*Tables, boxplots, and hypothesis testing*

**R**

**Practice**

Rp

ALY6010 Probability Theory and Introductory Statistics

Module 4 R Practice

**PREPERATION:**

By: John DiSessa

For: Professor Goulding

On: March 28th,2021

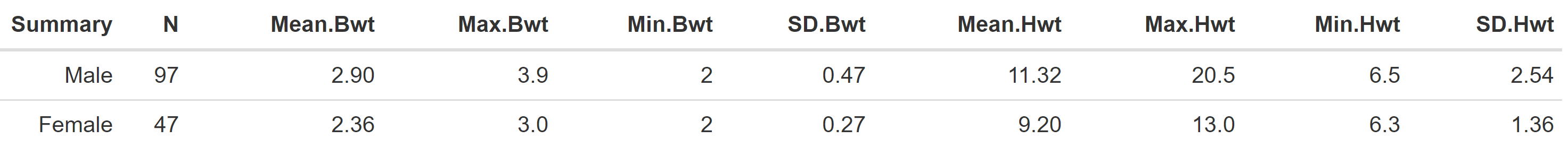
Part One

Introduction

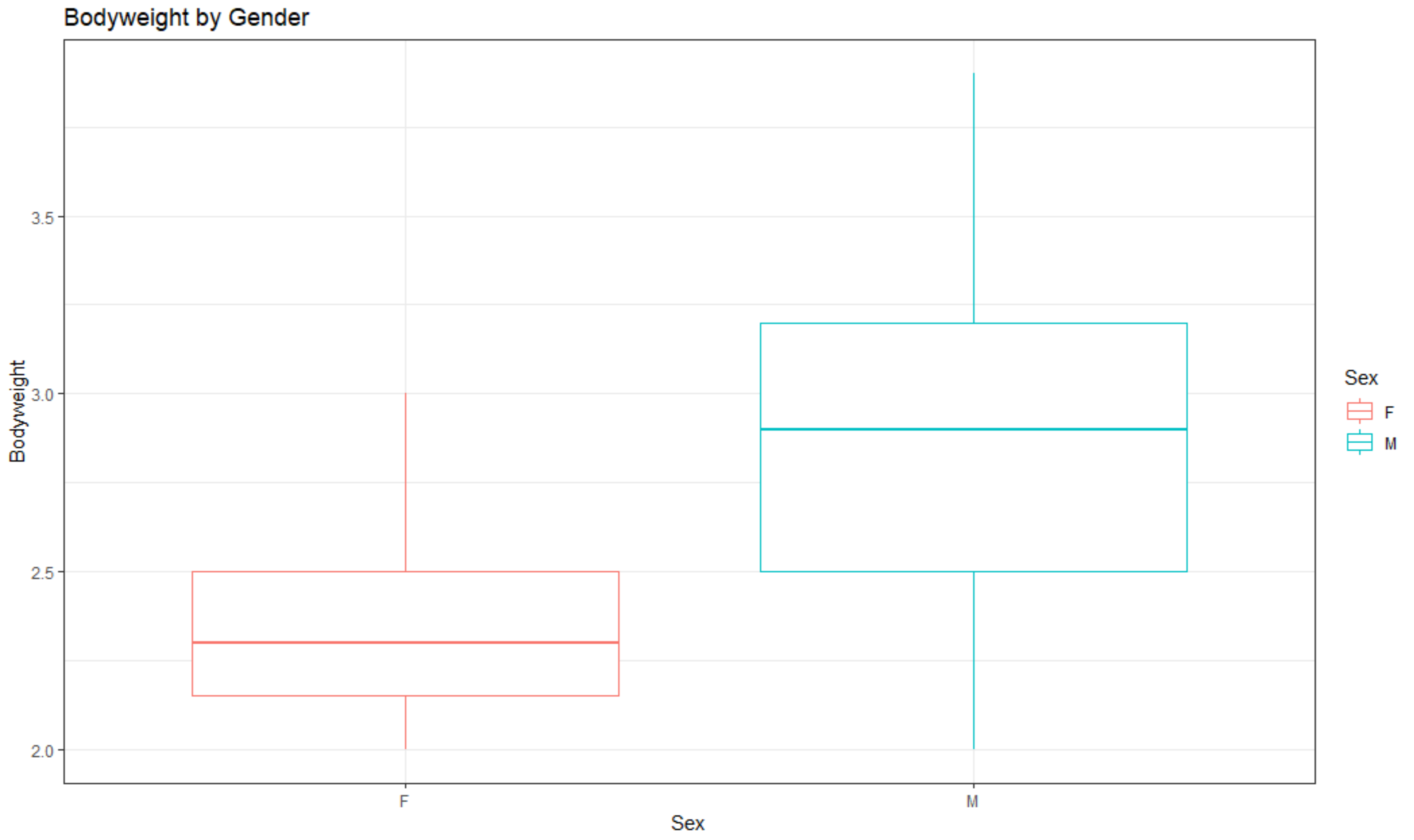
After downloading and installing the MASS package in R, I opened the cats dataset to see the different bodyweights and heights for male and female cats. I wanted to know if male and female cats have different bodyweights.

