

Project 2: Exploring Global Student Values with Descriptive Statistics

Contents:

- Description and Required Materials (Page 1)
- Project Instructions (Page 2)
- Grading Rubric (Page 4)
- Project Example (Partial) (Page 5)
- APPENDIX A: Variable Codebook (Page 7)
- APPENDIX B: SPSS Analysis Instructions (Page 11)

PROJECT 2 DESCRIPTION

For this project, we are exploring the global social values of our classmates. You will be asked to analyze three (3) different variables taken from Pew Research Center's survey on Global Attitudes and Trends (GAT) For each variable, you will need to compose a detailed statistical report, a frequency distribution (lesson 2), a graph/chart (lesson 2), a central tendency analysis (lesson 3), and a variability analysis (lesson 4). This project emphasizes learning goals and objectives 1, 2, 3, and 7 as listed in the syllabus.

1. the vocabulary and symbols used in social statistics
2. how to measure variables and test relationships at different levels
3. the basics of descriptive and inferential statistics
4. to become critical consumers of statistical information
5. about global systems and issues and associated dimensions (e.g., historical, political, economic, social, cultural, environmental, technological)
6. to analyze global interrelationships and interdependencies across place and time
7. to formally report findings from statistical analyses

PROJECT 2 MATERIALS

For this project, you will need the following items:

- Your NOTES from lessons one through four
- A word processing program to type up your final REPORT (e.g., Microsoft Word)
- The SPSS data analysis program
- The Project 2 data file (STA2122 Project 2 Data .sav) posted to Canvas
- The VARIABLE CODEBOOK (APPENDIX A) containing a list of the variables and values you may use
- The SPSS ANALYSIS INSTRUCTIONS (APPENDIX B) which explain how to use SPSS

PROJECT INSTRUCTIONS

NOTE: This project is unrelated to project #1.

Sample: Undergraduates enrolled in Dr. Toothman's online social science statistics course (n = 94)

Sampling Procedure: Convenience sample taken online with offer of extra credit.

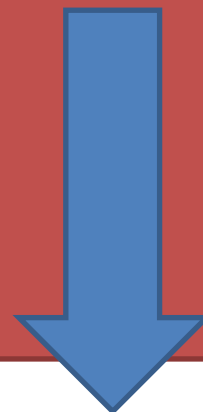
Target Population: Undergraduate students enrolled in U.S. colleges.

You should start this project by reviewing the codebook (APPENDIX A) and selecting the following types of variables: One variable measured at the *nominal level*, one variable measured at the *ordinal level*, and one variable measured at the *interval-ratio level*. You can determine the level of measurement for each variable by examining the "Measurement" column in the codebook or the questionnaire. Below, I outline how the report should be organized followed by a description of each component.

Project Organization At-a-Glance:

- **Part 1: Nominal Variable Analysis** (1-2 pages)
 - Executive summary (15 points)
 - Frequency table and a graph/chart (5 points)
 - Central tendency report (5 points)
 - Variability report (5 points)
- **Part 2: Ordinal Variable Analysis** (1-2 pages)
 - Executive summary (15 points)
 - Frequency table and a graph/chart (5 points)
 - Central tendency report (5 points)
 - Variability report (5 points)
- **Part 3: Interval/Ratio Variable Analysis** (1-2 pages)
 - Executive summary (15 points)
 - Frequency table and a graph/chart (5 points)
 - Central tendency report (5 points)
 - Variability report (5 points)

Each of these components
is described below!



Component Requirements:

- **Executive summary** (1/2 page or 2 full paragraphs). Often in research reports, you'll read about the data and then refer to the tables and charts for a better understanding of the numbers discussed. This executive summary is space for you to **introduce, describe, analyze** and **reflect** on your statistics. Your summary should be close to one-half of a page, or, two full paragraphs. Here is what all that jargon means:
 - **Introduce:** Introductions simply contain the following descriptions: the sample, the target population, the variable, its level of measurement, the question used to measure the variable, and its response categories (if applicable).
 - **Describe:** Descriptions might include a summary of the frequencies (e.g., how many people responded in each category/level), measures of central tendency, and measures of variability. You should specifically refer to your frequency tables and graphs/charts here. *HINT: If you're describing a scale (e.g., measured 1-10) you have to tell your audience what the average indicates on that scale – don't just report the value and assume they know what you mean!*

