**V506 Fall 24 R Lab 3 In-Lab Exercise**

Use Rmarkdown to document your response to each question.

1. Install DescTools library. Explore it. Read the help files for functions Freq and Mode.
2. Load “dplyr”,“ggplot2”, “here” and “DescTools” libraries using “pacman” into your environment.
3. Ggplot includes the dataset *diamonds* which has 53940 rows and 10 variables. **Load this data set and, using pipes, print out the summary statistics of this dataset.**
4. Use the Freq function to build a frequency table of the variable `cut`. Look at the output and interpret it.
5. Find the mode of variable cut. Interpret it using the output from the frequency table computed in the previous question.
6. Find the mean diamond price and the mean diamond price. Are they different?
7. Using ggplot(), create a histogram of the price distribution of the diamonds. Does the shape of the distribution make sense given the mean and median price? This chart should follow chart best practices—labels on the axes, a title, an appropriate binwidth, etc.
8. Use the geom\_vline layer of ggplot to add vertical lines to depict the mean and median price. Interpret the graph with these new features.
9. Save a new dataset called *diamonds2* where you have added a variable that is equal to 1 if the price of the diamond greater than or equal to the median price. This variable will be equal to 0 if the price of the diamond is below the median price. Call this new variable **high\_price.** (You should use ifelse(), mutate, and pipes to create this variable).
10. Using ggplot(), create a scatterplot with carat on the x-axis, price on the y-axis. Use the facet\_wrap() to create 5 separate plots for each “cut”. Again, include the axis titles and a main title.
11. Using ggplot(), create a bar chart where the x-axis is the “color” of the diamond. Again, include the axis titles and a main title.

There is nothing to turn in for this exercise. When you are finished, you are free to leave.