ANALYZING NUTRITIONAL INFORMATION OF STARBUCKS MENU ITEMS

George Mason University
AIT 664-DL2 | Prof.Dr.Ebrima Ceesay

Snehita Vattamwar
Department of Applied Information Technology
George Mason University
Fairfax,Virginia
svattamw@gmu.edu

Divya Nalluri
Department of Data Analytics Engineering
George Mason University
Fairfax,Virginia
dnalluri@gmu.edu

ABSTRACT

This research examines the nutritional values of drinks typically available in coffee shops and similar venues. It analyzes data covering aspects like calorie count, fat, carbohydrates, fiber, protein, and sodium content for each drink. The objective is to assess how this information could influence business tactics and public health policies.

By conducting a descriptive analysis, the research provides insights into the nutritional composition of various drinks and their potential health effects. The results show considerable differences in nutritional content across different beverage types, pointing out chances for businesses to introduce healthier choices and for government officials to establish guidelines that foster improved dietary habits.

In summary, this research enhances our comprehension of the nutritional characteristics of drinks and the broader implications for encouraging healthier eating habits. This highlights the importance of nutritional information in influencing corporate strategies and public health policies.

Keywords— *Starbucks, Nutritional Analysis, Beverages, Calorie Count, Fat Content, Carbohydrates, Fiber, Protein, Sodium, Public Health, Dietary Options, Data Analytics, Menu Development, Consumer Choices, Health Policies, Business Strategies, Dataset Analysis, Visualization.*

I. INTRODUCTION

Starbucks, initially established in Seattle, has evolved significantly from its humble beginnings to a global entity in the coffee sector. This company provides an extensive range of products, from the widely celebrated Frappuccinos® to classic espresso offerings and assorted cold beverages, meeting the diverse tastes of a global audience [18]. This diverse selection underscores the pressing need for clear and comprehensive nutritional information.Currently, heightened public awareness of nutrition and dietary choices, spurred by the increasing prevalence of health issues such as obesity, diabetes, and cardiovascular disease, has amplified the need for clear nutritional information. This increased demand aids in making more informed decisions regarding dietary habits. [18].

Despite Starbucks' efforts to provide such information, a gap persists in public comprehension and application of this data. The nutritional details, often limited to calorie information, do not convey the full implications necessary for making health-aligned choices [19]. The general populace lacks the specialized knowledge required for a deep analysis of such data, thereby impeding their ability to make choices that align with their health goals.

Furthermore, the deficiency in detailed nutritional analysis potentially undermines Starbucks' credibility. In an era where corporate transparency holds significant value, the expectation for comprehensive disclosure of nutritional facts is paramount. Inadequacies in this area could diminish consumer trust and loyalty, negatively impacting Starbucks' reputation and its initiatives aimed at promoting healthier dietary options [19].

This paper delves deep to resolve these challenges by conducting a thorough analysis of the nutritional content of Starbucks' menu items. The goal is to close the gap between the available data and consumer understanding, enabling customers to make knowledgeable dietary choices. These initiatives aim to cultivate a culture of wellness, transparency, and informed decision-making both within the Starbucks community and the wider society.

II. PROBLEM STATEMENT

Starbucks is known around the world for its wide variety of food and drinks, which appeal to many different tastes. As people become more health-conscious, they increasingly want clear information about the nutritional content of what they eat and drink. Although Starbucks does offer some nutritional information, it lacks a thorough analysis and explanation of this data. This issue leads to several problems:

1. *Lack of Proper Guidance for Customers:* Even though nutritional information is available, many customers don't have the knowledge or tools to understand this data properly. This can prevent them from making informed choices at Starbucks.

2. *Health Concerns*: With the rise of health issues related to diet, like obesity, diabetes, and heart disease, people are more careful about their food choices. Without detailed nutritional analysis from Starbucks, customers might end up choosing items that are not suitable for their health goals.

3. *Need for Transparency from Companies*: As customers demand more transparency about the nutritional content of their food, they expect companies like Starbucks to be open and responsible. If Starbucks does not provide detailed nutritional information, it could lose customers' trust and loyalty.

III. MOTIVATION

The motivation behind this project stems from the need to address these challenges and provide actionable insights for both consumers and Starbucks:

Empowering Consumer Choices: By conducting a comprehensive analysis of Starbucks' menu items, this project aims to empower consumers with the knowledge they need to make healthier choices. By understanding the nutritional profiles of various menu items, consumers can make informed decisions that align with their dietary preferences and health goals.

Enhancing Corporate Transparency: Through detailed nutritional analysis, Starbucks can demonstrate its commitment to corporate responsibility and transparency. By providing consumers with access to accurate and understandable nutritional information, Starbucks can strengthen its brand reputation and build trust with its customer base.

Promoting Healthier Eating Habits: Ultimately, the goal of this project is to contribute to the promotion of healthier eating habits among consumers. By raising awareness of the nutritional content of Starbucks' menu items, we aim to encourage healthier choices that support overall well-being and long-term health.

IV. LITERATURE REVIEW

Navigating the caloric seascape of urban coffee shops, Huang et al. in their 2009 and 2011 studies delved into the calorie content of beverages at major coffee chains in New York City. They found many offerings alarmingly high in calories, pointing out the associated obesity risks due to excessive caloric intake from popular drinks [1][11].

