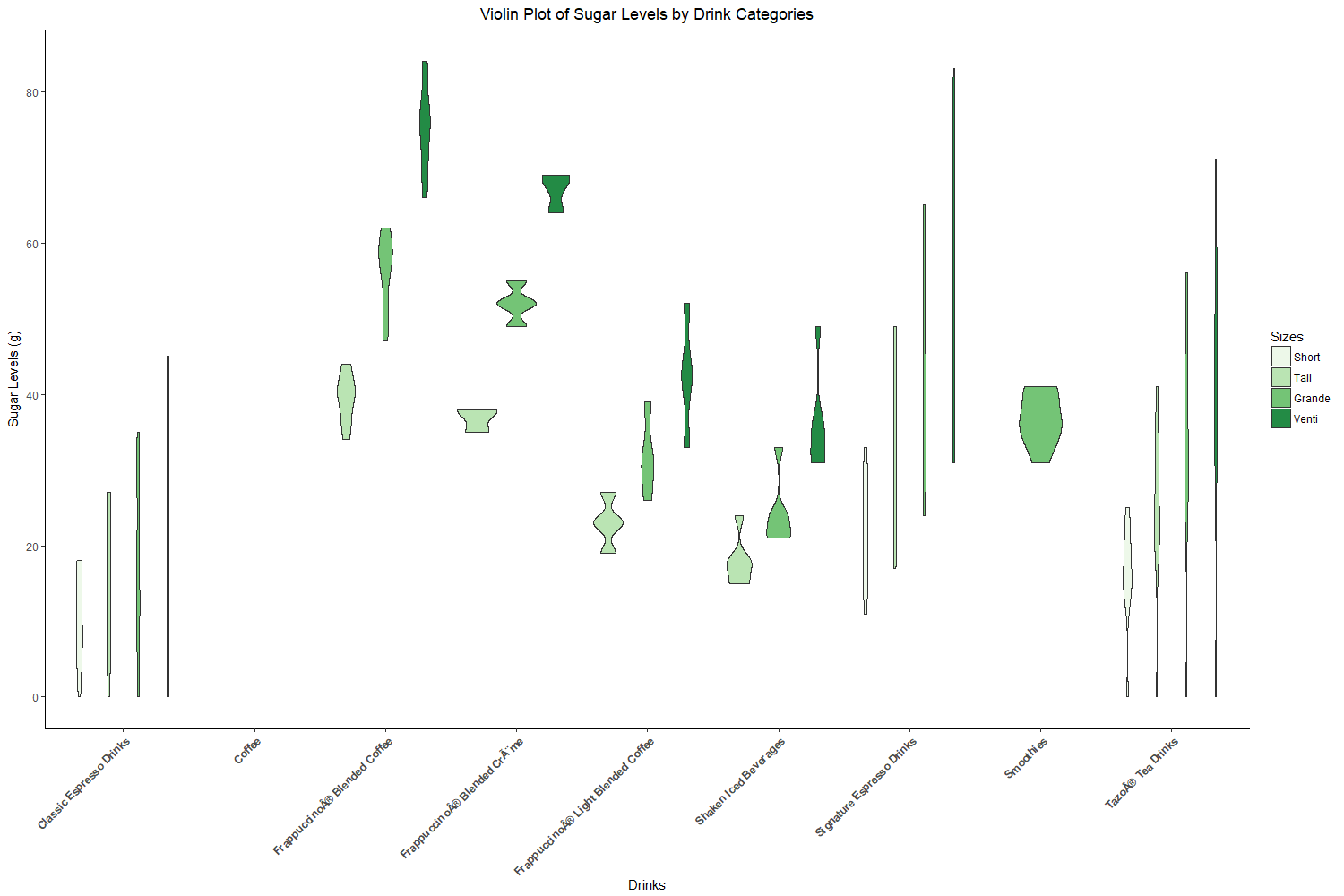
