

Tell A Story Using Descriptive Statistics

This assignment helps you to begin thinking about how data 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 Step 1-3 below.*
- *Pick 2 or 3 variables of related theme and find the sample means, standard deviations, medians, and histograms for them. Remember that to calculate a mean, you need a continuous variable, so be sure that at least one of your variables is continuous. This is Step 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 Step 6-8 below. [Be sure to read "Do's and Don'ts for Writing about Statistics" attached.]*

Do not hand in your print out! Just type up the results of your exercises and the answers to your discussion questions in memo format.

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. Also, each lab has SPSS manuals; learn to use the manual to answer your questions. Helpful hints for working with SPSS are available from the class web page. If you prefer to work at home, you can buy a copy of SPSS for your home computer at the University Bookstore Computer outlet (corner of University Way and 43rd).

If you would prefer to work somewhere other than the Evans School lab, bring a diskette to the lab to copy the data. Everyone will need a diskette or two to save output or new data sets you create from the original. Using a Zip disk is a good idea.

About the Data

The data for this assignment come from the 2004 Washington State Population Survey. It is a survey of the WA population similar to the national census. The original data set has 17,970 (people) observations from 7,050 households. This includes over-samples of households headed by African American, Asian, Hispanic, and Native American respondents.

Information on the data set is available at <http://www.ofm.wa.gov/sps/index.htm>. Read the "Letter to Participants" for an explanation of the structure and purpose of the study. Then look at the "Data Dictionary" to choose the variables you want to work with. The data report and the data dictionary are all Adobe Acrobat files that you can download, print, or read online. A data set with variable and value labels is available on the Evans School server in the folder \\Evanshost\CLASSES\PBAF 527- Stats\Data Sets\ containing a full dataset, a subset of this dataset, and a zipped version to make transport easier.

You'll work with these data for the policy memo due at the end of the course, so this is a good opportunity to get your feet wet.

Please note that while we provide detailed instructions for how to use SPSS, you should also make use of on-line help functions and tutorials.

For SPSS you can take a look at:

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

Get Acquainted with SPSS and the Data Set

1. Make a plan

In general, before you begin at the computer you should spend some time thinking about the data. Look at the list of variables 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

SPSS 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).

3. Read your data into SPSS

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 in
\\Evanshost\CLASSES\PBAF 527- Stats\Data Sets

First look to see if there is directory for pbaf527-stats (look for an icon on the desktop too). If so, click on it and look for the file.

If not, then 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 floppy or Zip disk, name the file something new, such as a: wapop2.sav.

You should see the data in a spreadsheet in the data window. Each row in the spreadsheet is a different survey respondent (case). Each column shows the value of a different variable; the labels are in the top row. See the variable list to match the labels to the variable definitions.

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. Confirm that there are in fact 17,970 observations. That way you can see if there are any weird data values or unexpected distributions.

SPSS 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 of the label. That variable name will match your variable list or data dictionary.

[If you want to only see the variable names (this may be easier), 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.

5. Getting Graphics

SPSS To see a *histogram* or a *bar chart* so you can 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 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 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. 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

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].

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. In addition, it is highly recommended that you save your work as you go in case disaster strikes.

SPSS 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 diskette, 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:SPS2000_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 attached "Do's and Don'ts for Writing about Statistics".]

When you're done with your initial computing, study the results and start writing. Usually, 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 variables you've just looked at. Your goal is to interpret your findings for a non-technical policy maker. As you write your memo, make sure you touch on the following points:

Explain the variables you are interested in exploring and your theories about how and why these are related to each other.

Explain what the descriptive statistics say about your particular variables and the distribution of their values.

What was your reasoning for choosing to use a mean versus a median or median rather than a mean? Why? For which of your variables is the standard deviation appropriate? Why? What other measures of dispersion might you use and why?

What do histograms or bar charts say about your variables? How do these graphics change or enforce the impression you got from the descriptive statistics?

What is the overall story that you can tell with your data? 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.

Do's and Don'ts for Writing About Statistics

Empirical information can be compelling and interesting when combined with a model and understanding of the world. However, it's your job to weave the statistics into a narrative. Data are pictures that help you "tell your story"; they are not the story.

DO:

Use numbers to support your argument, not to make it. Don't write about the numbers; write about ideas and hypotheses.

Use pretty pictures to keep your readers attention. Graphics (pie charts, histograms, bar charts, etc.) can help tell your story visually. Some people can most easily use information presented graphically.

Explain every statistic or graphic you use in the text. If you can't find room in the text for the explanation, ask yourself if the numbers belong there.

Provide enough information so a statistician can evaluate your work. Use appendices and footnotes for technical details of survey or analysis. This gives future researchers clues on how they can study the issue and on the meaning of your work.

Provide a clear summary of your conclusions drawn from empirical info. It's your job to synthesize all of the data and information. Your one page executive summary may be the only thing your client reads.

Acknowledge the shortcomings of your data and methods. No data or study is perfect for the purpose; yours will be no exception. Part of your job is to explain the uncertainties and caveats associated with your study. What can't we know from this study? What data, samples, or methodologies could provide better information?

DON'T:

Don't try to teach statistics. Busy policy-makers can't wade through a treatise on statistics. Your job is to provide the "translation", not a lecture on the definition of a confidence interval or a p-value.

Don't use the text to describe how you created the statistics. Most clients will not care about your process only your outcome. Data collection and your analysis methods should be described in an appendix, with only the important information in the text.

Don't use a cookie-cutter writing style to describe your outcomes. (E.g., For the eastern WA households, the mean was 10 and the standard deviation was 20. For the western WA households, the mean was 13 and the standard deviation was 15.) Vary your sentence construction and style to create interest. Concentrate on "telling your story" to facilitate this process.

General tips for memo writing:

Make your sentences short and direct. Policy makers are busy people--don't make them wade through endless clauses and qualifying statements. Make your sentences active, clear, and concise. This is not an academic paper and there is no extra credit for flowery phrases or complex clauses. It's your job to make the paper readable and easy to understand.

Headings and short sections give direction to reader. Bold or underlined headings allow the reader to skim and will emphasize important points. By creating short sections you will keep the reader interested and keep your analysis on track.

Put important points first. Don't hide the information your client needs by prefacing it with details. This is not a mystery--you don't have to build your argument for 3 paragraphs before presenting it. Also, extended quotations and repeated citations and references are unnecessary. Put sources in footnotes or appendices, unless important to client.

Don't forget your audience. Only provide information the client needs. Interesting facts don't belong in the analysis unless they will be useful to the client. Don't forget who your client is and what info they need.

Use appendices for tangential or extended materials and information. Your analysis should be lean--just the facts. Use appendices for information you think client could skip, but may want or need. That way, the flow of your narrative is not disturbed and the client can choose to read appendices now, later, or not at all.

Appearance counts. Make your paper look good by using sections, headings, a readable font, and useful graphics. And, don't forget to spell-check. It may seem trivial, but misspelled words won't engender confidence in your analytic abilities.

Writing takes practice. Don't expect your writing to be perfect on the first draft. It takes time to make words serve your purpose. Good rewriting is the key.