The effectiveness of calorie transparency was critically assessed by Bollinger et al. (2010) who explored how displaying calorie information impacts consumer choices in chain restaurants, including Starbucks. Their research indicated that although consumers were informed about calorie counts, this knowledge did not significantly influence their buying decisions, suggesting that mere calorie disclosure might not be enough to alter consumer behavior [2]. In a related vein, a 2021 study by Avery et al. noticed a nuanced shift where Starbucks' proactive calorie information disclosure led to some customers opting for lower-calorie options, highlighting variable responses based on personal health awareness [14].

The trend towards healthier choices was further supported by studies from Nguyen et al. and Verywell Fit, which observed an increasing consumer preference for healthier or lower-calorie beverage options at Starbucks, driven by a heightened public health consciousness [15][16]. Insights from Jephcote, Shmerling, and Han and Ren extend this discussion to the broader health implications of dietary selections. They delve into how dietary components like carbohydrates and sodium can influence heart disease, and how caloric limitation can offer significant health benefits, providing compelling reasons for consumers' preference for certain beverages [7][8][9].

In summary, these studies collectively underscore the importance of understanding the nutritional composition of food and beverage offerings in chain restaurants like Starbucks. Mandatory calorie posting has emerged as a potential tool for promoting healthier choices among consumers, with implications for public health policy and business practices in the foodservice industry. Future research in this area could explore the long-term effects of calorie posting on obesity rates and consumer behavior across diverse restaurant settings.

V. DATASET

The dataset is chosen from Data.world, titled "starbucks_drinkMenu_expanded.csv,"comprises information on various Starbucks beverages. It includes columns detailing beverage categories, names, preparation methods, and nutritional components such as calories, total fat, sodium, carbohydrates, cholesterol, protein, and vitamins. With data on key nutrients like saturated fat, sugars, and caffeine content, the dataset facilitates analyses to discern trends in beverage offerings, identify healthier options, and aid consumers in making informed choices aligning with their dietary needs and preferences. Moreover, this dataset serves as a valuable resource for menu development, nutritional analysis, and consumer education initiatives aimed at promoting healthier beverage consumption habits.

DATASETE URL:

https://data.world/raop72027/starbucks/workspace/file?filename=starbucks_drinkMenu_expanded.csv

Dataset contains 18 columns

Column Name	Data Type	Description
Beverage_Category	String	Indicates the category to which the beverage belongs.
Beverage	String	Specifies the name or type of beverage.
Beverage Prep	String	Describes the preparation method or variant of the beverage, if applicable.
Calories	Integer	Represents the number of calories in the beverage.
Total Fat	String	Indicates the total fat content in the beverage.
Trans Fat	Integer	Denotes the amount of trans fat present in the beverage.
Saturated Fat	Integer	Specifies the quantity of saturated fat in the beverage.
Sodium_mg	Integer	Indicates the sodium content in milligrams (mg) in the beverage.
Total_carbohydrates_g	Integer	Specifies the total carbohydrates content in grams (g) in the beverage.
Cholestrol_mg	Integer	Represents the cholesterol content in milligrams (mg) in the beverage.
Dietary_Fibre_g	Integer	Indicates the dietary fiber content in grams (g) in the beverage.
Sugars_g	Integer	Specifies the sugar content in grams (g) in the beverage.
Protien_g	Integer	Represents the protein content in grams (g) in the beverage.
Vitamin_a_dv	Integer	Denotes the percentage of the daily value (%DV) of Vitamin A provided by the beverage.
Vitamin_c_dv	Integer	Denotes the percentage of the daily value (%DV) of Vitamin C provided by the beverage.
Calcium_dv	Integer	Denotes the percentage of the daily value (%DV) of Calcium provided by the beverage.
Iron_dv	Integer	Denotes the percentage of the daily value (%DV) of Iron provided by the beverage.
Caffeine_mg	Text	Specifies the amount of caffeine in milligrams (mg) in the beverage.

VI. PROPOSED APPROACH

A. Data Collection and Preparation

- Acquiring Data: Secure the dataset named "starbucks_drinkMenu_expanded.csv" from Data.world, which is rich in nutritional specifics for a wide array of Starbucks beverages.
- Purifying Data: Engage in an exhaustive review of the dataset to address and correct any instances of incomplete data, irregularities, or inaccuracies in data presentation. This process might include filling in missing values, standardizing data representations, and purging redundant entries.
- Refining Features: Identify and prioritize the nutritional factors critical for analysis, focusing on elements such as calorie counts, fat percentages, sugar levels, and caffeine content.

B. Initial Data Exploration

- Statistical Summary: Generate and review statistical summaries for the data's numerical aspects to gauge its distribution, variability, and central metrics.
- Data Visualization: Craft a range of visual aids (like histograms, box plots, and scatter plots) to visually assess the spread and patterns of important nutritional metrics throughout different types of beverages.
- Trend and Anomaly Detection: Apply exploratory data analysis techniques to spot and highlight notable patterns or deviations, especially concerning calorie and sugar concentrations.

C. Categorizing Data

- Applying Segmentation Logic: Use specific benchmarks to categorize beverages into distinct groups based on their nutritional content, employing set boundaries for metrics like calories and sugar content.
- Identifying Patterns: Perform a thorough investigation of these segments to unearth prevalent trends or significant characteristics, especially the occurrence of items with elevated sugar or calorie figures in certain segments.

