



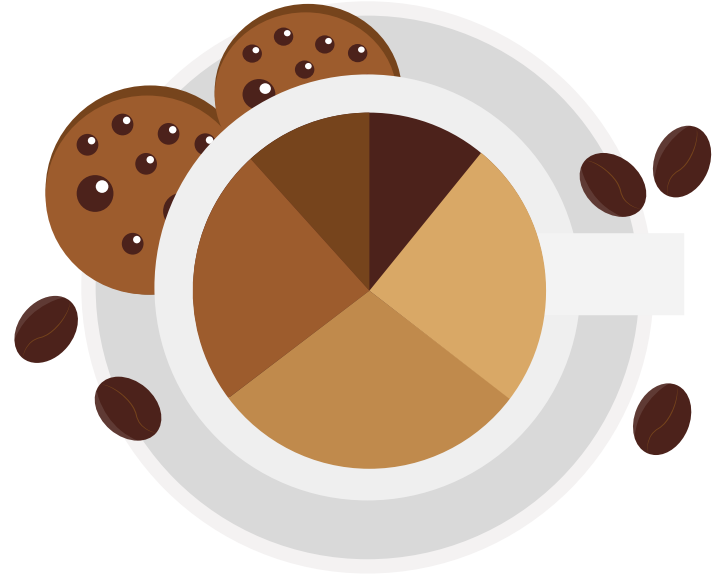
Stats on Starbucks

By
Sydney, Sindhu,
John, and Zane

Introduction

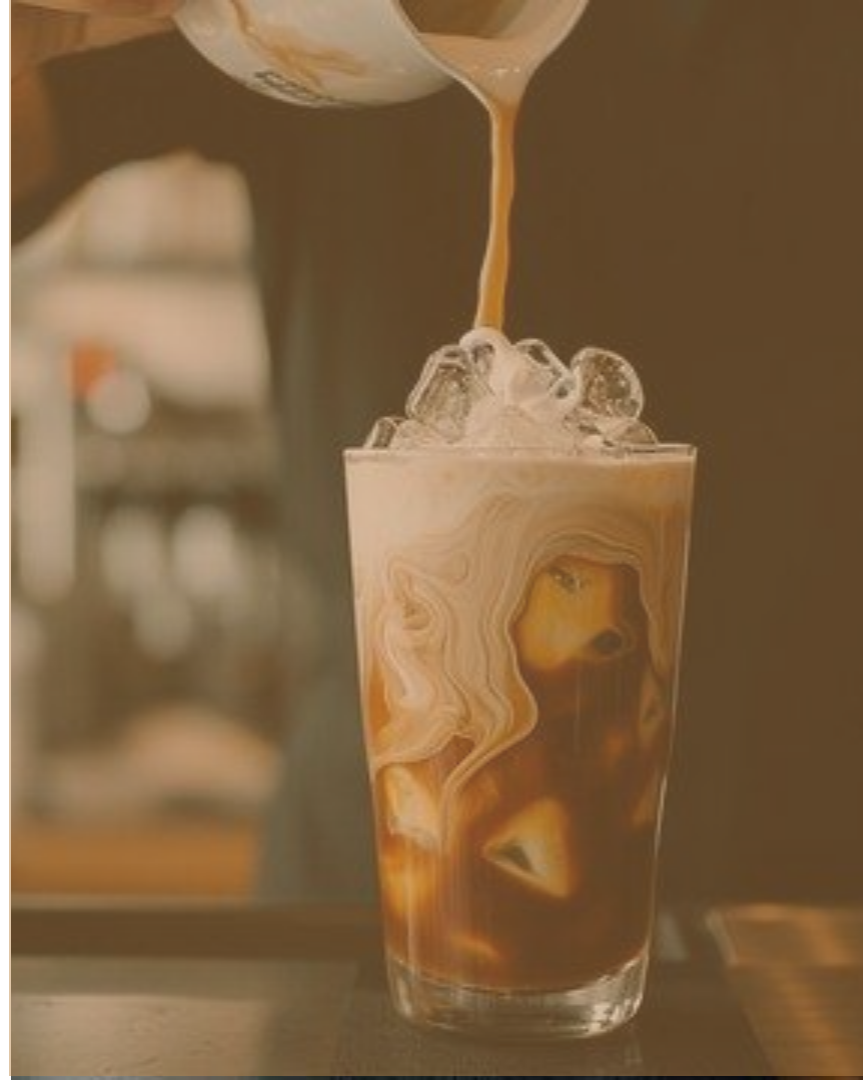
Choosing a topic:

