

MGT 9050
Project #1

MODELS, ERROR, AND SIMPLE DESCRIPTIVE STATISTICS

1. I give a 10-item test of Clemson University trivia to 13 people. Their scores are as follows:

12 16 18 14 13 15 12 13 11 17 10 13 14

Before we turn over our calculations to the computer, answer these questions using hand, calculator, or Excel calculations.

- a. Calculate the mean, median, and the mode.
 - b. Using the mean as your parameter estimate for your model, calculate the number of errors, the sum of the absolute errors, and the sum of the squared errors.
 - c. Now use the median as your parameter estimate and calculate the same three error terms.
 - d. Use the mode as your parameter estimate and calculate the same three error terms.
 - e. Look across your answers to (b), (c), and (d) to see for each error term, which estimate gives the lowest error? That is, which of the mean, median, and mode minimizes the sum of squared errors and by how much? Which one minimizes the sum of absolute errors and by how much? And which one minimizes the count of errors and by how much?
2. Use RStudio or Jamovi to obtain the mean and median for the data in the previous question and share a screenshot of the results.
 3. "HR Data.csv" contains 311 rows and 36 columns. First, open this data file in RStudio or Jamovi. Next, obtain estimates of the mean, median, variance, and standard deviation of two variables (SALARY, ENGAGEMENTSURVEY). Report these along with a one sentence interpretation of what the values mean to you.
 4. Compare the mean SALARY for Men and Women. Briefly describe these results, being sure to indicate whether it *APPEARS* (we don't know how to do the formal test yet) that it would be useful to make predictions of SALARY conditional on employee sex (M or F).