Name: _____ **Date:** _____ **Class:** _____

Analyzing Backpack Weights

Part A:

1. Use the backpack weights data to calculate the following values for each of the variables ("Guess", "Actual", "Difference"):
   a. mean and standard deviation
   b. five-number summary

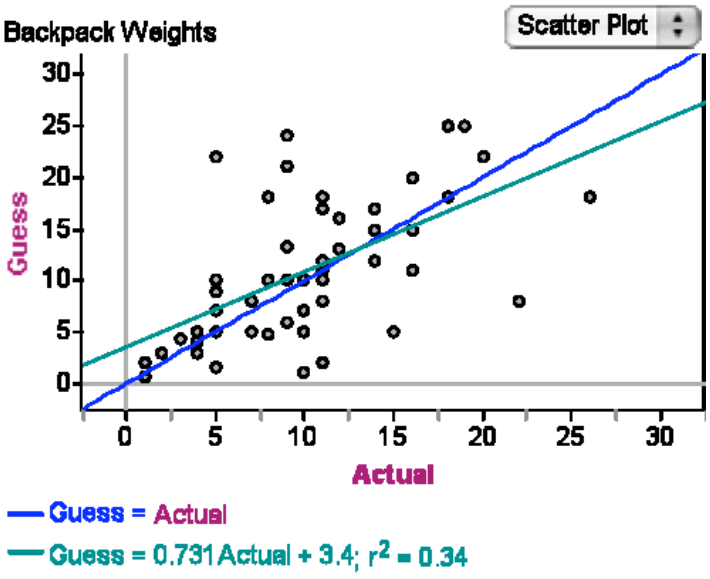
2. Use Fathom to generate a summary table for the values requested in Number 1.

3. Compare your results to the values provided by the summary table. If you find any discrepancies, determine the nature of the errors found.

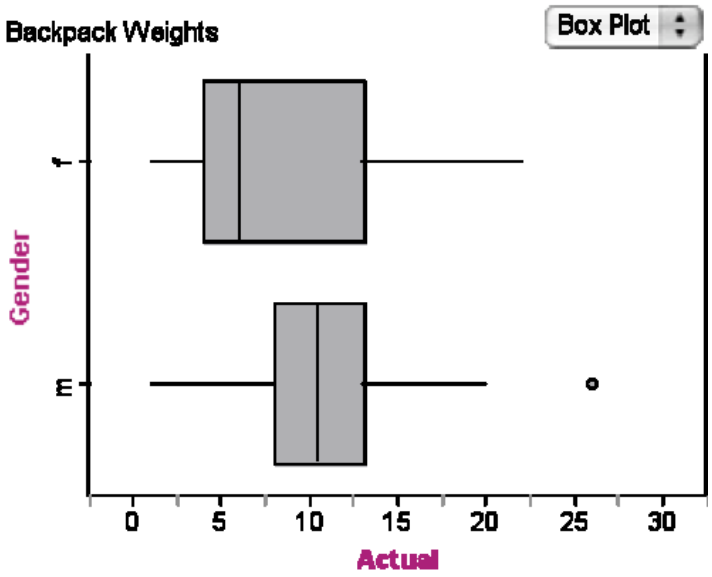
Part B:

1. Consider the following two graphs generated using Fathom.

GRAPH 1:



GRAPH 2:



a. Provide a list of questions that could be answered through the analysis of the graphs above.
b. Include complete and clear justifications to the questions posed.

2. Generate further graphs using Fathom that could best be used to answer the following questions. Be sure to provide full justifications to each question.
   a. Did this group tend to overestimate or underestimate backpack weights?
   b. Which group, males or females, tend to guess their backpack weights better?
   c. Are genders equally represented in this study?
   d. How do backpack weights vary in this study and between genders?

Part C:

1. Generate a histogram using "Difference" as the horizontal axis and "Relative Frequency of Difference" as the vertical axis.
   a. How well does the normal cumulative distribution model the data?
   b. Provide justifications to support your answer in Number 1a.
   c. Suggest different approaches to obtain a better fit of the data.

2. Generate a percentile plot using "Difference" as the horizontal axis and "Percentile" as the vertical axis.
   a. Create a vertical line that represents the mean of "Differences"
   b. Create vertical lines that represent (mean) $\pm$ ($k \times$ standard deviation), for $k = 1,2,3$.
   c. Use a vertical moving line (or other animation), to estimate the percentile of data that are:

i. Less than (mean) $\pm$ ($k \times$ standard deviation), for $k = 0,1,2,3$.
ii. In the interval [(-$k \times$ standard deviation), (-$k \times$ standard deviation)] for $k = 1,2,3$.

d. Plot the cumulative normal distribution and determine how well it models the data.

Part D:

1. In your own words, define the term “standard deviation”.

2. Suppose that the scale used to weigh the backpacks was not calibrated correctly and the true “Actual” weights of the backpacks were misrepresented by *p* lbs. What effects would *p* have on the tables, graphs, distributions, models and calculated statistical values?