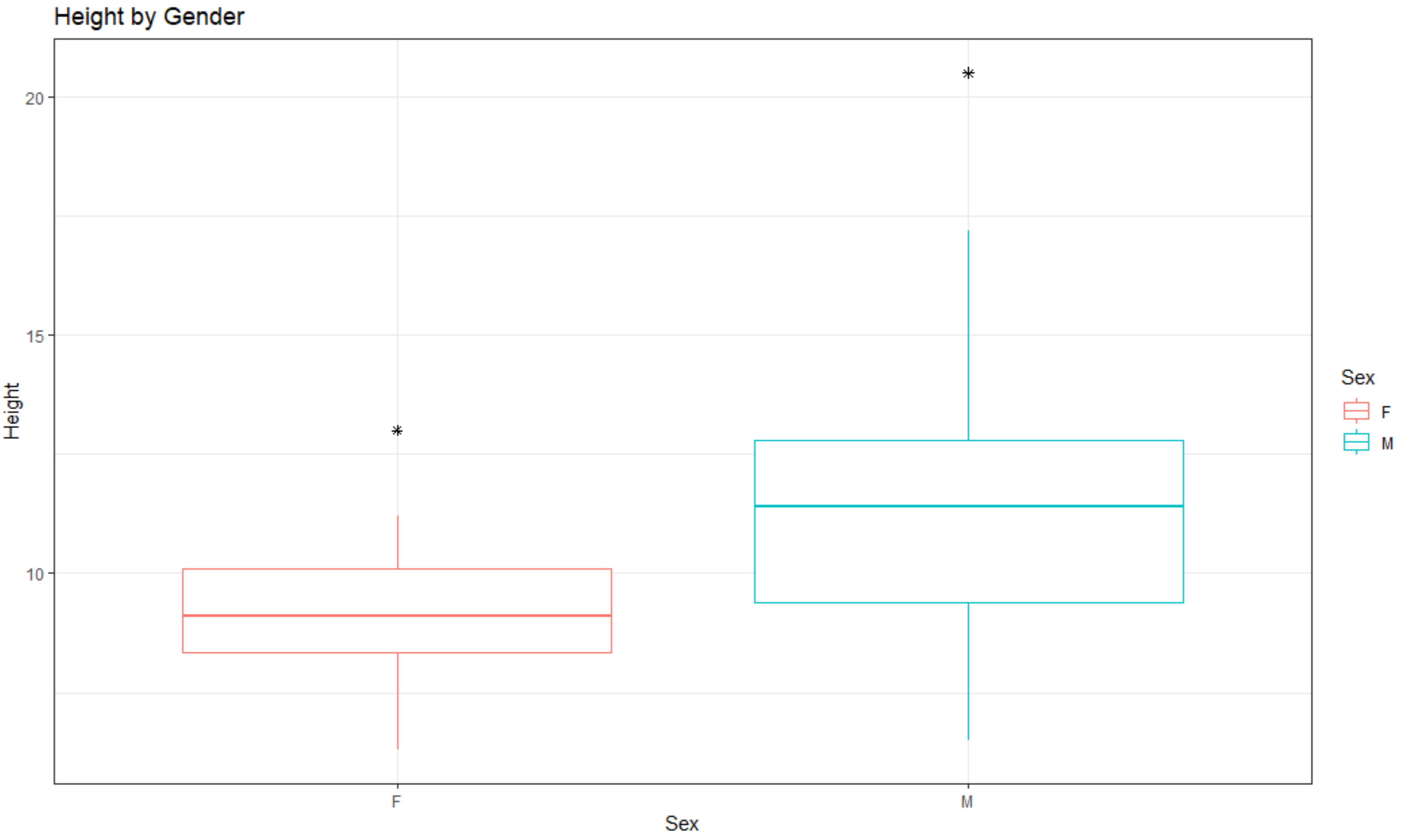
Analysis

Before I moved further, I created subsets of only male and only female cats in order to simply calculate the descriptive statistics, which is summarized is the table below.