D. Advanced Analytical Techniques

- Examining Correlations: Delve into the relationships among different nutritional dimensions through correlation analysis to pinpoint elements that significantly influence the nutritional profile of beverages.
- Illustrating Relationships: Utilize graphical tools such as heatmaps or scatter plots to visually convey the strength and direction of correlations among various nutritional metrics.

E. Deriving and Sharing Insights

- Unearthing Insights: Extract meaningful insights from the exploratory analysis, segmentation, and correlation study, highlighting options that are nutritionally favorable or rich in specific nutrients.
- Creating Educational Material: Convert these insights into accessible content aimed at educating consumers, thereby facilitating informed nutritional choices based on their health objectives.

F. Tool Design and User Feedback

- Prototype Design: Architect and prototype an interactive digital interface that allows users to sift through the menu based on their nutritional preferences, including features for filtering and sorting.
- Integrating Consumer Feedback: Gather user input on the prototype and refine its functionality and user interface accordingly to enhance the user experience.

G. Implementation and Launch

- Final Enhancements: Implement final adjustments to the interactive platform based on extensive user feedback, optimizing it for user engagement and efficiency.
- Integration and Rollout: Incorporate the polished tool into Starbucks' primary digital channels (both web and mobile platforms), formally introducing it to provide a tailored, enriching browsing experience.

VII. ANALYSIS AND RESULTS

Fig 1: Displaying few rows of the dataset

	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)	Calcium (% DV)	Iron (% DV)	Caffeine (mg)
0	Coffee	Brewed Coffee	Short	3	0.1	0.0	0.0	0	5	0	0	0	0.3	0%	0%	0%	0%	175
1	Coffee	Brewed Coffee	Tall	4	0.1	0.0	0.0	0	10	0	0	0	0.5	0%	0%	0%	0%	260
2	Coffee	Brewed Coffee	Grande	5	0.1	0.0	0.0	0	10	0	0	0	1.0	0%	0%	0%	0%	330
3	Coffee	Brewed Coffee	Venti	5	0.1	0.0	0.0	0	10	0	0	0	1.0	0%	0%	2%	0%	410
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%	20%	0%	75

The fig 1 displays a data table from a dataset that captures nutritional information of various beverages. It consists of several columns, including "Beverage category," "Beverage type," "Beverage prep," "Calories," "Total 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)," "Calcium (% DV)," and "Iron (% DV)." Each row represents a different beverage, providing details such as the beverage type (e.g., Coffee), its preparation style (e.g., Short, Grande), and its comprehensive nutritional content including macronutrients and micronutrients. The table shows partial data for five beverages, including Coffee and Caffè Espresso, highlighting variations in nutritional values based on preparation style.

Fig 2. Summary Statistics of the numerical attributes

	Calories	Trans Fat (g)	Saturated Fat (g)	Sodium (mg)	Total Carbohydrates (g)	Cholesterol (mg)	Dietary Fibre (g)	Sugars (g)	Protein (g)
count	242.000000	242.000000	242.000000	242.000000	242.000000	242.000000	242.000000	242.000000	242.000000
mean	193.871901	1.307025	0.037603	6.363636	128.884298	35.991736	0.805785	32.962810	6.978512
std	102.863303	1.640259	0.071377	8.639257	62.303223	20.795186	1.445944	19.730199	4.871659
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	120.000000	0.100000	0.000000	0.000000	70.000000	21.000000	0.000000	16.000000	3.000000
50%	185.000000	0.500000	0.000000	5.000000	125.000000	34.000000	0.000000	32.000000	6.000000
75%	260.000000	2.000000	0.100000	10.000000	170.000000	50.750000	1.000000	43.750000	10.000000
max	510.000000	9.000000	0.300000	40.000000	340.000000	90.000000	8.000000	84.000000	20.000000

The fig 2 presents a statistical summary of nutritional data for beverages in a color-coded table format. It includes rows for various statistical metrics like count, mean, standard deviation (std), minimum (min), 25th percentile (25%), 50th percentile (median), 75th percentile (75%), and maximum (max) across columns such as Calories, Trans Fat (g), Saturated Fat (g), Sodium (mg), Total Carbohydrates (g), Cholesterol (mg), Dietary Fibre (g), Sugars (g), and Protein (g). Each cell displays specific values, with a total count of 242 data points for each nutritional category. The table provides a comprehensive overview, allowing for quick insights into the range and distribution of each nutrient within the dataset. The color coding suggests a gradient, indicating higher values with darker shades gradually moving to lighter shades implies lower values.

Fig 3: Attribute corressponding null values

	count	Nulls	nulls%	cardinality
Beverage_category	242	0	0	9
Beverage	242	0	0	33
Beverage_prep	242	0	0	13
Calories	242	0	0	48
Total Fat (g)	242	0	0	24
Trans Fat (g)	242	0	0	18
Saturated Fat (g)	242	0	0	4
Sodium (mg)	242	0	0	9
Total Carbohydrates (g)	242	0	0	51
Cholesterol (mg)	242	0	0	75
Dietary Fibre (g)	242	0	0	8
Sugars (g)	242	0	0	70
Protein (g)	242	0	0	26
Vitamin A (% DV)	242	0	0	11
Vitamin C (% DV)	242	0	0	10
Calcium (% DV)	242	0	0	14
Iron (% DV)	242	0	0	18
Caffeine (mg)	242	1	100	36

The fig 3 is a table that displays statistical data on various nutritional attributes of beverages. It includes columns like 'count', 'Nulls', 'nulls%', and 'cardinality' across multiple nutritional categories such as calories, fats, vitamins, and minerals for 242 beverage entries. Most categories show zero null values except for caffeine, which has one null value representing 100% of its entries. The table also indicates the variety in beverage types, preparations, and nutritional content through its 'cardinality' values.

