Sean McLean

ALY 6010

# Module 4 R Practice

<u>Part 1</u>

The first part of the assignment was uploading a data set that showed the bodyweights of 50 male cats and 50 female cats. The question asked was whether the average body weight was the same between the two genders using two sample t-tests. The first step when comparing the means is to separate the two genders into different variables and then summarize them individually. There is a much wider range in bodyweights for males than females and is about on par with the mean of the cat's bodyweights by gender.

When comparing the summaries of both genders and their body weights, there is a noticeable difference in the averages, with male cats being about a half a pound bigger. The biggest male cat is also about a half pound larger than the largest female cat, and when ranking the top 10 cats in terms of bodyweights between genders the differences in size are even bigger than the overall mean. On the other side when looking at the top 10 smallest cats per gender, the differences in size between the genders still lean toward the males but it is a much smaller difference than the overall average. And despite female cats being smaller in size, the one outlier in the dataset when looking at the gender means that stands out is that the smallest female cat is bigger than the smallest male cat.

I incorporated several t-tests in the coding to compare the means of the bodyweights between the two genders. The independent samples t-test and the Welch's t-test calculated the exact same results for the dataset, indicating a high t-score and very small p-value which means that there is a considerable difference in gender means and that if there was a null and alternate hypothesis established on whether there was a difference between means that the null hypothesis would be rejected. The one-sided t-test was used to show if the average bodyweight of male cats was greater than the average bodyweight of the female cats. This test would be conducive to choosing a gender that is believed to be bigger in average bodyweight than the other gender, with males being larger than females as a hypothesis that could be established before the test is done.

The comparison of the cat bodyweight means per gender is shown in a box plot and what sticks out visually is the range of the male which is clearly bigger than the female range. This offers the conclusion from the dataset that male cats overall tend to get larger in average size than female cats and also have a wider range in bodyweight. Both genders do appear to get to about the exact same minimum bodyweight, but male cats on average will have a bigger range in their bodyweight.

Comparison of Cat Bodyweight Means between Genders
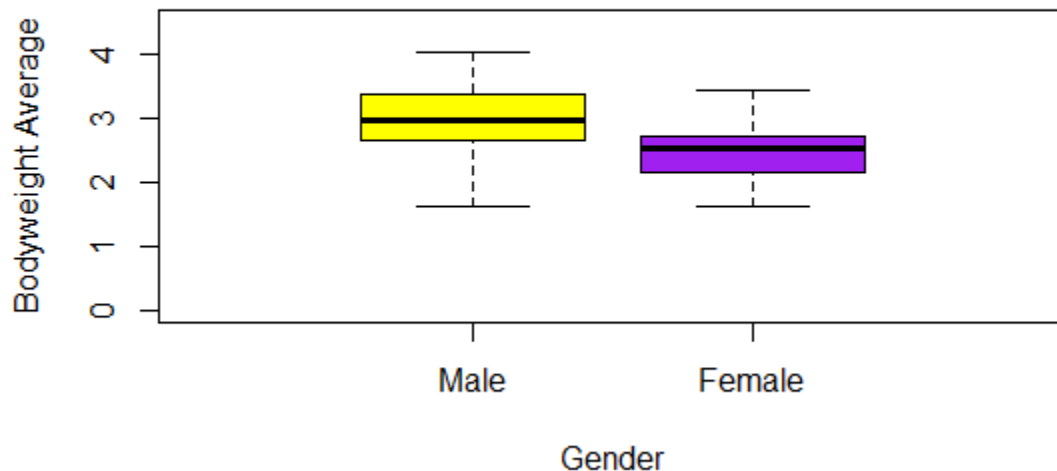
Part 2

The second part contains a dataset with before and after samples from a meditation workshop that looks to improve sleeping quality. Hypotheses for this dataset are created as instructed before looking at the data for analysis.

Null Hypothesis: The average sleep quality score before the workshop is the same as the average sleep quality score following the workshop.

Alternate Hypothesis: The average sleep quality score before the workshop is not the same as the average sleep quality score following the workshop.

Variables are created for the before scores and after scores that will show if there are any differences between the two samples. When performing t-tests on the two samples, there is substantial evidence that from the low t-statistic values and high p-values that there is not enough evidence to reject the null hypothesis. The paired t-test when changing the significance levels from 0.05 to 0.1 also showed no differences in whether the alternate hypothesis should be favored, with the p-value and t-statistic not changing in their values.

The averages of the sleep quality scores following the workshop are higher than the sleep quality score before the workshop, but only by 0.62. This could indicate that this is not a large enough improvement to conclude that the null hypothesis should be rejected for the alternate hypothesis. The t-tests conducted provide the support necessary to show this improvement is not large enough and that the meditation workshop doesn't improve sleep quality scores. Using the paired samples t-test is the best way to show the means between the two samples because of the simplicity of the data set with only 10 scores in the before and after test that shows the means

between the pairs. The results of the other t-tests were similar to the paired samples t-test, so using this t-test made the most sense. The bar chart below provides a visual look to how little difference there is between the two samples.

Sleep Quality From Meditation
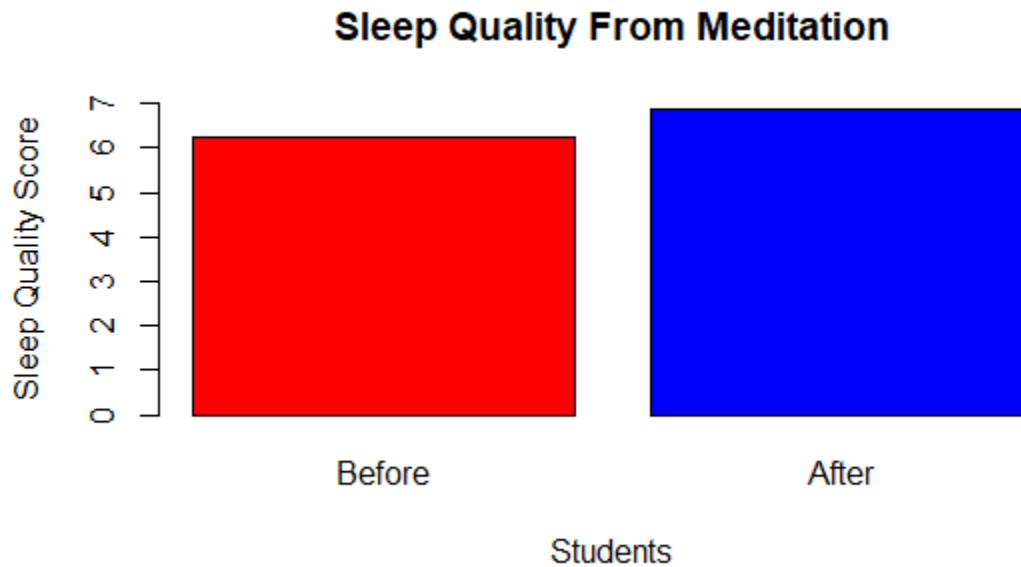
References:

Bluman, A. G. (2017). Elementary Statistics: A Step-by-Step Approach. 10th edition. McGraw-Hill Education.

Chat GPT. (2023, December 1st). Default (GPT 3.5).