As you can see, males have a larger mean bodyweight and height than females. However, the standard deviations are also higher so there is a bigger spread in male weights and heights. The boxplots below visually display these differences. Both males and females have similar minimum values, but males have larger maximum values. I calculated the medians of male and female weights and heights to see if the data was skewed by these larger values. However, the medians basically equaled the means so both male and female data seemed normally distributed.

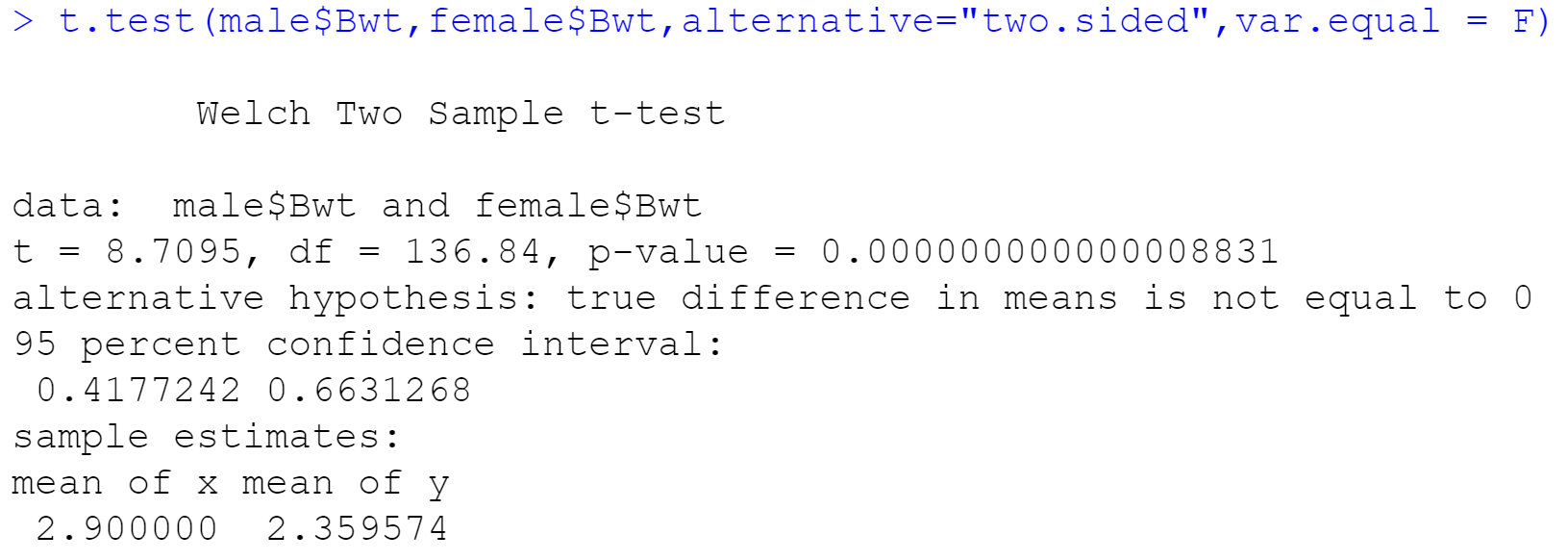




Next, I conducted hypothesis testing at 95% confidence levels (α = .05) to see if the perceived differences in mean weight and mean height are statistically significant. First, I ran an unpaired two sample t-test to see if the mean weight is the same or different between males and females. I chose the unpaired (independent) test because the sample values from males and females are separate, independent groups.