Fig 4: Table displaying the beverage having at least 80 gm of sugars

	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)	Calcium (% DV)	Iron (% DV)	Caffeine (mg)
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%	0%	0%	0%	0
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%	0%	20%	35%	145
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%	0%	15%	35%	145

The fig 4 displays a data table from a beverage dataset, specifically filtered to show entries where the sugar content exceeds 80 grams. This table shows values of three beverage one "Caramel Apple Spice (Without Whipped Cream)" from the "Signature Espresso Drinks" category, and two "Java Chip Frappuccino® Blended Coffee" variations from "Frappuccino® Blended Coffee" category, differing only in the type of milk used (nonfat milk and whole milk, respectively). Each row includes nutritional details such as calories, total fat, trans fat, saturated fat, sodium, total carbohydrates, cholesterol, dietary fiber, sugars, protein, vitamins A and C, calcium, iron, and caffeine content. This table provides a overview of higher-sugar beverage options within the dataset, showcasing substantial variations in nutritional content based on minor recipe adjustments.

Fig 5: Table depicts the summary of beverage having minimum 25 gm of sugars and atleast 5gms of protein

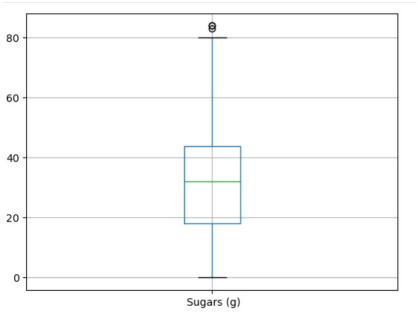
	Calories	Trans Fat (g)	Saturated Fat (g)	Sodium (mg)
count	39.000000	39.000000	39.000000	39.000000
mean	131.025641	1.094872	0.041026	8.076923
std	39.589663	1.111175	0.075107	8.930682
min	60.000000	0.100000	0.000000	0.000000
25%	100.000000	0.200000	0.000000	0.000000
50%	130.000000	0.500000	0.000000	5.000000
75%	150.000000	1.500000	0.100000	12.500000
max	240.000000	4.500000	0.300000	35.000000

	Total Carbohydrates (g)	Cholesterol (mg)	Dietary Fibre (g)
count	39.000000	39.000000	39.000000
mean	107.435897	17.897436	0.512821
std	41.531088	6.012246	0.683328
min	55.000000	9.000000	0.000000
25%	77.500000	12.500000	0.000000
50%	95.000000	18.000000	0.000000
75%	127.500000	23.000000	1.000000
max	220.000000	29.000000	2.000000

	Sugars (g)	Protein (g)
count	39.000000	39.000000
mean	15.205128	8.641026
std	5.620355	2.906320
min	5.000000	6.000000
25%	10.000000	6.000000
50%	16.000000	8.000000
75%	20.000000	10.000000
max	24.000000	16.000000

The fig 5 shows a statistical summary table of a filtered subset of beverage data where each entry has less than 25 grams of sugars and more than 5 grams of protein. The table covers various nutritional metrics such as calories, trans fat, saturated fat, sodium, total carbohydrates, cholesterol, dietary fiber, sugars, and protein for 39 beverage entries. Notable statistics include an average calorie count of approximately 131.03, a mean of 15.20 grams of sugars, and about 8.64 grams of protein per beverage. Other key figures are the maximum protein content at 16 grams and the highest sugar amount at 24 grams. This summary helps in understanding the nutritional profile of beverages that are relatively moderate in sugars but higher in protein.

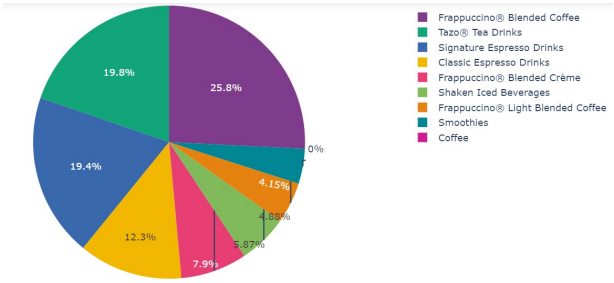
Fig 6: Table depicts the summary of beverage having minimum 25 gm of sugars and at least 5gms of protein



The fig 6 shows a box plot representing the distribution of sugar content (in grams) across a selection of beverages. The plot provides key statistical

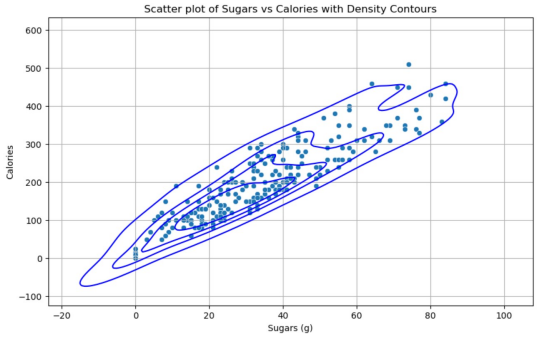
insights: the median sugar content, indicated by the green line within the box, appears to be around 30 grams. The box itself, representing the interquartile range (IQR), spans approximately from 20 to 40 grams, showing the middle 50% of the data. The whiskers extend from about 10 grams to just over 60 grams, indicating the range for the bulk of the data, while an outlier is observed near 80 grams, suggesting an unusually high sugar content compared to the rest. This box plot is useful for understanding the central tendency and variability of sugar content in the dataset.

