

Telling A Story Using Descriptive Statistics

This assignment helps you to begin thinking about how you can tell a story using descriptive statistics such as means, medians, standard deviations, percentiles, and frequencies. You'll use statistical software to summarize and display information. The work you do here may be useful for the policy report; you should think ahead to that assignment.

Assignment Requirements

- *Read the data into SPSS. [This is Steps 1-3 below.]*
- *Pick 2 or 3 variables on a related theme and find the appropriate descriptive statistics for them (e.g., sample means, standard deviations, medians, frequencies, histograms or bar charts. [This is Steps 4-5 below.]*
- *Write a 2-3 page memo (single-spaced) describing your findings. Incorporate your graphics or tables into the memo as you tell your story. [This is Steps 6-8 below]. [Be sure to read "Do's and Don'ts for Writing about Statistics".]*

Do not hand in your print out! Just hand in your memo.

About SPSS

SPSS is available in the Evans School computer lab and at the Center for Social Science Computation and Research (CSSCR, in the basement of Savery Hall). Each location has lab consultants, so don't be afraid to ask questions. If you prefer to work on your own computer, you can buy a copy of SPSS for your home computer at the University Bookstore Computer outlet (corner of University Way and 43rd). Later versions of SPSS (v. 12 and 13) allow you to use long variable names (more than 9 characters), but these are not compatible with earlier versions of SPSS so you might not want to use them. You should use the on-line SPSS resources, help menu, and "results coach" to help you with SPSS. You can also look at:

<http://julius.csscr.washington.edu/pdf/spss10.pdf> .

About the Data

The data for this assignment come from the 2004 Washington State Population Survey. It is a random sample of the WA population similar to the national Current Population Survey. The data set has 17,788 (people) observations from 7,095 households. This includes an oversample of households from King County (region 3) in order to obtain larger samples of people of color.

The data set in SPSS format with variable and value labels is available on the Evans School server in the folder: \\Evanshost\CLASSES\PBAF 527- Stats\Data Sets\

Information on the data set is available at <http://www.ofm.wa.gov/sps/2004/default.asp> and on the course web page. The "Data Dictionary" describes each of the variables. You'll work with these data for the policy memo and later assignments, so this is a good time to invest in exploring the data.

Get Acquainted with SPSS and the Data Set

1. Make a plan

Before you begin at the computer you should spend some time thinking about the data and your research question. Look at the list of variables in the data dictionary and decide which variables interest you and how you want to explore them (which descriptive statistics). This will save you time at the computer and keep you focused.

2. Start your software

In Windows, you can start SPSS by double clicking on the SPSS symbol. Alternatively, you can click on the icon for an SPSS data file to open the file within SPSS.

The screen will change and you'll see a screen with the data editor (it looks like a spreadsheet). Later, you can also see screens for the "SPSS for Windows Viewer" (the results you'll generate), and syntax (SPSS commands). You can change between these windows by clicking on the Window menu and selecting it from the list of open windows. There are a number of ways to give SPSS your commands, but I'll give you the instructions for only one method (by using pull-down menus). Later learning to use syntax (command) files will save you time.

3. Read your data into SPSS

You need to tell SPSS where your data are and what kind of file it is. To do this, click on the FILE menu, choose OPEN choose DATA.

The file is called SPS04.sav and is in: \\Evanshost\CLASSES\PBAF 527- Stats\Data Sets

In the Evans lab look for pbaf527-stats (look for an icon on the desktop too) or click on "Shortcut to Classes" folder, then "PBAF 527-stats", then "Data sets".

[Note: SPSS knows the data is SPSS data because of the ".sav" extension on the file. You could also read in data from another program by changing the FILE TYPE.]

Later, if you create new variables save them in a new SPSS data set on your disk and name the file something new, such as a: wapop2.sav so you have the variables next time.

You should see the data in a spreadsheet in the DATA EDITOR window. Each row in the DATA VIEW spreadsheet is a different person (case). Each column shows the values of a different variable; the labels are in the top row. Click on the VARIABLE VIEW tab to see more details on each variable (labels, missing values, etc.).

4. Get some statistics

When you are using a new data set, it's good to get in the habit of looking at the means, standard deviations, and ranges for each variable. That way you can see if all the variables and observations are there or if there are any unexpected data values or distributions.

To get **descriptive statistics**, click on the ANALYZE menu,
choose DESCRIPTIVE STATISTICS,
then DESCRIPTIVES

The window will show you a list of all the variables on the left. Select the variables you'd like descriptives for by clicking on them. If you want more than one, hold the <ctrl> key down as you click on additional variables. You can use your cursor arrow keys or mouse to move down the list. By default, the list of variables shows the variable labels, rather than their names. If you run your mouse pointer over the list, you'll see the variable name at the end that will match your data dictionary.

