinkMenu	Calories
84	Signature Espresso Drinks	510
215	Frappuccino® Blended Coffee	460
85	Signature Espresso Drinks	460
83	Signature Espresso Drinks	450
128	Tazo® Tea Drinks	450
...	...	...
0	Coffee	3
105	Tazo® Tea Drinks	0
104	Tazo® Tea Drinks	0
103	Tazo® Tea Drinks	0
102	Tazo® Tea Drinks	0

242 rows × 2 columns

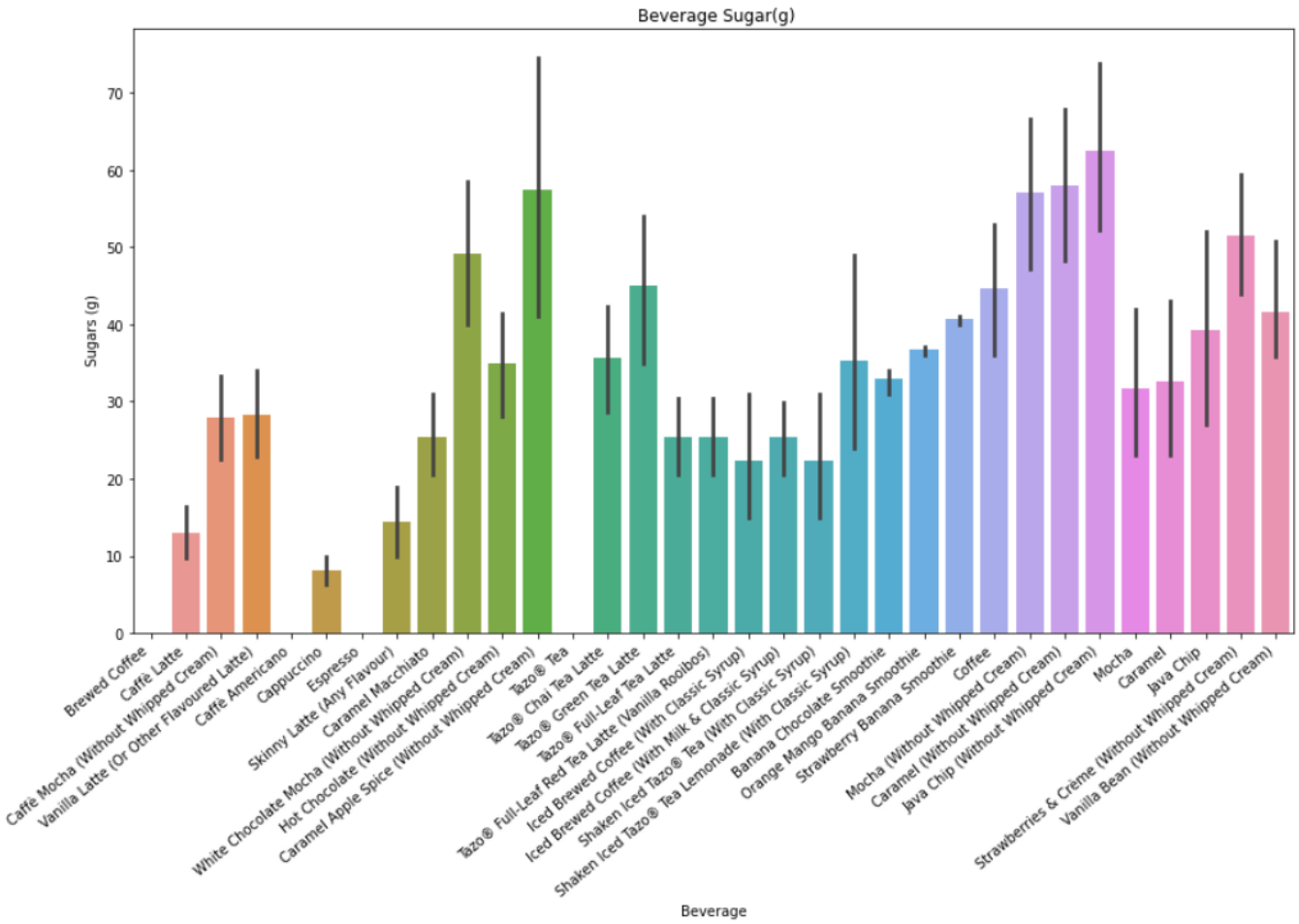

Beverage Calories
Calories
400
350
300
250
200
150
100
50
0
Brewed Coffee
Caffè Latte
Caffè Mocha (Without Whipped Cream)
Vanilla Latte (Or Other Flavoured Latte)
Caffè Americano
Cappuccino
Espresso
Skinny Latte (Any Flavour)
Caramel Macchiato
White Chocolate Mocha (Without Whipped Cream)
Hot Chocolate (Without Whipped Cream)
Caramel Apple Spice (Without Whipped Cream)
Tazo® Tea
Tazo® Chai Tea Latte
Tazo® Green Tea Latte
Tazo® Full-Leaf Tea Latte
Tazo® Full-Leaf Red Tea Latte (Vanilla Rooibos)
Iced Brewed Coffee (With Classic Syrup)
Iced Brewed Coffee (With Milk & Classic Syrup)
Shaken Iced Tazo® Tea (With Classic Syrup)
Shaken Iced Tazo® Tea Lemonade (With Classic Syrup)
Banana Chocolate Smoothie
Orange Mango Banana Smoothie
Strawberry Banana Smoothie
Coffee
Mocha (Without Whipped Cream)
Caramel (Without Whipped Cream)
Java Chip (Without Whipped Cream)
Mocha
Caramel
Java Chip
Strawberries & Crème (Without Whipped Cream)
Vanilla Bean (Without Whipped Cream)
Beverage

2.3) Plot of the drink that has the highest amount of Sugar

The following table shows the drinks in a descending order according to the sugar

Out[19]:

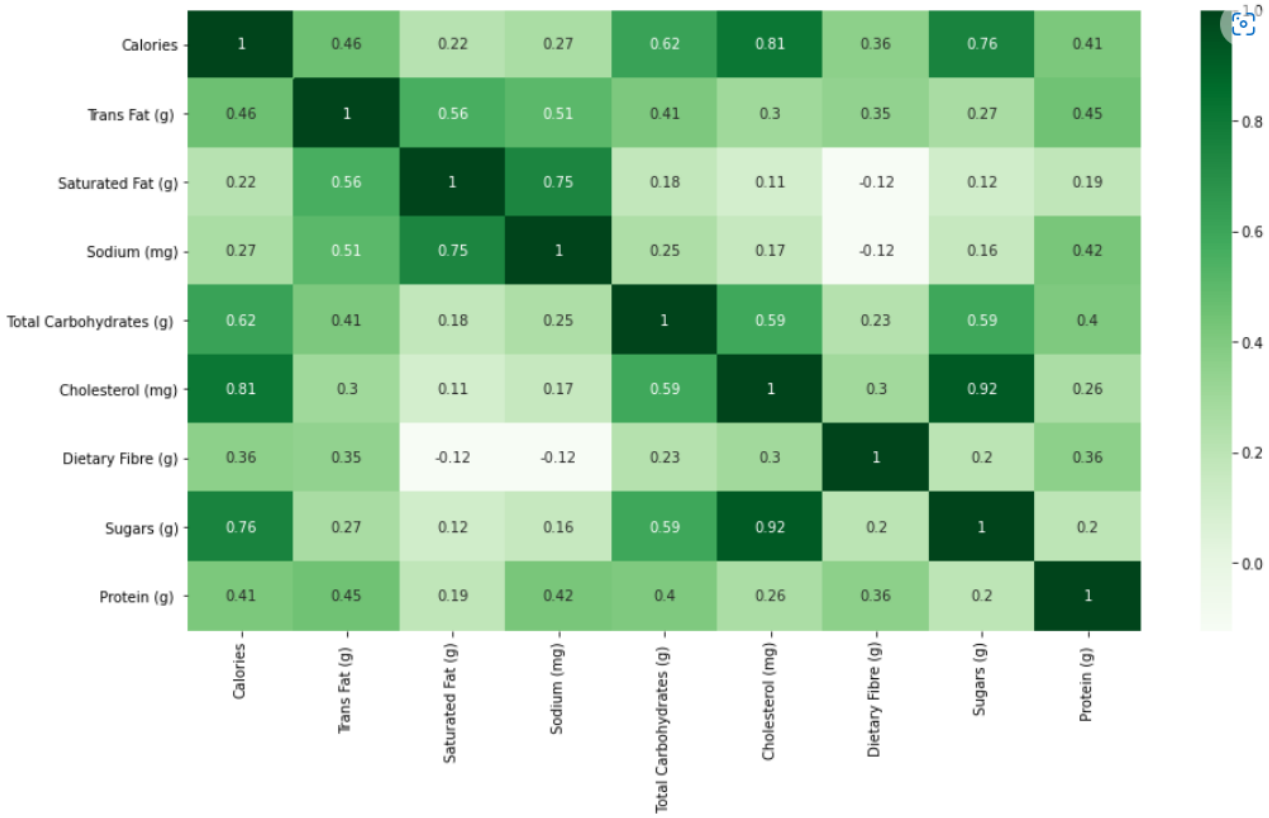
	Beverage_category	Beverage	Beverage_prep	Calories	Total Fat (g)	Trans Fat (g)	Saturated Fat (g)	Sodium (mg)	Total Carbohydrates (g)	Cholesterol (mg)	Dietary Fibre (g)	Sugars (g)	Protein (g)	Vitamin A (% DV)	V
214	Frappuccino® Blended Coffee	Java Chip (Without Whipped Cream)	Venti Nonfat Milk	420	5	4.0	0.0	5	340	90	2	84	7.0	10%	
215	Frappuccino® Blended Coffee	Java Chip (Without Whipped Cream)	Whole Milk	460	10	7.0	0.2	15	340	90	2	84	7.0	6%	
101	Signature Espresso Drinks	Caramel Apple Spice (Without Whipped Cream)	Venti	360	0	0.0	0.0	0	25	89	0	83	0.0	0%	
216	Frappuccino® Blended Coffee	Java Chip (Without Whipped Cream)	Soymilk	430	8	4.5	0.0	0	330	88	3	80	6.0	6%	
206	Frappuccino® Blended Coffee	Caramel (Without Whipped Cream)	Whole Milk	370	5	3.0	0.1	15	300	78	0	77	5.0	8%	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
56	Classic Espresso Drinks	Espresso	Solo	5	0	0.0	0.0	0	0	1	0	0	0.4	0%	
57	Classic Espresso Drinks	Espresso	Doppio	10	0	0.0	0.0	0	1	2	0	0	1.0	0%	
102	Tazo® Tea Drinks	Tazo® Tea	Short	0	0	0.0	0.0	0	0	0	0	0	0.0	0%	