- ❖ Part of our daily routine
- ❖ Relation to college life
- ❖ Caffeinated drinks and health
- ❖ Starbucks as a leader in the coffee market



The Data Set

- ❖ Google Dataset Search
- ❖ Uploaded on Kaggle by Henry Shan
- ❖ Collected in 2024, updated in 2025
- ❖ Variables:
 - Beverage Category
 - Beverage (name)
 - Calories
 - Total Fat (g)
 - Trans Fat (g)
 - Saturated Fat (g)
 - Sodium (mg)
 - Total Carbohydrates (g)
 - Cholesterol (mg)
 - Beverage Size



The Question

When looking at the dataset, we were deciding between two questions:

- Option 1: How does the beverage preparation way such as iced, blended and different types of milk affect the calorie and fat content of the drink?
- **Option 2: Which nutritional factors such as trans fat or sodium are the most indicative of the calorie content of the drinks?**

Why we settled on Option 2:

- ❖ Visualize correlation with linear trend lines on scatterplots
- ❖ Compare the strength of variable relationships
- ❖ Examine four nutrition facts as the independent variable
- ❖ Calories as the dependent variable
- ❖ Determine strongest indicator of calorie count by comparing correlation values



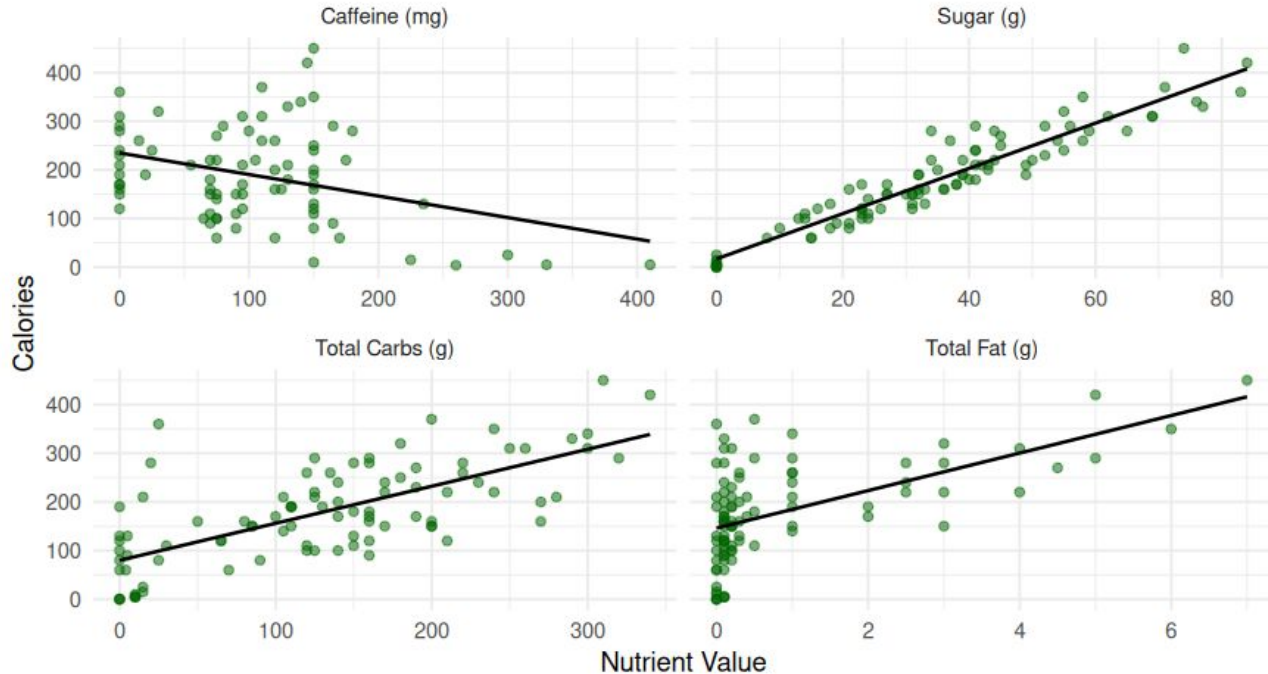
Cleaning the Data

- ❖ Beverage_prep to beverage_size
- ❖ Mutate columns
- ❖ Str_detect sizes
- ❖ Filter () out Tall, Grande, Venti
- ❖ Select(- beverage_prep)

Scatter Plots and Trend Lines

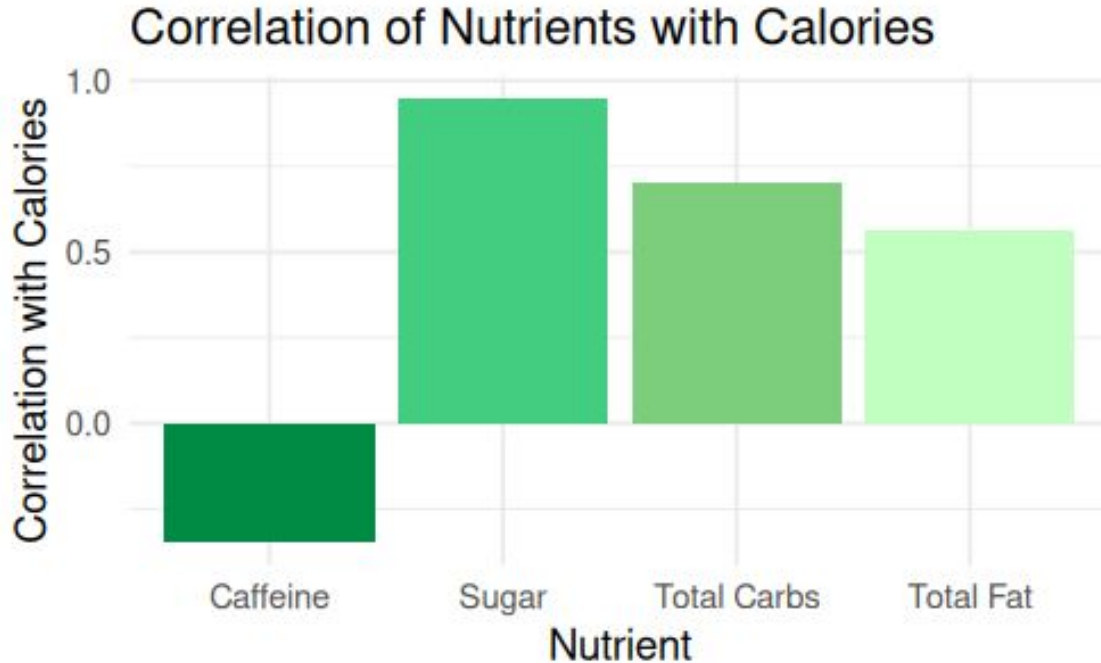
Relationship Between Nutritional Factors and Calories

Starbucks Beverage Menu



Factor	Correlation Values
Caffeine	-0.349
Sugar	0.948
Total Carbs	0.7
Total Fat	0.565

