Project 1: Chapter 2-3
80 points

This project must be typed, and you may work in groups for this project (a maximum of three people). Answer the questions with complete sentences and use all the appropriate terms given in the notes or in the book. Include all the graphs and state what software, calculator, or any other website you are using to do your calculation and graphs. If you use the excel, list the functions you use to get your results.

The project will be due on Monday 01/30/21 by 11:59 pm. You may submit your project until Tuesday by 5 pm with a 5% penalty.

A popular U.S. automobile manufacturer has 10,000 dealerships located throughout the country. The automobile manufacturer has multiple brands within its portfolio: a value brand that caters to younger clientele, a moderate brand that caters to middle class customers and finally, a premium brand which is marketed to wealthy clientele. The company's leadership, located at corporate headquarters, is very interested in the relationship between the median salary of potential customers and the company's revenue. Specifically, the company is concerned that if potential customers' salaries continue to not increase in the future, the company's revenue will remain stagnant, which will in turn steer away potential investors and shareholders. The company's research department recently collected data for analysis in order to support leadership's upcoming discussion with shareholders and investors about the company's future revenue forecast. Sales figures from a random sample of 1000 dealerships were collected. The research division also conducted statistical analysis, using data provided by the Bureau of Labor and Statistics, to calculate the median salary of people living in the vicinity of these 1,000 dealerships. The Dealership Number, State, Median Salary, Annual Sales, Number of Vehicles Sold, Square Footage and Quality Award Winner data were collected for these 1000 dealerships.

For the raw data, use the excel file given on Canvas.

1. Create a histogram using a bin width of 5000, a stem and leaf plot, a boxplot, and a dotplot of the Median Salary data. From a visual analysis, describe the shape and center of the distribution. Which graph do you think best depicts the data? Explain.

2. Based upon your visual analysis of the Median Salary histogram, what is the most appropriate measure of center and variation to calculate and interpret?

3. Write an overall summary in the context of the Median Salary data including the shape, center and spread of the distribution. Use the mean, standard deviation, median, and Interquartile Range in the summary.

4. Corporate headquarters is very interested in analyzing dealership performance in order to identify best sales practices and potentially eliminate underperforming dealerships. Using the Number of Vehicles Sold data, create a boxplot and using the 1.5IQR Rule determine whether there are any dealerships that sold an unusually large or small number of vehicles last year.

    a. What is the shape of the distribution?

    b. Are there any outliers in this sample? If so, identify the dealership by identification number.

Make sure to include the questions and any other relevant files on your project.