ST 560

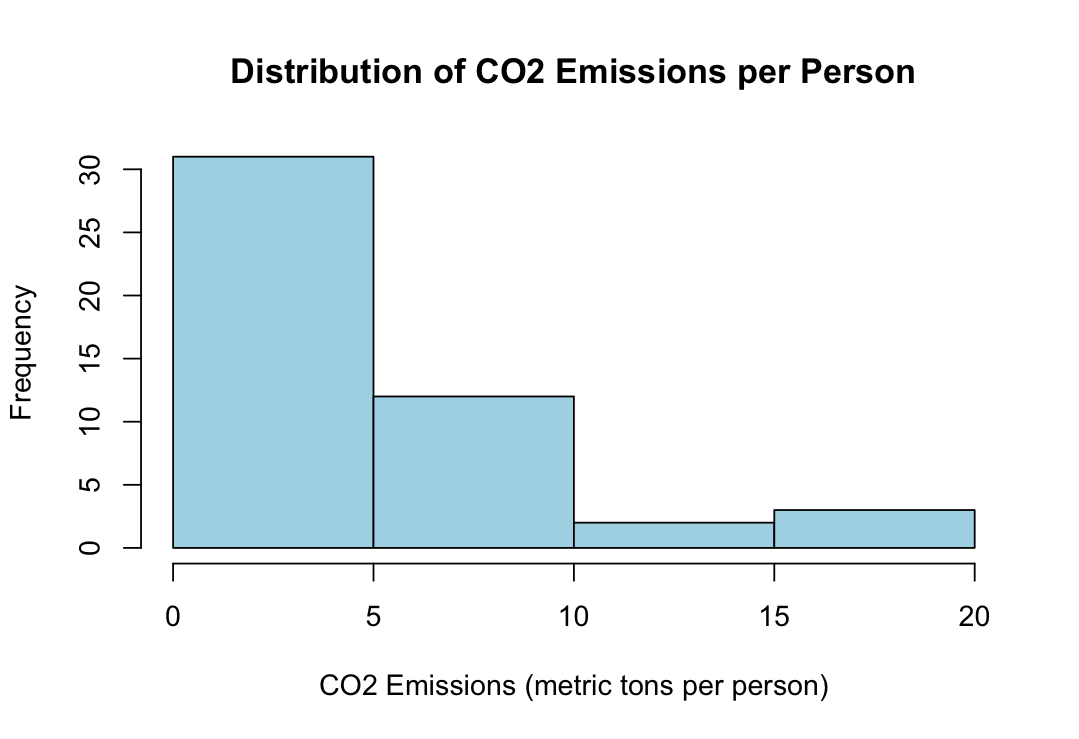
Gulmira Amangalieva

Homework 1

1. **CO2 emissions per person from 20 countries with population at least 20 million.**
2. Using the data set co2.txt, make an appropriate graph for the variable CO2.

A graph of carbon dioxide emissions

Description automatically generated



1. Describe the shape, center, and spread of the distribution.

The distribution of CO2 emissions per person seems right-skewed, with most countries having low to moderate emissions (median value Mdn = 3.2), and a few countries having much higher emissions. Maybe histogram?

The spread is quite large, with CO2 emissions ranging from 0.0 (for countries like Congo and Ethiopia) to 19.9 metric tons per person (for the United States). The relatively large standard deviation (ST = 4.8) suggests that there is considerable variability in CO2 emissions across the countries, and the variance (var = 23.2) confirms the high variability in CO2 emissions.

1. Countries such as Australia (C02 = 17.0), Canada (CO2 = 16.0), and the United States (CO2 = 19.9) can be considered outliers, having much larger CO2 emissions compared to other countries. Congo and Ethiopia (CO2 = 0.0 each) can be also considered outliers with the lowest values equal to zero.
2. **Educational data for 50 seventh grade students.**
3. Make an appropriate graph for the variable Gender.

A pie chart with a pink and blue circle

Description automatically generated

As can be seen from the pie chart, females constitute around one-third of the class population.

1. Make an appropriate graph for the variable GPA.

A graph of a number of bars

Description automatically generatedAs can be seen from the histogram, the distribution appears to be left-skewed, with fewer students having a higher GPA and the most frequent range between 7 and 8 points.

1. Compute the mean and median for the variable GPA. Comment on the difference between the mean and the median based on the shape of your graph in (b).

The mean value of GPA (M = 7.7) is lower than the median value (Mdn = 8), meaning that the majority of the data are closer to the median due to the left-skewness of the data.

1. Compute the standard deviation and interquartile range for GPA.

The standard deviation (SD = 2.19) indicates that there is moderate variability in the GPAs of the students. The interquartile Range (IQR = 2.18) indicates that most of the students’ GPAs fall within a fairly concentrated middle range.