

**LAB 3: Measures of Central Tendency and Dispersion; Graphical Displays**

Name: \_\_\_\_\_

Date: \_\_\_\_\_

1. For which of the following variables would it be appropriate to obtain a histogram?
  - a. Miles driven to work each week
  - b. Attitude toward affirmative action
  - c. Blood pressure
  - d. Annual household income
  - e. Happiness of marriage
  - f. Region of country
  - g. Subjective social class identification
  - h. Age in years
  
2. For each of the following variables, indicate which measure(s) of central tendency would be permissible (place an X in the appropriate blank).

	Mode	Median	Mean
a. Political party identification	_____	_____	_____
b. Level of occupational prestige	_____	_____	_____
c. Annual number of deaths from AIDS	_____	_____	_____
d. Religious identification	_____	_____	_____
e. Violent crime rate	_____	_____	_____
f. Age in years	_____	_____	_____
g. Level of support for US foreign aid	_____	_____	_____
h. Number of consecutive alcoholic drinks	_____	_____	_____

*For questions 3-5, use the Explore procedure to produce descriptive statistics and plots. You are required to print out and hand in the output from your work.*

3. Find the variable AGE in the dataset named CH10END.
  - a. Examine the stem-and-leaf plot. How would you describe the distribution of this variable? Is it approximately normal, or is it skewed to the left or right?
  
  - b. Report the mean and the median for AGE. Which of these measures would you use to describe central tendency? Why did you make this choice?
  
  - c. Report the range, interquartile range, and standard deviation for AGE. Based on the standard deviation, roughly two-thirds of all cases fall between which ages?

4. Find the variable CHILDS in the dataset named FAMILY96.
  - a. Examine the boxplot. How would you describe the distribution of this variable? Is it approximately normal, or is it skewed to the left or right? Are there any outliers?
  - b. Report the mean and median for CHILDS. Which of these measures would you use to describe central tendency? Why did you make this choice?
  - c. Report the range, interquartile range, and standard deviation for CHILDS. Which of these measures best summarizes the variability in number of children?
  
5. Find the variable TVHOURS in the dataset named CH10END.
  - a. Examine the boxplot. How would you describe the distribution of this variable? Is it approximately normal, or is it skewed to the left or right? Are there any extreme outliers?
  - b. Report the mean and median for TVHOURS. Which of these measures would you use to describe central tendency? Why did you make this choice?
  - c. Report the range, interquartile range, and standard deviation for TVHOURS. Which of these measures best summarizes the variability in number of hours watching television?
  
6. You can use data from previous years of the GSS to answer questions about change over time. First, find the variable EDUC in the dataset CH10END. Calculate the mean and standard deviation, and report both statistics below. Then, find the same variable in the dataset from six years earlier, GSS90a. Calculate and report the mean and standard deviation again. Is there any difference between 1990 and 1996? If we used GSS data from different years in this way, which type of longitudinal design would it be?