H0: µ1 = µ2

Ha: µ1 ≠ µ2

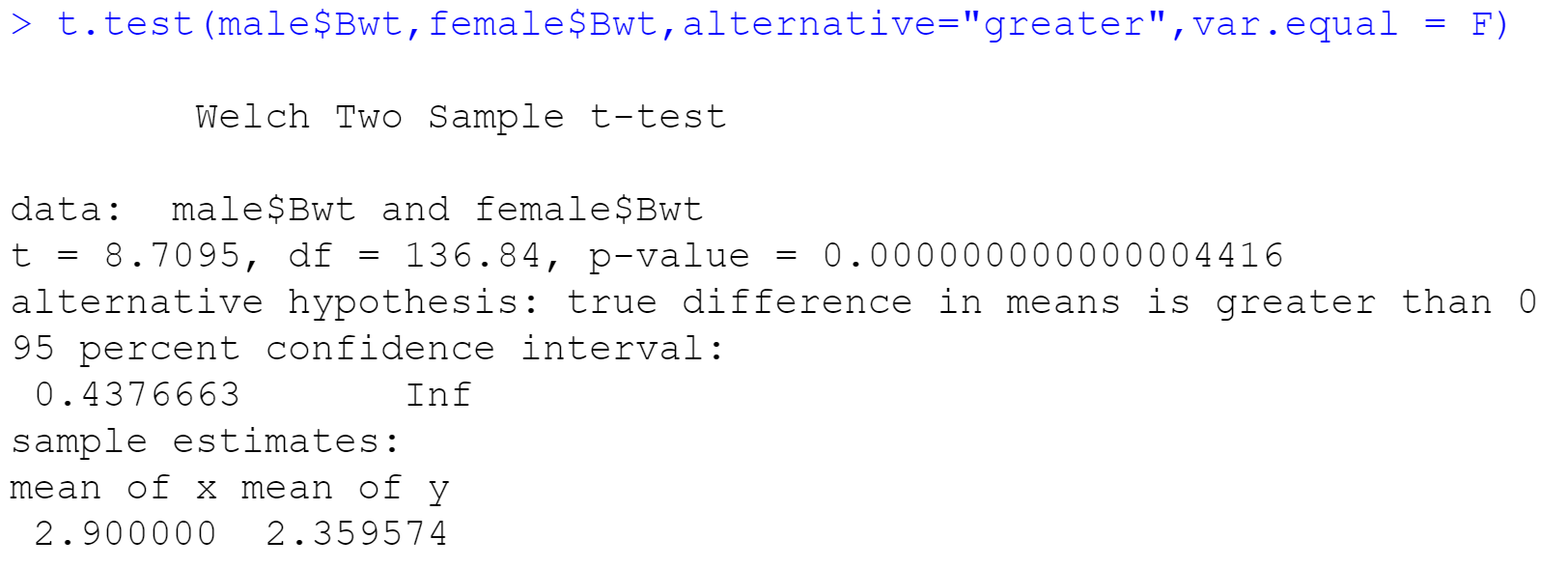


The test was two-sided because I wanted to see any difference, greater than or less than, between the means. The p-value is basically 0, which is less than our alpha of .05, so we can reject the null hypothesis in favor of the alternative. We reject that the mean body weights of males and females are the same and accept that they are different.

Next, I conducted a one-sided t-test to see if males are heavier than females. We are not concerned that males are actually lighter than females because all signs point to males being heavier. Even though the last test confirmed that their means are different, that test did not specifically say if males were heavier.

H0: µ1 ≤ µ2

Ha: µ1 > µ2



Since the p-value is less than .05, we reject the null that males are lighter or equal to females and accept that males are in fact heavier than females.

I conducted the same two unpaired two sample t-tests with the height variable. The p-values of both tests were essentially 0 (less than .05) so we can confirm that males are taller than females.

Summary

After the initial data points and boxplots seemed to point towards male cats being heaver and taller than female cats, the hypothesis tests checked to see if they were statistically significant. All four hypothesis tests allow us to conclude that male cats are heavier and taller than female cats.

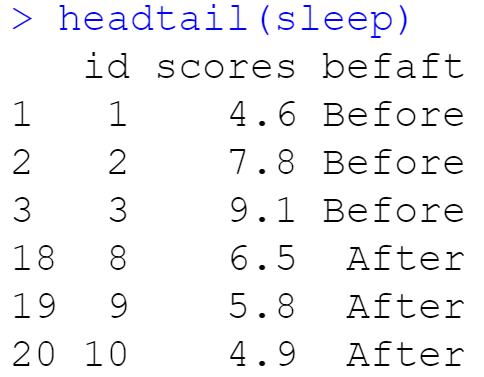
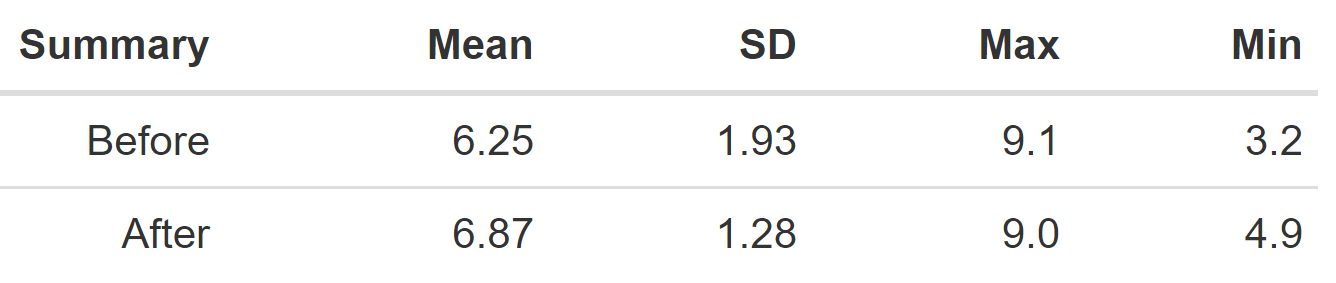
Part Two

