**Project 1 – Written Report**

**Kim’s Hypothesis**

There are a total of four flavors that we used from our dataset. We compared these four flavors to their respective rating, effects, location, and breeder. After cleaning the data, we realized that flavor four only had 43 datapoints out of 2351 entries. Our team decided to drop this flavor four dataset due to it being so small.

Another research questions our team posed is if the sweet flavor cannabis has the highest ratings overall. Typically, when looking at flavors you would think that the population would lean towards favoring sweet, flavored products.

We cleaned the data to breakout the flavors into flavor 1,2, and 3 by taking the first listed flavor in the entire column. Those flavors make up the flavor 1 profile. We then repeated the same steps to make the flavor 2 and 3 profiles. This data is what was used for our statistical analysis. The flavors ranged from skunk and pungent to sweet.

Flavor profile 1 ended up having its ratings span across a range 4.1-4.9. There were 47 flavors counted for the first column. The minimum rating went to a flavor called blue and the highest rating was for Tea.

The flavor profiles are large and we thought the cleanest way to visually depict the avg and outliers per flavor would be to use a box plot chart. At a quick glance you can see the mean and median lines pretty much line up across the board and that were a limited number of outliers. The y axis shows the rating and the x axis lists out each flavor.

Flavor profile 2 ended up having its ratings span across a range 4.2-4.6. There were 46 flavors counted for the first column. The minimum rating went to a flavor called apricot and the highest rating was for violet.

Flavor profile 3 ended up having its ratings span across a range 4.2-4.6. There were 48 flavors counted for the first column. The minimum rating went to a flavor called blue and the highest rating was for coffee.

Comparing across all three flavor profiles, the mode and mean values are extremely similar which inclines me to trust comparing values across all the flavors.

We were surprised to find that tea was the highest with coffee closely behind as having the highest ratings. I would not naturally associate cannabis flavors being popular for mimicking beverage flavors, but coffee and tea are drank globally so maybe consumers are accustomed to those flavors. F