Fig 7: Distribution of Sugar Intake vs Beverage Category



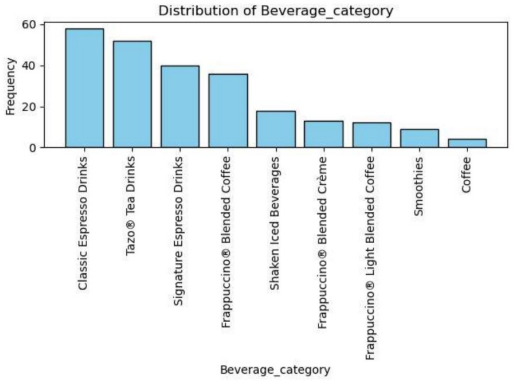
The fig 7 depicts a pie chart that illustrates how various beverage categories are distributed according to a given factor, such as popularity or probable sales. Frappuccino® Blended Coffee holds the biggest market share at 25.8%, with Tazo® Tea Drinks coming in second at 19.8%. The figure also shows the proportions of the other categories, which are Signature Espresso Drinks, Classic Espresso Drinks, Frappuccino® Light Blended Coffee, Shaken Iced Beverages, Smoothies, and Coffee. The categories are distinguished by various shades in the pie chart, making a visual comparison simple.

Fig 8: Relationship between sugar content and calories



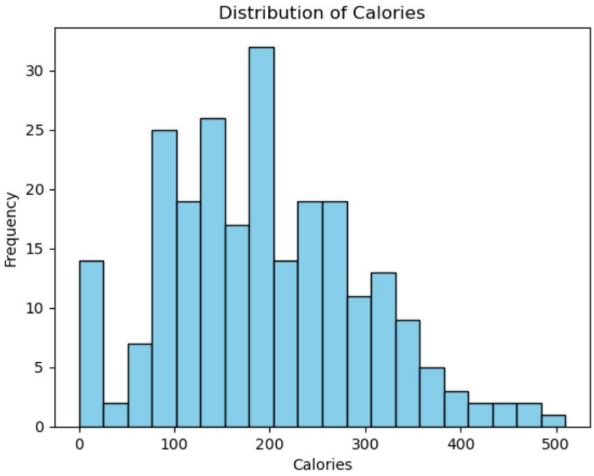
The fig 8 shows a scatter plot graph illustrating the relationship between the sugar content (in grams) of beverages and their caloric content. The x-axis represents sugars ranging from -20 to over 100 grams, and the y-axis represents calories ranging from -100 to 600. Each dot on the plot represents a beverage, with the distribution suggesting a positive correlation between the amount of sugar and calorie content. Concentric ellipses centered around clusters of data points indicate areas of higher density, showing that most beverages cluster within specific sugar and calorie ranges, with the majority having relatively moderate sugar and calorie values, but with some outliers having significantly higher values in both metrics.

Fig 9: Distribution of Beverage Category



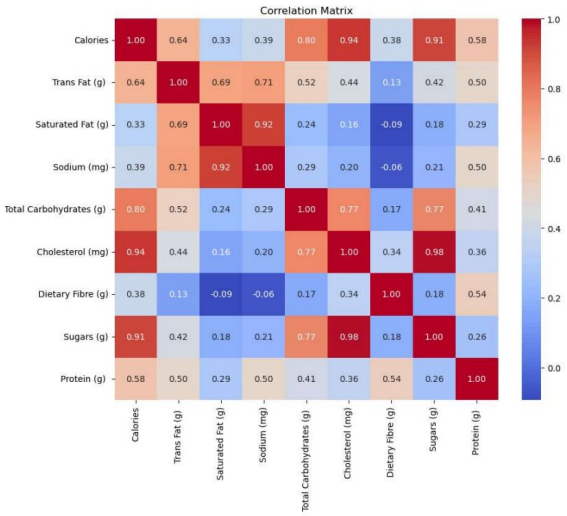
The fig 9 depicts a bar chart representing the distribution of different beverage categories. Classic Espresso Drinks dominate the chart, indicating the highest frequency among the categories listed. This is followed by Tazo® Tea Drinks and Brewed Tea Drinks, showing a slightly lower yet substantial presence. The chart progresses to display categories like Signature Espresso Drinks, Frappuccino® Blended Coffee, Shaken Iced Beverages, Frappuccino® Light Blended Coffee, and Smoothies with decreasing frequencies. The category "Other" has the fewest entries, suggesting a minor variety of beverages falling outside the main specified categories. This visualization provides a clear overview of the popularity and availability of various types of beverages in the dataset.