Introduction

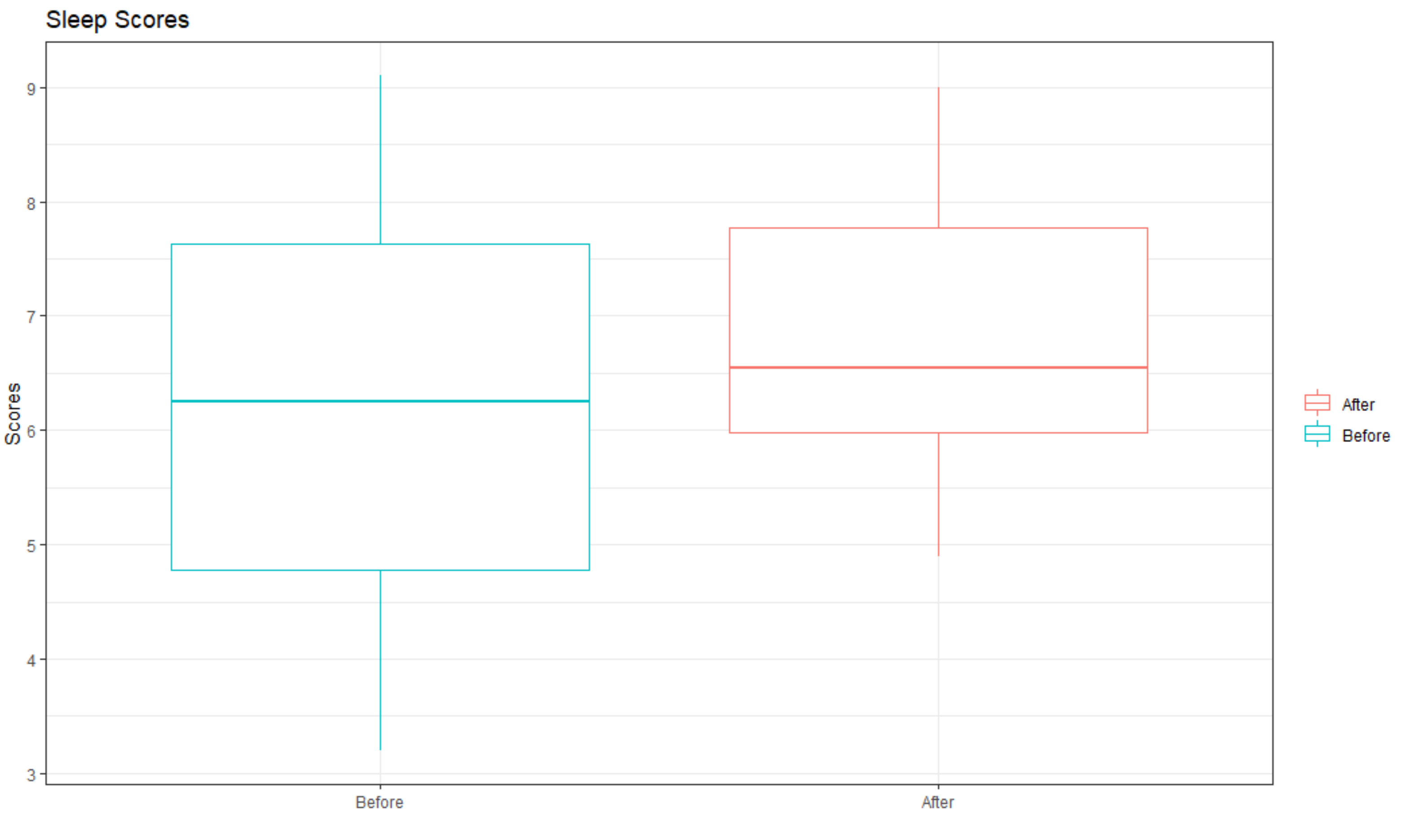
I began my sleep quality analysis by first inputting the raw data into a table in R such that I could analyze the affects of meditation on sleep. 10 students wore sleeping evaluators and had their scores recorded (0-10 scale, 10 = best) before the meditation workshop and after the meditation workshop. The goal was to determine if meditation, or at least this meditation workshop, had any effects on sleeping quality.

