

**Homework #2 (Descriptive Statistics)**  
**PSCI 2300 - Fall 2019 (Dr. Hensel)**

This homework assignment is due at the beginning of class (i.e., before lecture starts) on **Monday, October 14**. Any work turned in after lecture begins that day (even if only a few minutes into class) will be assessed a late penalty. Also, note that all work must be your own -- students found to have copied their answers from other students (or to have had their answers copied by other students) will receive automatic zero grades on this assignment, and may face further disciplinary action.

Be sure to show your work wherever relevant; correct answers that do not show their work will only receive half credit.

*For the first four questions, you are to analyze the following data set, which shows the percentage of votes cast for Donald Trump in the 2016 presidential election in each of the four Census Bureau regions:*

Northeast: 40  
Midwest: 53  
South: 55  
West: 46

1. What is the mean level of support for the president across these four regions? Be sure to show your work. Briefly interpret what this value tells you (i.e., explain the interpretation of this value in 1-2 sentences).
2. What is the standard deviation of support for the president across these four regions? Be sure to show your work. *(Remember, you will need to calculate each score's deviation from the mean, square the deviations, sum the squared deviations, divide by N to find the variance, and take the square root to find the standard deviation. Feel free to use a calculator, Excel, or some other tool for each of these calculations, but you still need to do each of these five steps -- you will only receive half credit if you just find the correct answer by using some other tool to calculate the entire standard deviation within showing your work on each step.)*

*The remaining questions involve basic calculations using data on the 50 U.S. states compiled by the author of our textbook, which is available from the same portion of the class web site where you access the homework assignments. You will analyze this data using SPSS or PSPP statistical software, as described in the class SPSS Guide that is provided on the class web site (and more briefly at the end of this assignment sheet).*

Nominal-level Variables

3. Using SPSS or PSPP, produce a frequency table for the variable **region** (which may be listed in some places as **Census Region**; this is the region of the country where the state is located).

Be sure to copy and paste the output into your homework, or else print out the output in the computing lab and attach it to your homework when you hand it in.

4. Using SPSS or PSPP, produce a bar graph for the variable **region**.

Be sure to copy and paste the output into your homework, or else print out the output in the computing lab and attach it to your homework when you hand it in.

Ordinal-level Variables

5. Using SPSS, produce a frequency table for the variable **Gun\_rank3** (which may be listed in some places as **RECODE of Gun-rank11**; this is the state's categorization of gun restrictions, from more restrictive to less restrictive).

Be sure to copy and paste the output into your homework, or else print out the output in the computing lab and attach it to your homework when you hand it in.

6. Using SPSS, produce a bar graph for the variable **Gun\_rank3**.

Be sure to copy and paste the output into your homework, or else print out the output in the computing lab and attach it to your homework when you hand it in.

7. What are the mode and median categories for **Gun\_rank3**? (*1/2 point each*) How did you determine each of these?

#### Interval/Ratio-level Variables

8. Using SPSS or PSPP, calculate the measures of central tendency and dispersion for the variable **hs\_or\_more\_2015** (which may be listed in some places as **Percent population high school education or higher**; this is the % of state population with at least a high school education). This should include the following measures: Mean, Median, Mode, Standard deviation, Minimum, Maximum, and Range.

Be sure to copy and paste the output into your homework, or else print out the output in the computing lab and attach it to your homework when you hand it in.

9. Using SPSS or PSPP, produce a histogram for the variable **hs\_or\_more\_2015**.

Be sure to copy and paste the output into your homework, or else print out the output in the computing lab and attach it to your homework when you hand it in.

10. Based on the median and mean that SPSS/PSPP produced, as well as the results of the histogram, what can you say about the skewness of this variable? How did you determine this? (Be sure to refer to how both the median/mean and the histogram led you to this conclusion)

### **SPSS Instructions for This Assignment**

• For much more detail about using SPSS, see my SPSS Guidelines and Instructions document that is posted on the class web site.:

<http://www.paulhensel.org/Teaching/spss.pdf>

• Download the data set that you will need for this assignment (perhaps to a flash drive if you are not working on your own personal computer). From a computer that has SPSS, open this data file in SPSS, using one of the following two options:

--Double-click on the data set, which should automatically open SPSS and then open the data set in SPSS. [*Note that this option does not work in PSPP*]

--Open SPSS or PSPP manually by clicking on the icon or on an alias/shortcut to it. Once it is open, select and open the data set using **File > Open** from the menu bar at the top of the screen. [*Note that this is the only option that works in PSPP*]

• Once the data set opens in SPSS, you should see a spreadsheet with a list of variable names across the top. It is advisable to produce a list of variables, so you know what is in the data set. This can be done using the following command:

--DISPLAY DICTIONARY: **File > Display Data File Information > Working File**

#### *Calculating Frequencies*

• To calculate the frequency distribution for any variable in the data set, go to **Analyze > Descriptive Statistics > Frequencies**. This will open a dialog box that allows you to choose one or more variables for analysis.

• Select the variable you want to examine on the left side of this box, push the arrow button in the middle of the screen, and the variable should then appear in the window at the right side of the box; once this happens, click OK. (If you want to remove a variable from the list, select the variable in the right side of the box, and click the arrow in the middle of the screen.)

• This will give you a frequency table with a variety of frequency-related details for each value of the variable:

the frequency count (the number of cases that have this value), the percentage of cases, the valid percentage of cases, and the cumulative percentage of cases.

- Once the output appears in the Output Viewer window, be sure to print it out (or copy and paste it into a word processing document), because you will need to turn this in to get full credit for your assignment.

#### *Producing Bar Graphs and Histograms*

- To generate bar graphs/charts, pie graphs/charts, or histograms for any variable in the data set, go to **Analyze > Descriptive Statistics > Frequencies**. This will open a dialog box that allows you to choose one or more variables for analysis.
- Select the variable you want to examine on the left side of this box, push the arrow button in the middle of the screen, and the variable should then appear in the window at the right side of the box; once this happens, click the Charts button and select the correct type of graph/chart and any desired options. (Be sure that the "Chart Values" option is set to Frequencies, not Percentages.) Click the Continue button to go back to the main Frequencies window, and click OK.
- Once the graph displays in the Output Viewer window, be sure to print it out (or copy and paste it into a word processing document), because you will need to turn this in to get full credit for your assignment.

#### *Calculating Central Tendency and Dispersion*

- To calculate measures of central tendency and measures of dispersion or variability for any variable in the data set, go to **Analyze > Descriptive Statistics > Frequencies**. This will open a dialog box that allows you to choose one or more variables for analysis.
- Select the variable you want to examine on the left side of this box, push the arrow button in the middle of the screen, and the variable should then appear in the window at the right side of the box; once this happens, click the Statistics button and select the desired statistics that you want to be calculated. For the purposes of this assignment, you should select Quartiles, Mean, Median, Mode, Std. deviation, and Range (all of which we have used in class). Click the Continue button to go back to the main Frequencies window, and click OK.
- Once the output appears in the Output Viewer window, be sure to print it out (or copy and paste it into a word processing document), because you will need to turn this in to get full credit for your assignment.