Fig 10. Distribution of Calories



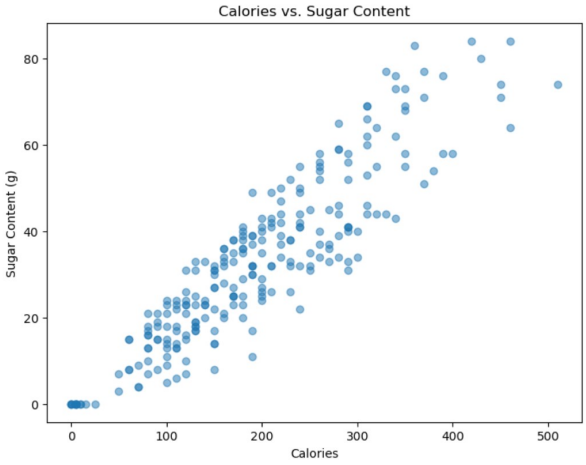
The distribution of calories in a dataset. The x-axis indicates the calorie count, which ranges from 0 to 500 calories, while the y-axis reflects the number of observations within each calorie range. The bars represent the amount of things that fall into various calorie categories. There appears to be a concentration of goods between 100 and 300 calories, with the largest frequency in the 200 to 250 calorie range. The distribution appears to be broadly unimodal, with a peak around 200-250 calories and tapering off as the calorie count approaches 500 as shown in fig 10.

Fig 11. Correlation Matrix



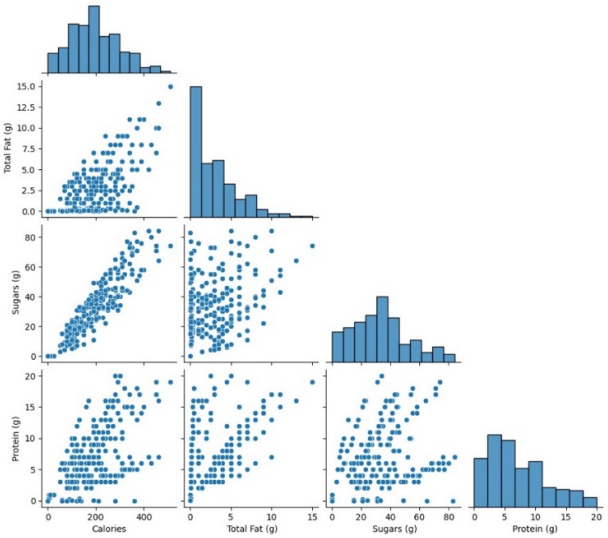
The fig 11 shows correlation matrix heatmap depicted in the image illustrates the connections between a number of nutritional components, including calories, dietary fiber, sugars, sodium, total carbs, cholesterol, and saturated and total fat. The correlation coefficient, which ranges from -1 to 1, is provided by each cell in the matrix and indicates the strength of the relationship between two variables. Positive correlations are represented by red tones, negative correlations by blue tones, and the degree of the correlation is shown by the color's intensity. Perfectly positively correlated cells have a correlation of 1, and they are usually located along the diagonal, where each variable correlates with itself.

Fig 12. Relationship between Calories and Sugar Intake



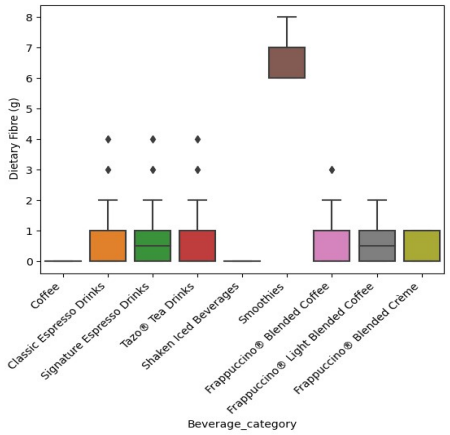
The fig 12 displays a scatter plot titled "Calories vs. Sugar Content," illustrating the relationship between the sugar content (in grams) and calorie content of various beverages. The plot features data points distributed over a calorie range of 0 to 500 and a sugar content range of 0 to 80 grams. The majority of the points tend to cluster in the lower calorie and sugar ranges, but a noticeable trend shows that as the calorie content increases, the sugar content generally increases as well. This indicates a positive correlation between the amount of sugar in beverages and their caloric value, suggesting that beverages with higher sugar content tend to be higher in calories.

Fig 13. Pair Plot Correlational Analysis of Nutritional Metrics: Calorie, Total Fat, Sugar, & Protein



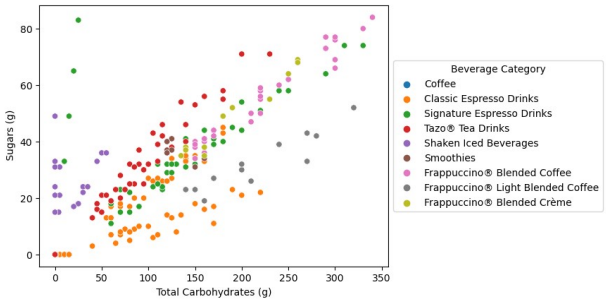
This fig 13 depicts graph is known as a pair plot or scatterplot matrix, displays a scatter plot and histogram matrix. It illustrates the relationships between the various nutritional elements—calories, total fat, carbohydrates, and protein. The distribution histograms for each individual variable are displayed on the diagonal. Scatter plots illustrating the relationship between two variables are displayed in the off-diagonal plots. In this case, a positive relationship may be seen in the scatter plot of calories and total fat, indicating that total fat tends to rise along with calories. In statistics and data analysis, this type of visualization is commonly employed to investigate possible relationships and distributions of continuous variables.