Analysis

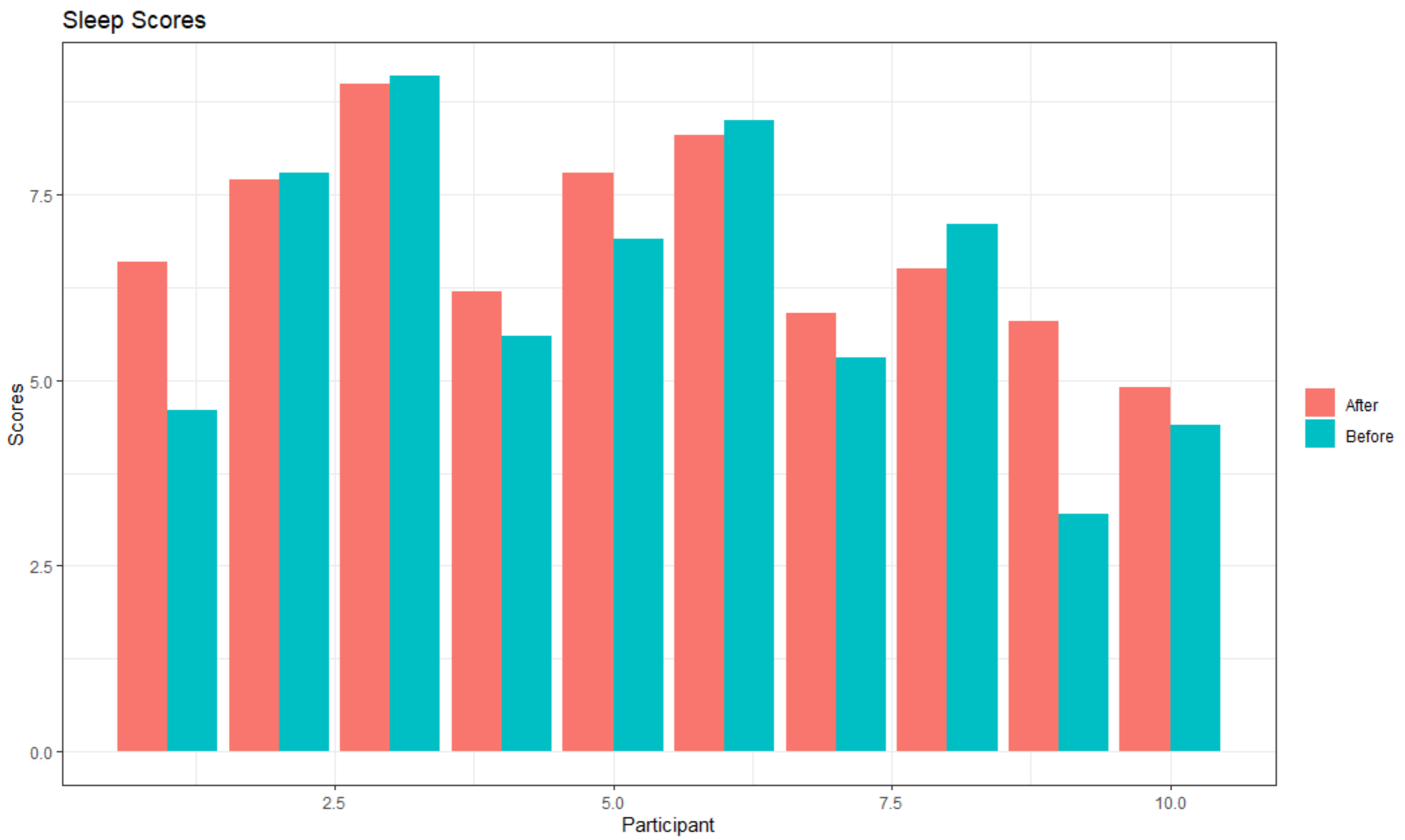
In addition to creating the before and after score vectors in R, I created the following data table to help with analysis. I also created the following summary table to quickly show some descriptive statistics.

As you can see, the average after score is higher than the average before score. The max before and after scores are about the same, but the minimum after score is a whopping 1.7 points higher than the minimum before score. I created the following boxplot to display these descriptive statistics. The most notable difference is the variance in after scores is dramatically smaller than before, even with an increased mean.