103	Tazo® Tea Drinks	Tazo® Tea	Tall	0	0	0.0	0.0	0	0	0	0	0	0.0	0%	
0	Coffee	Brewed Coffee	Short	3	0.1	0.0	0.0	0	5	0	0	0	0.3	0%	



2.3) **Plot of the drink that has the highest amount of Sugar**

the following plot shows the collrelation between the columns after removing the unnecessary columns.

It is iportatant to consider thatThe relationship between two variables is generally considered strong when their r value is **larger than 0.7**. The correlation r measures the strength of the linear relationship between two quantitative variables. Pearson r: r is always a number between -1 and 1.


Calories
Trans Fat (g)
Saturated Fat (g)
Sodium (mg)
Total Carbohydrates (g)
Cholesterol (mg)
Dietary Fibre (g)
Sugars (g)
Protein (g)
1 0.46 0.22 0.27 0.62 0.81 0.36 0.76 0.41
0.46 1 0.56 0.51 0.41 0.3 0.35 0.27 0.45
0.22 0.56 1 0.75 0.18 0.11 -0.12 0.12 0.19
0.27 0.51 0.75 1 0.25 0.17 -0.12 0.16 0.42
0.62 0.41 0.18 0.25 1 0.59 0.23 0.59 0.4
0.81 0.3 0.11 0.17 0.59 1 0.3 0.92 0.26
0.36 0.35 -0.12 -0.12 0.23 0.3 1 0.2 0.36
0.76 0.27 0.12 0.16 0.59 0.92 0.2 1 0.2
0.41 0.45 0.19 0.42 0.4 0.26 0.36 0.2 1
1.0
0.8
0.6
0.4
0.2
0.0