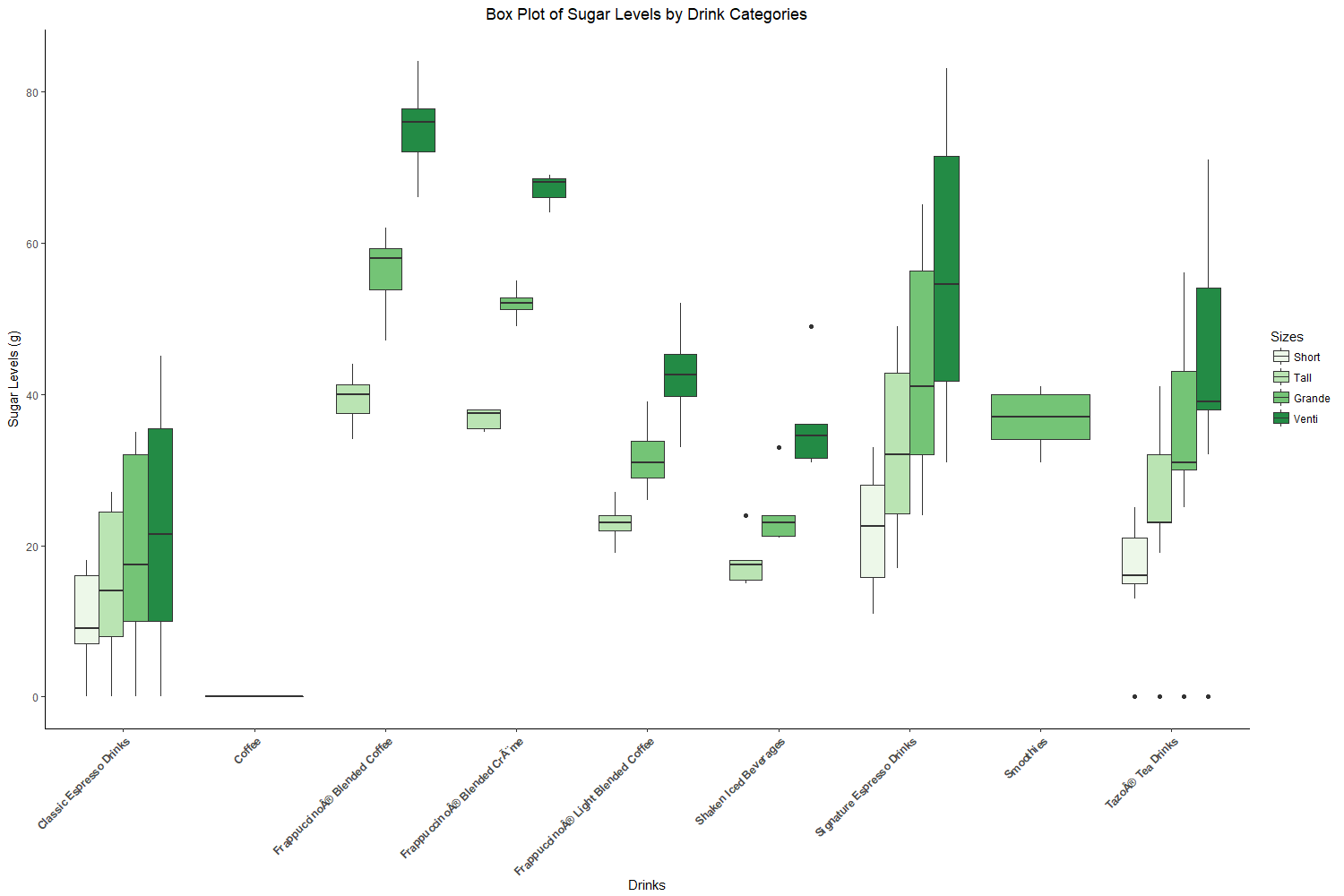
ENSF 619.25: Lab 1





