Sociology 3155 – Quantitative Research Methods & Analysis  
Homework #2: Due Thursday, September 29, at the beginning of class. PLEASE: (1) Type all answers and (2) attach all calculations and SPSS output. **Worth 31 points**.

**FOR PROBABILITIES AND AREAS UNDER THE NORMAL CURVE, ROUND TO 4 DECIMALS; FOR PERCENTAGES, ROUND TO 2 DECIMALS.**

**PART A, *Complete the following problems (Based loosely on the Healey book). Note: the answers for the odd number problems for Healy are in the back of the book if you want to practice these types of problems.***

1. (2 pts): After taking the state merit examinations for the positions of sanitary engineer and fart catcher, you receive the following information on the tests and on your performance. On which of the tests did you do better?

Sanitary Engineer: mean = 118, s = 17, your score = 127  
Fart Catcher: mean = 27, s = 3, your score = 29

2. (2 pts): To be accepted into the royal order of snobby students, a person must have a GPA in the top 10% of their school class. If the mean GPA is 2.78 and the standard deviation is .33, which of the following GPAs would qualify?

3.0, 3.20, 3.21, 3.25, 3.40

**PART B**

**1. (3 pts) The following data represent the percentage of workers living in each city who used public transportation to commute to work in 2000.**

|  |  |  |
| --- | --- | --- |
| **City** | **%** | **a. Calculate the mean and median, and identify the mode of this distribution.  b. Compare the mean and median. Which is the higher value? Why?  c. Exclude New York from this distribution and recalculate the mean and median. Which one decreases more? Why?** |
| Boston | 32 |
| Chicago | 26 |
| Detroit | 9 |
| Houston | 6 |
| Los Angeles | 10 |
| Minneapolis | 15 |
| New Orleans | 14 |
| New York | 53 |
| Philadelphia | 25 |
| Phoenix | 3 |
| San Antonio | 4 |
| San Diego | 4 |
| San Francisco | 31 |
| Seattle | 18 |
| Washington, DC | 33 |

**2. (3 pts) Consider the following summary statistics for school districts' annual per student expenditures (in dollars) for education from 12 cities for two different years, 1997 and 2007:**

1997: mean = $7,050; median=$7,215 standard deviation (s) = 804

2007: mean = $9,009; median=$7,516; s=1,960

***Using the above information, describe the changes in these distributions in terms of their typical scores, spread of scores and shape (e.g., normal, skew?). Use full sentences and be as descriptive as possible.***

|  |  |  |
| --- | --- | --- |
| **3. (6 pts) Two hundred applicants to a police department have taken its aptitude test, and the average score was 35 with a standard deviation of 4. Assuming the distribution is normally distributed, what percentage of applicants have a score:** | | |
| a. between 31 and 37? b. between 41 and 46? c. more than 40? d. more than 33? e. less than 45? f. less than 29? |  |  |
| **4. (6 pts) For the past 10 years, your city fire station has averaged 47 calls for a “cat in a tree” per year, with a standard deviation of 5. Assuming the factors that cause cats to run up trees are constant, and that the distribution in normally distributed, what is the probability that next year the number of treed cat calls (snicker…) will be:** | | |
| a. less than 60? b. less than 47? c. between 42 and 52? d. more than 30? e. more than 55? f. between 45 and 50? |  |  |

**PART C, SPSS**

**Use SPSS and the 2010 GSS data set to create a frequency table for the variable HoursTV (hours per day watching TV).**

***In the Frequencies box, after you have selected this variable, click on “Statistics…” Then, check the boxes next to: Mean, Median, Mode, Quartiles, Cut points for 10 equal groups and Std. deviation. Then click on “Continue,” and click “OK” to run the frequency.***Use the output from this analysis of HoursTV to answer the following questions (please use full sentences) (5 pts)

1. What is the inter-quartile range (Q)?

2. What is the standard deviation?

3. What is the 70th percentile value for hours spent watching TV?

4. What is the direction of the skew for this" variable? Would you consider this skew to be problematic? Why or why not?

**In the GSS Data, examine the variables “colscinm” and “sibs.” Which measure of central tendency would you use for each of these variables? Report the value of this measure. Why would you use this measure?**

5. colscinm (2 pt)

6. sibs (2 pt)

**REMINDER: STAPLE SPSS OUTPUT TO YOUR ASSIGNMENT**