MTH 1210
Lab 2 Boxplots, the Empirical Rule, and Z-Scores

Part A: Comparing Two Data Sets

The file **studying.mpx** contains the results of a survey that asked students in a large college how many minutes they study on a typical weeknight. The two columns in the spreadsheet correspond to **women's** and **men's** responses.

1. Open the *MINITAB worksheet* **studying.mpx**.
2. Create comparative dot plots of these data sets:

   GRAPH > DOT PLOT
   In the first box: Select the type of plot: Multiple Y's (Simple).
   < OK >.

   In the second box: Select the variables (double click).
   < OK >

   This will produce a dot plot in a graphics window. You can interactively reset the title and the labeling in the graphics window (double click on them).

3. **Copy and paste the dot plots into the Answer Sheet**.
4. Now construct side-by-side boxplots:

   GRAPH > BOX PLOT
   In the first dialogue box: Choose: Multiple Y's (Simple) for two or more boxplots side-by-side.
   < OK >
   In the second dialog box: Select the variables (double click).
   < OK >
   < OK >

   The resulting boxplots should appear side-by-side in the same graph.

5. **Copy and paste the side-by-side boxplots into the Answer Sheet.**
6. **Please answer Question 1 on the Answer Sheet.**
7. Now compute the sample mean for women and the sample mean for men ratings (STAT > BASIC STATISTICS > DISPLAY DESCRIPTIVE STATISTICS).
8. **Please answer Question 2 on the Answer Sheet**.
9. Compute the sample standard deviation for women and the sample standard deviation for men.
10. **Please answer Question 3 on the Answer Sheet.**

Part B: The Empirical Rule (68-95-99.7 Rule)

The file **octane.mpx** contains data on motor octane ratings for various blends of gasoline. A higher octane rating means the gasoline will produce less engine "knocking."

1. Open the MINITAB *worksheet* **octane.mpx**.
2. Compute the mean and standard deviation of the octane ratings (STAT > BASIC STATISTICS > DISPLAY DESCRIPTIVE STATISTICS).
3. **Please answer Questions 1 and 2 on the Answer Sheet.**
4. Construct a histogram of the data (GRAPH > HISTOGRAM).
5. **Copy and paste the histogram into the Answer Sheet**.
6. The Empirical Rule predicts that *approximately* 68% of the data will fall within one standard deviation of the mean, and *approximately* 95% will fall within two standard deviations of the mean.

   Using the values you got in **Question 2 of the Answer Sheet**, determine the *actual* percentage of data that lies in these ranges. You may find it useful to *sort* the data, and then count how many data values lie within one and two standard deviations of the mean:

   DATA > SORT
   In the dialog box: SORT COLUMNS: Select the variable
   BY COLUMN: Select the variable again
   < OK >

   This will produce the sorted data in a new worksheet.

7. **Please answer Question 3 on the Answer Sheet.**

Part C: Z-Scores

Recall that the *z-score* (or *standardized value*) for a variable x is

$$z = \frac{x - \mu}{\sigma} \qquad \text{or, when } \mu \text{ and } \sigma \text{ aren't known,} \qquad z = \frac{x - \bar{x}}{s}$$

1. First, you need the mean $\bar{x}$ and standard deviation s of the octane ratings from **Part B**. Use these values to compute the z-score for an octane rating $\boldsymbol{x} = \mathbf{88.5}$.
2. **Please answer Question 1 on the Answer Sheet.**
3. Minitab will compute z-scores automatically. At the top of a new column in the **octane.mpx** worksheet, write the heading Z-Score. Now compute the z-scores of all the octane ratings:

   CALC > STANDARDIZE

In the dialog box: INPUT COLUMN(S): Select the variable to be standardized
STORE RESULTS IN: Select the column in which to store the standardized values

< OK >

This will produce the z-scores in the new column.

4. **Please answer Question 2 on the Answer Sheet.**