

Lab Exercise #2 Descriptive Statistics and Distributions

Exercise Objectives

- 1) Learn how to describe data (variables) in statistical language.
- 2) Understand normal distributions and how to test for normality.

Using the SPSS data set from the Anchorage Parks and Open Space survey, answer the following questions:

1. Examine the frequency distributions for the variables in Q-21 (Q21_1, Q21_2, Q21_3, Q21_4, Q21_5, Q21_6, Q21_7, and Q21_8). Note: these variables should have values within the range of numbers from 0 to 7.4. If a number is out of range, there has been a data entry error. Are there any obvious data entry errors? On which variables?

2. After correcting the data errors (you can simply delete the values), generate the following statistics for each variable:

mean, median, mode, standard deviation, variance, range, standard error of the mean, skewness, and kurtosis.

- a. Is Q-21_1 positively or negatively skewed? What about Q-21_6?
- b. Is Q-21_2 more pointy or flat distribution (kurtosis)? What about Q-21_7?
- c. Calculate the Z-scores for skewness and kurtosis for Q-21_4 (see page 72 in Field text). Does skewness appear to be a problem for this variable?
- d. Which variable has the highest level of variability?
- e. Which variable has the least level of variability?
- f. Which variable has the highest median value?
- g. What is the modal value for all the variables?
- h. Which variable has the highest mean value?
- i. Which variable has the lowest standard error of the mean?

3. Generate histograms with a normal curve superimposed. Of the eight variables, which variables appear to be normally distributed?

4. Following the procedure in the textbook (pp. 93-96), conduct a statistical test to determine the normality of the variable distributions.

a. Which variables, if any, are normally distributed?

5. If variables are not normally distributed, one of the parametric assumptions is not satisfied. To determine whether responses from two different groups are statistically different (they will be always numerically different) when the dependent variable is not normally distributed, one should run a non-parametric statistical test. Follow the procedure outlined on pp. 528-531 to run the non-parametric statistical test (Mann-Whitney) to determine whether the responses of women differ from men (Q-27) on the eight variables (Q-21_1 through Q-21_2).

a. Which variables do the responses of men differ from women (assuming a statistical significance level of .05)?