Fig 14. Dietary Fiber Content of Each Beverage Category



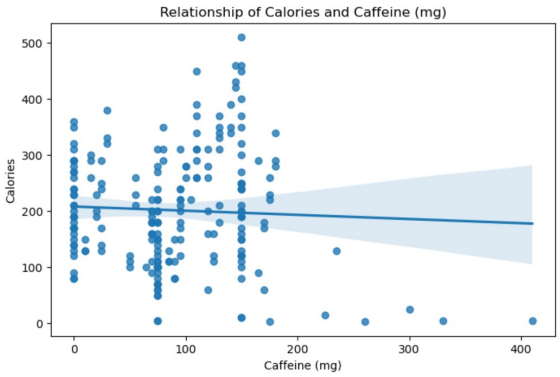
A box plot comparing the dietary fiber content of various beverage categories, including coffee, classic espresso drinks, signature espresso drinks, Tazo® tea drinks, shaken iced beverages, smoothies, Frappuccino® blended coffee, Frappuccino® light blended coffee, and Frappuccino® blended crème. Each box plot depicts the distribution of dietary fiber within a category, with the central line denoting the median, the box edges reflecting the interquartile range (25th to 75th percentile), and the whiskers extending to the farthest points that are not considered outliers. Outliers are represented as individual points. The Tazo® Tea Drinks category appears to have the highest median dietary fiber content, while the other categories have lower medians, with some variation and outliers as shown in fig 14.

Fig 15. Scatter Plot for Total Carbohydrates vs. Sugars



The fig 15 shows correlation between sugar content (grams) and total carbohydrates (grams) in various beverage categories. Color-coded beverages include coffee, classic espresso drinks, signature espresso drinks, tazo® tea drinks, shaken iced beverages, smoothies, frappuccino® light blended coffee, and frappuccino® blended coffee. The point distribution indicates a positive correlation between sugar concentration and total carbohydrates across beverage categories, as expected given that sugars are a form of carbohydrate. The beverages in each category are depicted as clusters of dots, with some variance in sugar and carbohydrate content both within and between categories.

Fig 16. Relationship of Calories and Caffeine (mg)



The fig 16 illustrates a scatter plot titled "Relationship of Calories and Caffeine (mg)," which maps the correlation between caffeine content (in milligrams) and the caloric values of various beverages. The x-axis represents caffeine content ranging from 0 to 400 mg, while the y-axis depicts calories ranging from near 0 up to 500. The plot points are densely clustered, especially around lower caffeine levels, with a noticeable spread as caffeine content increases. A trend line is included in the graph, accompanied by a shaded area representing the confidence interval, which shows a slight upward trend, suggesting a weak positive correlation between caffeine content and calorie count. This implies that beverages with higher caffeine content might tend to have slightly higher caloric values, but the correlation is not strongly pronounced.

IX. LIMITATIONS

There are three critical challenges for this paper analyzing the nutritional content of Starbucks menu items:

1. Completeness and Accuracy of Data: The available nutritional data might not encompass all items or their variations found at Starbucks, especially those that are seasonal or specific to certain regions. Additionally, any inaccuracies or outdated figures in the data could distort the analysis, thus undermining the credibility of any conclusions regarding the nutritional implications.

2. Applicability of Results: The insights obtained from the analysis of Starbucks' menu items may not be transferable to other cafes or restaurant chains, which could have distinct nutritional criteria or different menu offerings. This restricts the scope for generalizing the study's findings across the broader fast-food sector or other food service establishments.

3. Consistency in Preparation and Portion Sizes: The nutritional values can differ markedly based on the method of preparation and service at various Starbucks outlets, influenced by differences in ingredient types and amounts, as well as any extra toppings or condiments that may be added. Such variations can result in inaccuracies when estimating the true nutritional consumption of customers, potentially leading to misguided advice for consumers and policy recommendations deriving from this research.

X. FUTURE SCOPE

Creating interactive digital tools, such as calorie calculators integrated with menu items, provides customization options for people with dietary preferences. These tools can be included in applications or websites, enhancing user engagement and meeting the specific needs of individuals seeking nutritional information tailored to their dietary requirements.

Expanded menu selections, such as vegan, gluten-free, and keto diets, increase the number of health-conscious options available to accommodate various dietary requirements. This also includes offering seasonal menus that emphasize fresh and local ingredients.

Customer feedback involves implementing a system to gather feedback from customers on new initiatives. By analyzing the insights provided, improvements can be effectively made to the offerings. To facilitate this process, the system could include tools such as surveys or focus groups, ensuring a broad range of customer perspectives is considered.

XI. ACKNOWLEDGEMENT

We are profoundly grateful to Prof. Dr. Ebrima Ceesay for his invaluable mentorship and profound expertise, which have greatly influenced our research and deepened our comprehension of the topic. Our thanks also go to both the Department of Applied Information Technology and the Department of Data Analytics Engineering at George Mason University for their essential support and resources that made our research possible. We appreciate the constructive feedback from our peers and colleagues, which has substantially improved the manuscript's quality. Finally, we would like to thank Data.world for granting us access to the "starbucks_drinkMenu_expanded.csv" dataset, a crucial resource in our nutritional analysis of Starbucks' menu items.

XI. CONCLUSION

The visualizations show that different types of beverages include varying amounts of nutrients such as sugar, carbs, dietary fiber, and calories. Furthermore, certain beverage categories, such as Frappuccinos®, appear to have higher sugar and carbohydrate content, which corresponds to higher caloric content, whereas Tazo® Tea Drinks may include more nutritional fiber. Overall, these visual data insights can help customers make informed decisions based

on the nutritional composition of beverages.

