STA 275 Practice Problems 1

1. The accompanying table gives a breakdown of the total U.S. oil consumption, in relative frequencies, categorized by the various purposes for which the oil is used.

Use	Relative Frequency
Gasoline	0.43
Industrial fuel oil	0.12
Heating oil	0.17
Jet fuel	0.07
Diesel fuel	0.05
Other	0.16

a. What is the categorical variable under study?
b. What are the categories of the categorical variable?
c. Construct a bar chart of the data. Include appropriate labels on the graph. Please write a sentence or two describing the bar chart.

2. The number of siblings in 20 families were determined and the result for each family is recorded below:

2, 2, 1, 3, 3, 4, 5, 0, 1, 1, 3, 4, 5, 0, 1, 2, 2, 2, 3, 4

a. What is the discrete variable under study?
b. Construct a frequency table for the discrete variable.
c. Construct a frequency histogram for the discrete variable. Include appropriate labels for the graph. Please write a sentence or two describing the histogram.

3. The accompanying 24 ozone readings (in parts per million) were taken in Los Angeles before the 1984 Olympics in order to study the ozone levels at the various sporting venues.

10, 14, 13, 18, 12, 22, 14, 19, 22, 13, 14, 16, 3, 6, 7, 19, 8, 13, 17, 28, 6, 23, 24, 35

a. Construct a stem-and-leaf display. Does this display adequately summarize that data? Please explain.
b. Construct a density histogram for the continuous variable. Include appropriate labels for the graph. Please write a sentence or two describing the histogram.

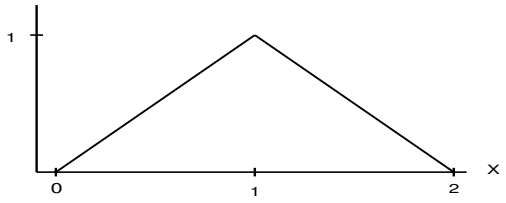
4. a. Invent a sample of size 5 for which four of the deviations, defined by $(x_i - \overline{x})$, are -4, 0, 1, and 2, respectively.
b. Use the definition of s to compute the standard deviation of your sample.

5. The accompanying 24 ozone readings (in parts per million) were taken in Los Angeles before the 1984 Olympics in order to study the ozone levels at the various sporting venues.

10, 14, 13, 18, 12, 22, 14, 19, 22, 13, 14, 16, 3, 6, 7, 19, 8, 13, 17, 28, 6, 23, 24, 35

a. Compute the sample mean ($\overline{x}$) and standard deviation (s). Please interpret the meaning of $\overline{x}$ and s in the context of the problem.
b. Determine the median.
c. Determine the upper quartile, lower quartile, and the interquartile range.
d. Construct a standard boxplot of the data. Are the data symmetric? Please explain.

6. The density curve of a continuous variable x is given below.



a. Find the mean of the continuous variable x.
b. Find the lower quartile, median, and upper quartile of the continuous variable x.

7. A major news corporation conducts an opinion poll by asking a random sample of 1500 U.S. citizens the following question: "Who do you plan to vote for in the next Presidential election?"

a. What is the response variable?
b. Is this an observational study or a designed experiment? Please explain.
c. What is the population?
d. What is the sample?
e. If you were the one in charge of collecting the data, explain how you would try to obtain a sample that is representative of the population.
f. Did your answer to **e.** involve random selection? Please explain.

8. Explain the relationship between the empirical rule and z-values (also known as z-scores).