Homework - Week 4

Your name here

1. Consider the builtin dataset `iris`.
    a. What is the structure of the `iris` data frame?
    b. Create a histogram of the `Sepal.Width` variable.
    c. Create a histogram of the `Petal.Width` variable.
    d. For both histograms, does the data appear normally distributed? Are they skewed?
    e. For both histograms, does it appear that the data come from more than one populations?
    f. What is the mean and median of `Sepal.Width`? What is the variance and standard deviation?
    g. What is the mean and median of `Petal.Width`? What is the variance and standard deviation?
2. Consider the builtin dataset `trees`.
    a. What is the structure of the `trees` data frame?
    b. Create a histogram of the `Height` variable.
    c. Create a histogram of the `Volume` variable.
    d. For both histograms, does the data appear normally distributed? Are they skewed?
    e. For both histograms, does it appear that there are outliers in the data?
    f. What is the mean and median of `Height`? What is the variance and standard deviation?
    g. What is the mean and median of `Volume`? What is the variance and standard deviation?
3. Load the dataset `bears.csv` from D2L.
    a. What is the structure of the `bears` data frame?
    b. Create a frequency table for the variable `MONTH`. What is the mode, if any?
    c. Create a histogram of the `WEIGHT` variable.
    d. Is there distribution of `WEIGHT` data normal? Is it skewed? Are there outliers?
    e. Based on your answers to part (d), do you expect the mean and median to be the same (or very close)? If not, which do you expect to be greater?
    f. What is the mean and median of `WEIGHT`?
    g. Based on the histogram in part (c), what would you expect the mode to be, approximately?
    h. What is the mode, if any?