Correlation Visualization

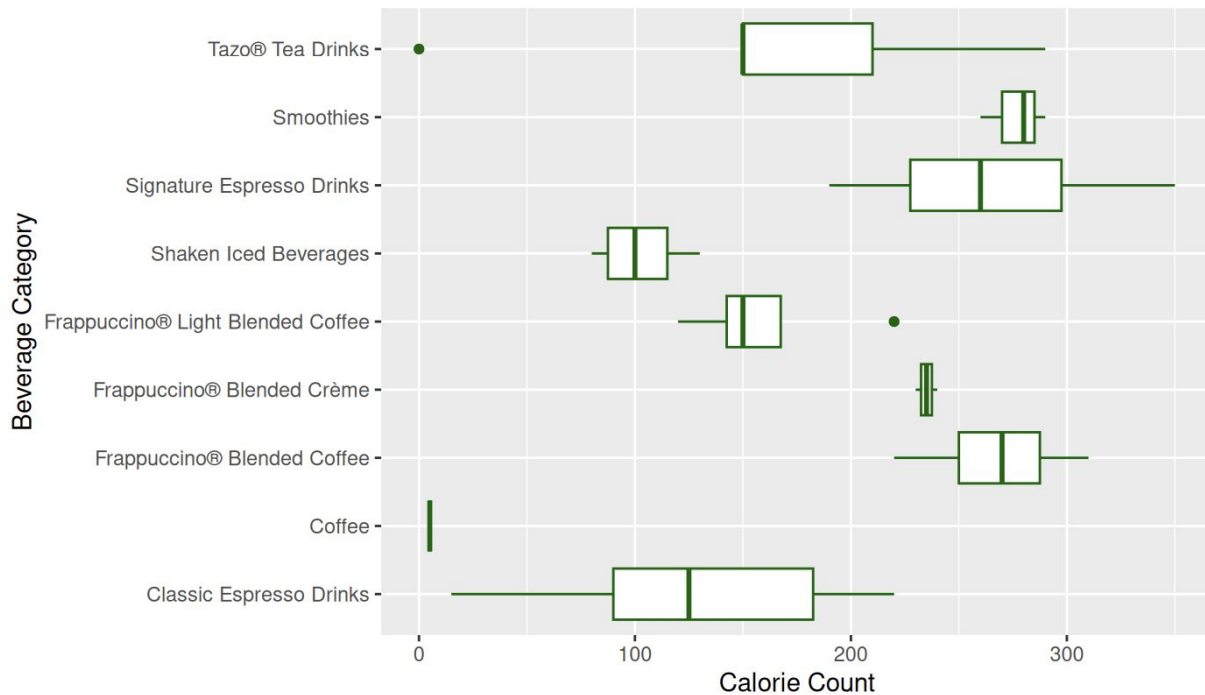


Takeaway:

- ❖ Sugar has the highest correlation coefficient of 0.948
- ❖ A correlation coefficient of 0.8–1.0 indicates a strong relationship
- ❖ Total carbs (0.700) and total fat (0.565) show moderate positive associations
- ❖ Caffeine (-0.349) is the only factor with a negative correlation

Boxplot

Calorie Count by Beverage Category for a Grande



beverage_category	mean_calorie_count
<chr>	<dbl>
Classic Espresso Drinks	127.5000
Coffee	5.0000
Frappuccino® Blended Coffee	267.5000
Frappuccino® Blended Crème	235.0000
Frappuccino® Light Blended Coffee	160.0000
Shaken Iced Beverages	102.5000
Signature Espresso Drinks	265.0000
Smoothies	276.6667
Tazo® Tea Drinks	160.0000

Conclusion

- ❖ Sugar is the strongest indicator of calorie content in Starbucks drinks
- ❖ Total carbs and total fat also positively associated with calories
- ❖ Caffeine is weakly and negatively associated
- ❖ Understanding the relationship guides consumers and public health movements to healthier choices
 - Calorie-dense drinks are driven by sweeteners, syrups, and milk-based ingredients, not caffeine

Limitations

- ❖ Data reflects published menu info, not actual customer orders with customizations
- ❖ Correlation shows association, NOT causation
- ❖ Cleaning process simplified sizes to Tall, Grande, Venti, excluding some drinks
- ❖ Only examined four nutritional factors out of many available
- ❖ Specialty drinks can skew the measurements

