R Dataframes Assignment – text for exercise 5 part 3 – Sam Duffield and Stacey Stevens

In order to ensure that each line of the code operated as expected after we had fixed the file, we ran them one at a time and inspected the results displayed or changes to the dataframe. First, we changed the code for importing the original frame so that it could be easily adjusted to fit with various working directory locations as long as a data folder with the data in question was present. We then adjusted the first summarize function code to calculate the desired mean instead of the maximum value for the groups in the data set. We also added another pipe command to the second code to ensure that the data of interest would be grouped and summarized properly, even though there were other ways to fix this problem.

This ensured that the two codes had parallel structure and could be compared to each other. The only differences were the desired grouping variable, making it easier to compare the codes. If it would be impossible to calculate the values for the dataframe without code, we would have taken a subset of data with the identical structure and ensure that the code functioned properly. In order to compensate for any NAs in the dataframe. we could also add mean(volume, na.rm=TRUE) to the code to remove these so that the averages would only represent data that had been collected successfully.