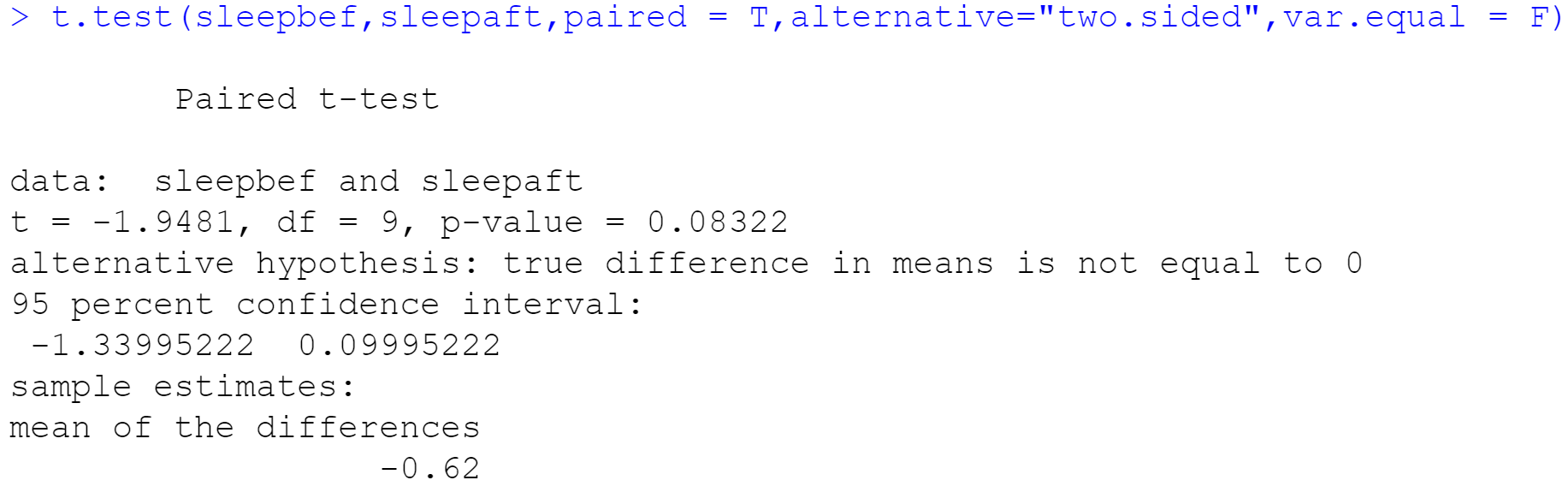
I created a double bar chart to show each student’s after and before scores. Even though only 6 students increased their scores (and 4 decreased), the increased scores were rather large compared to the tiny decreases.



Next, I conducted a two-sided paired sample t-test (at a 95% confidence level or α = .05) in order to see if the mean before and after scores are different. I chose a paired (dependent) test because each score is from the same student (one sample), just recorded before and after the workshop. The scores are also continuous, not discrete, so I ruled out any proportion tests.