[If you want to only see the variable names, pull down the EDIT menu, choose OPTIONS, and under the "variable lists" click the button in front of "display names." You could also alphabetize the list if you want by clicking on the circle in front of "alphabetize." You'll have to close SPSS and reopen it for the option to take effect.]

After you have selected your variables, click on the arrow button to move them over to the right hand list, for analysis.

Which statistics do you want to see? Click on the OPTIONS button to select the statistics you want from among mean, standard deviation, etc. Choose CONTINUE to get back to the previous window.

Ready to run? Choose OK to tell SPSS to compute the statistics.

How does it look? SPSS will compute the statistics and show you the results in the Output Navigator. You can use cursor arrows and page up and down to look at the results or you can click on the titles in the left side of the navigator.

To get **percentiles or medians**, pull down the ANALYZE menu
choose DESCRIPTIVE STATISTICS
then choose FREQUENCIES

Select your variable from the list, and click on the arrow to move it over to the list. You can get percentiles or a median for your variable by clicking on ANALYZE

within the FREQUENCIES window and selecting the appropriate boxes. CONTINUE, and then click on OK when you're ready to run. [What kinds of variables do you want frequencies for?]

5. Getting Graphics

To see a **histogram or a bar chart** to see the shape of your distribution

Pull down the ANALYZE menu,
choose DESCRIPTIVE STATISTICS,
then choose FREQUENCIES

select your variable from the list and click on the arrow to move it over to the list

choose CHARTS
then select HISTOGRAM or BAR CHART

Bar Charts are appropriate for discrete variables. A histogram will group values into categories, whereas a bar chart will show each value separately (disastrous is hundreds of different variable values!). You probably want to click on "percentages" instead of "frequencies" and to uncheck the box to DISPLAY FREQUENCY TABLES unless your variable has only a few values. Click OK to get your results.

To see your histogram or bar chart, go to the SPSS for Windows Output Viewer. You can edit the labels and axes of the chart by double clicking on the chart. Also, you can cut and past a chart to a word processing package by clicking EDIT and COPY CHART, go to your word processing package and PASTE. PC users can highlight the chart in the SPSS for Windows Viewer and choose COPY, switch to the word processing package, and EDIT, then PASTE. [You probably want to past special as PICTURE to avoid huge files.] When you're done looking at your charts in SPSS, minimize the window to get back to the regular SPSS screen.

To create a **pie chart** for a discrete variable choose:

GRAPHS
PIE
SUMMARIES OF GROUPS OF CASES (Hit DEFINE)

Choose what SLICES REPRESENT as % OF CASES (for relative frequencies). Pick your variable from the list and move (with arrow) to DEFINE SLICE BY. [You may want to click on OPTIONS, then unclick on "DISPLAY GROUPS WITH MISSING VALUES" so that the missing values don't show.] Now click on OK and check the results.

6. Moving Output from SPSS

Copy the results from the output window to your word processing package by using the mouse to select the results you want to copy, then clicking on EDIT menu and choosing COPY. Switch to the word processing software, and paste your results into a new document [put your cursor where you want the results to appear, pull down the EDIT menu, and then PASTE]. You might also want to

paste summaries (e.g., relative frequencies) into Excel and create better charts there.

Print results directly from SPSS by selecting FILE menu from within the SPSS for Windows Viewer, then PRINT, and choosing the print all of the output file or just a selection.

7. Saving your work

After you're exhausted from your data extravaganza, you may want to save some souvenirs. Also, save your work as you go in case disaster strikes.

To **save your output file**, click on the output window to bring it to the top, click on the FILE menu, choose SAVE AS, and type in a new filename for your output (e.g.: ASSIGN1.SPO). You can later cut and paste from this file by re-opening the file from within SPSS. Also, it is always a good idea to save this in case you forget later exactly what you did!

To **save your data** on your disk, after you create new variables, you can create a new version of the data set by pulling down the FILE menu from the data editor, select SAVE AS, and type in the name of your new data set with the appropriate drive (e.g. a:SPS2004_2.sav). Next time read in this data set instead of the version on the network or hard drive.

8. Write the Story

Write a 2-3 page memo (single-spaced) describing your findings. Incorporate your graphics or tables into the memo as you tell your story. [Be sure to read the "Do's and Don'ts for Writing about Statistics".]

When you're done with your initial computing, study the results and start writing. Writing about statistics is an iterative process, because as you write you'll find you need a different graph or statistic to make your argument. The goal of your memo is to tell a story supported by data to a policy maker who is interested in the issue you've just looked at. You need to interpret your findings for a non-technical policy maker—not explain how you got them. As you write your memo, make sure you touch on the following points:

- What research question are you are exploring? Explain your theories about how and why factors are related to each other.
- What is the overall story that you can tell with your data? Discuss the evidence on the research question provided by your results.
- Explain what you've learned without resorting to statistical jargon. Be sure to think about the constraints of the data, issues that need further exploration, and the needs of your audience.