**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 question our team posed was that the sweet flavor cannabis has the highest ratings. Typically, when looking at flavors you would think that the population would lean towards favoring sweet, flavored products as seen with popular vape flavors.

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.

**SLIDE**

Flavor profile 1 ended up having its ratings span across a range of 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.

Mean 1: 4.43

Mode 1: 4.4

The flavor profiles are large and we thought the cleanest way to visually depict the datapoints per flavor would be to use a box plot chart. At a quick glance, the boxplot shows the minimum, first quartile, median, third quartile, and maximum values. You can see the 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.

**SLIDE**

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

Mean 2: 4.45

Mode 2: 4.5

**SLIDE**

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

Mean 3: 4.45

Mode 3: 4.4

**SLIDE**

Comparing across all three flavor profiles, the datapoints values are extremely similar with tight ranges, similar modes and means which inclines me to trust comparing values across all the flavors.

**SLIDE**

In conclusion, our hypothesis was incorrect with assuming sweet as having the highest rating. 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.

I would want to group the flavors even more for future research such as grouping them by the five different types of taste our tongue receptors can detect which are salt, sweet, sour, bitter, and umami.

Our next hypothesis will be presented by Lattiana