H0: µ1 = µ2

Ha: µ1 ≠ µ2

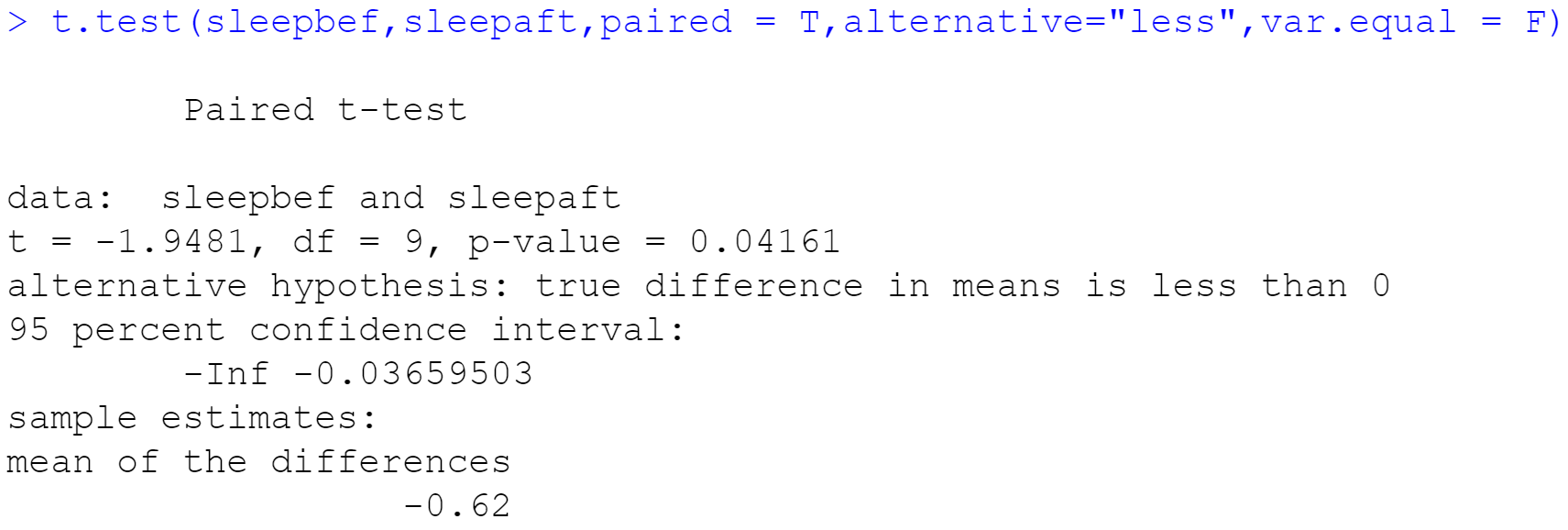


Since the p-value is greater than .05, we fail to reject the null hypothesis and cannot say that there is a significant difference in means. The p-value of .083 means there is an 8.3% chance that there is an actual difference in before and after scores, but this is too much uncertainty to confidently reject the null. If we decreased our confidence level (increase α from .05 to .1), we would reject the null hypothesis with the same results. By rejecting the null we would say there is a significant difference in before and after mean scores. However, it is important to establish our α ahead of time. The α is our cutoff point for our willingness to accept a type I error (false positive). A higher α should mean that false positive results would not have serious consequences. Medical testing, for instance, would have severe false positive consequences so they generally keep their alphas low. Increasing our alpha after seeing the results in order to change out conclusions is a misleading statistical practice. The consequences of a false positive (saying that the meditation camp was effective when it really wasn’t) could cause wasted time and investments in the program, but overall would not have severe repercussions.

Similar to part one, I was not concerned that after scores were lower than before scores. So, I conducted a one-sided paired sample t-test at a 95% confidence level to see if the before scores were less than the after scores.

H0: µ1 ≥ µ2

Ha: µ1 < µ2



This hypothesis test had half the p-value of the previous test because we only looked at if the after values were greater than the before values. This test did not look at if the after values were greater than or less than the before values. If we only ran this one-sided test we could say that after scores were greater than before scores and conclude that the meditation workshop was helpful. However, considering the uncertainty around the two-sided test, the different conclusions between the two tests with the same data, and the fact that we only studied 10 students, I cannot confidently make any assumptions about the effectiveness of the meditation workshop.

Summary

A quick glance at the initial data seemed to show that the meditation workshop had a small improvement on sleep scores. However, the hypothesis tests checked their statistical significances. I chose a 95% confidence level because I did not want to be too strict or too lenient in my conclusions. The negative impacts of false positives are not that severe so I would understand if someone else used the same dataset and conducted hypothesis testing at a 90% confidence level. However, I felt more comfortable keeping my alpha at .05. The two-sided t-test had a p-value of .083 so we could not say there was a difference in mean before and after scores. The one-sided t-test only had a p-value of .042 so it did say that after scores were greater than before scores. Taking all this into account, I cannot comfortably make any determinations about the effectiveness of the meditation workshop. I would like to conduct more testing with more sample groups and larger sample sizes before making any conclusions.

Citations

“Unpaired Two-Samples T-Test in R.” STHDA, www.sthda.com/english/wiki/unpaired-two-samples-t-test-in-r.

“Paired Samples T-Test in R.” STHDA, www.sthda.com/english/wiki/paired-samples-t-test-in-r.

“ggplot2 Scatter Plots : Quick Start Guide - R Software and Data Visualization.” STHDA, www.sthda.com/english/wiki/ggplot2-scatter-plots-quick-start-guide-r-software-and-data-visualization.

“How to Graph ‘before and after’ Measures Using Ggplot with Connecting Lines and Subsets?” Stack Overflow, 1 Apr. 1967, stackoverflow.com/questions/50910044/how-to-graph-before-and-after-measures-using-ggplot-with-connecting-lines-and.

Cmdline. “How To Reorder a Boxplot in R? Hint: Use Forcats.” Python and R Tips, 4 Mar. 2020, cmdlinetips.com/2019/02/how-to-reorder-a-boxplot-in-r/.

Saleh, Khalid. “One-Tailed vs. Two-Tailed Tests – Everything You Possibly Need to...” Invesp, 29 Apr. 2020, www.invespcro.com/blog/one-tailed-vs-two-tailed-a-b-testing-everything-you-possibly-need-to-know/.

Alboukadel, and Leonardo. “How To Easily Customize GGPlot Legend for Great Graphics.” Datanovia, 25 Dec. 2019, www.datanovia.com/en/blog/ggplot-legend-title-position-and-labels/.

“ggplot2 Barplots : Quick Start Guide - R Software and Data Visualization.” STHDA, www.sthda.com/english/wiki/ggplot2-barplots-quick-start-guide-r-software-and-data-visualization.