The above plots contain drink data from menu items on a coffee shop, with their corresponding sugar levels. Sugar, measured in grams, was chosen as the target variable for deeper analysis due to its negative impact and public stigma on today’s generation. As we know, sugar is the leading cause of obesity, which leads to diabetes and eventual heart problems. These are some of the reasons why many individuals have begun to avoid ingesting foods with high sugar content to maintain a healthy lifestyle. The data analysis done on this data is targeted at individuals who frequent coffee shops often, such as students, and want to gain a better understanding of the amount of sugar in their favorite drinks.

As can be seen from the violin plot, coffee is the only drink with zero sugar (assuming sugar isn’t added by the customer), making it the favorable drink for people looking to cut sugar from their diet completely. The “Classic Espresso Drinks”, as well as the “Tazo Tea Drinks” have a similar distribution of sugar among its many varieties, with “Tazo Tea Drinks” having the higher overall sugar levels. The ‘violin shapes’ are narrow for both sets, meaning that the number of varieties is distributed relatively evenly. The Classic Espresso Drinks range between 0 to ~40g of sugar per serving, while the Tazo Tea Drinks range between 0 to ~70g of sugar per serving. This also means that customers should be more aware of the sugar level when ordering the Tazo Tea Drinks since some varieties have no sugar, while others can have up to about 70g! The same conclusion can be made for the Classic Espresso Drinks, however because the upper limit is 40g of sugar, it is a relatively sugar-friendly option in comparison. Looking at the boxplot, some of these findings can be analyzed, and corrected even further. The Tazo Tea drink varieties seem to lie mostly in the upper range of 0-70g, with the actual 0g drinks being outliers in the data. This means that although the specific Tazo Tea drinks with 0g of sugar do exist on the menu, but are very few compared to the higher sugar content ones. As a result, unknowingly customers are more likely to choose a Tazo Tea drink in the upper echelon of sugar levels, as compared to the Classic Espresso Drinks, have a much more even distribution, verified further by the boxplot. Similar to the Classic Espresso Drinks, the Signature Espresso Drinks show the same distribution trend, with a narrow violin shape. However, the peak sugar levels hit over 80g for the Venti size.

The next observation that can be made are for the Smoothies, which are only served in the Grande size in this particular coffee shop. The shape of the plot suggests that most of the data lies in the ~35-40g of sugar range, which gives customers a clear idea of exactly how much sugar is in a smoothie, regardless of variety.

For the blended and shaken beverages (Frappaccino Blended Coffee, Frappaccino Blended Cream, Frappacino Light blended Coffee, and Shaken Iced Beverages), they are ordered from highest to lowest relative sugar ranges. For each size, its sugar content has a relatively central value among its varieties, due to the length of the violin shapes being so small in height. For the blended beverages, most varieties lie close to the median sugar content of their respective categories, while the shaken beverages lie closer to the lower end of sugar levels, as shown by the violin shapes being wider towards the bottom. This is also shown by the box plots, where the median line lies relatively closer to the upper quartile as compared to the other drinks.

Overall, the highest sugar content drink categories are the Signature Espresso, and the Frappacino Blended coffee, with peak sugar levels of ~85g of sugar in certain varieties. The sugar levels of the smoothies, as well as the blended and shaken drinks gravitate towards a more central value, so their many varieties will be similar in terms of sugar content. Coffee without added sugar is the only sugar free option on the menu, while the Tea and Espresso drink types have the largest ranges of sugar, making it important for customers to be aware of the sugar content of their particular drink from those categories.

The violin plot was chosen to understand how the sugar contents were distributed among drinks, and to clearly visualize where the highest density of data lied, so that customers could understand where most of the drinks in each category are in terms of sugar level. The box plot serves to correct and verify aspects of the violin plot, and understand where the median value occurred, and whether the majority of datapoints sat above or below it.