By including more health and dietary related options to the menu and introducing interactive nutritional tools, Starbucks is enhancing its commitment to good health and making everything transparent. Starbucks can keep improving and adapting to suit the changing requirements of its customers by putting in place a strong system for collecting and analyzing client feedback. In addition to boosting consumer confidence, this strategic focus establishes Starbucks as a icon in the food and beverage sector when it comes to promoting health and wellness.

XII. REFERENCES

[1] C. Huang, T. Dumanovsky, L. D. Silver, C. Nonas, and M. T. Bassett, "Calories from beverages purchased at 2 major coffee chains in New York City, 2007," Preventing Chronic Disease, vol. 6, no. 4, p. A118, Oct. 2009, Accessed: Mar. 30, 2024. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/19754994/

[2] Bollinger, Bryan, et al. Calorie Posting in Chain Restaurants. 1 Jan. 2010, https://doi.org/10.3386/w15648 . Accessed 23 May 2023.

[3] Wu, A. & Mission San Jose High School. (2018). Starbucks Coffee statistical analysis. In Proceedings of the International Conference on Industrial Engineering and OperationsManagement. https://www.ieomsociety.org/paris2018/papers/424.pdf

[4] M. Kim, B. -T. Zhang and J. -S. Lee, "Dynamic and Static Influence Models on Starbucks Networks," 2009 International Conference on Advances in Social Network Analysis and Mining, Athens, Greece, 2009, pp. 344-347, doi:10.1109/ASONAM.2009.50.

[5] M. Kim, B. -T. Zhang and J. -S. Lee, "Social Influence Models Based on Starbucks Networks," 2009 International Conference on Computational Aspects of Social Networks, Fontainebleau, France, 2009, pp. 3-9, doi: 10.1109/CASoN.2009.26.

[6] R. Sarki and Y. -M. Kwon, "Towards buying experience with things: Experimental study @Starbucks," 2015 17th International Conference on Advanced Communication Technology (ICACT), PyeongChang, Korea (South), 2015, pp. 267-270, doi: 10.1109/ICACT.2015.7224800.

[7] Han, X., Ren, J., Caloric Restriction and Heart Function: Is There a Sensible Link?, NCBI, Available: https://ww.ncbi.nlmnih.gov/pmc/articles/PMC4002317/

[8] Jephcote, B., The More Carbs You Eat, the Higher the Risk of Heart Disease, Diabetes.co, Available: https://www.diabetes.co.uk/in-depth/carbs-higher-risk-heart-disease-states-leading-cardiologist-dr-salimyusuf/ February 20, 2017.

[9] Shmerling, R., Sodium Studies Blur the Picture on What is Heart Healthy, Harvard Health Publishing, Available: https://www.health.harvard.edu/blog/sodium-studies-blur-picture-heart-healthy-20140815736 August15,2014.

[10] Three to Four Cups of Coffee a Day Linked to Longer Life,ScienceDaily, Available https://www.sciencedaily.com/releases/2017/11/171122190659.html November 22, 2017.

[11] C. Huang, T. Dumanovsky, L. D. Silver, C. Nonas, and M. T. Bassett, "Calories From Beverages Purchased at 2 Major Coffee Chains in New York City, 2007," Preventing Chronic Disease, vol. 6, no. 4, p. A118, Sep. 2009, Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2774632/

[12] "Nutrition by the Cup." Available: https://stories.starbucks.com/uploads/2019/01/nutrition-1.pdf

[13] S. Berinato, "Calorie Labeling at Starbucks (727 Words, 6 Minutes to Read)," Harvard Business Review, Feb. 04, 2010. https://hbr.org/2010/02/starbucks-calories

[14] R. J. Avery, J. Cawley, J. Eddelbuettel, M. D. Eisenberg, C. Mann, and A. D. Mathios, "Consumer Responses to Firms' Voluntary Disclosure of Information: Evidence from Calorie Labeling by Starbucks," SSRN Electronic Journal, 2021, doi: https://doi.org/10.2139/ssrn.3893361.

[15] V. Nguyen, O. Alstott, D. Sanders, L. Beveridge, and L. Du, "Health Trends and Their Effects On Consumers' Starbucks Order Decisions INTRODUCTION AND BACKGROUND." Accessed: May 08, 2024. [Online]. Available: https://theolivealstott.wordpress.com/wp-content/uploads/2019/06/health-trends-and-their-effects-on-consumerse28099-starbucks-order-decisions.pdf

[16] "Healthiest Starbucks Drinks and Food," Verywell Fit. https://www.verywellfit.com/starbucks-how-to-order-low-calorie-coffee-drinks-3495691

[17] N. Geereddy, "Strategic Analysis of Starbucks Corporation," 2023. Available: https://scholar.harvard.edu/files/nithingeereddy/files/starbucks_case_analysis.pdf

[18] Starbucks, “Timeline,” 2020. Available: https://stories.starbucks.com/uploads/2019/01/AboutUs-Company-Timeline-1.6.21-FINAL.pdf

[19] Starbucks, “Timeline,” 2020. Available: https://stories.starbucks.com/uploads/2019/01/AboutUs-Company-Timeline-1.6.21-FINAL.pdf