- **Analyze.** Analyses include in-depth discussions of smaller bits of information. Consider the following here: How and why is one category more popular than another (could be the sample or the way it's measured)? What does the measure of central tendency indicate about your sample/population? What about the measure of variability? What are the benefits/drawbacks of each of these measures? Can you infer information about any type of population from this sample? Why? Why not?
 - **Reflect.** Reflections allow us to step outside of the parameters of the data and tools we have in front of us and think more broadly. This is a space where you can assert ideas about what things could be improved, what future tests might reveal, *et cetera*. This is a great space for you to demonstrate the things you've learned so far in the class (i.e., lessons 1 through 6).
-
- **Frequency table and a graph/chart.** You will need to reproduce (make your own) frequency table that contains the following:
 - A title (e.g., "Table 1: Nominal Variable – Favorite Political Candidate, N=50").
 - All variable categories, frequencies, percentages, and cumulative percentages.
 - For a review of frequency tables, refer to LESSON 2.
 - For instructions on how to make a frequency table in SPSS, refer to APPENDIX B.
 - You MAY NOT copy and paste this table from SPSS.
 - And you also will need to produce the best type of graph or chart to display the data you've collected on each variable. Each graph should:
 - Be the most appropriate type of graph for each variable's level of measurement
 - Contain a title (e.g., "Figure 1: Nominal Variable – Favorite Political Candidate, N=50")
 - For instructions on how to make a frequency table in SPSS, refer to APPENDIX B.
 - You MAY copy and paste this table from SPSS.
-
- **Central tendency report.** Identify and describe the most appropriate measure of central tendency. The selected measure of central tendency should be the most precise or optimal for each variable based on level of measurement. This part of your report should contain the following:
 - Three-to-five sentences describing the best-fitting measure of central tendency (e.g., mean, median, or mode), what the value you calculated means, and why it's appropriate for this type of variable.
 - For a review of central tendency measures, refer to LESSON 3.
 - For instructions on how to compute central tendency in SPSS, refer to APPENDIX B.
-
- **Variability report.** Identify and describe the most appropriate measure of variability. Select and report the most appropriate measure of variability for each of your three variables, if applicable. This part of your report should contain the following:
 - Three-to-five sentences describing the best-fitting measure of variability you selected (e.g., IQV, IQR, range, or standard deviation), what the value you calculated means, and why it's appropriate for this type of variable.
 - For a review of variability measures, refer to LESSON 4.
 - For instructions on how to compute variability in SPSS, refer to APPENDIX B.

PROJECT 2 GRADING RUBRIC

Executive summary (15 points x3)

- **Exceeds expectations.** Summary demonstrates mastery of subject matter and is over two paragraphs in length. This part includes statistical language that is specific, clear, and thorough. The variable is described in detail, the statistics (e.g., frequency, graph, central tendency, and variability) are correctly explained and analyzed, and a reflection is provided that demonstrates deep thinking about measurement, data, survey, sample, etc.
 - **Meets expectations.** Summary demonstrates basic understanding of subject matter and is mostly two paragraphs in length. This part will contain some basic statistical language. The variable is described in, the statistics (e.g., frequency, graph, central tendency, and variability) are generally explained and analyzed, and a reflection is provided. Some of this part may be stilted or superficial.
 - **Approaches/below expectations.** Summary demonstrates little-to-no understanding of subject matter or is mostly under paragraphs in length. The variable may only be mentioned, the statistics may be provided but with no explanation, an analysis may only touch on one aspect of the variable, or the reflection is missing or brief.
-

Frequency distribution and graph/chart (5 points x3)

- **Exceeds expectations.** The frequency distribution and the graph/chart include all required components (e.g., titles, labels, frequencies, percentages). If class-intervals should be used, they have been used appropriately. The selected graph is the MOST appropriate for the variable's level of measurement. Both are informative. Polished effort.
 - **Meets expectations.** The frequency distribution and the graph/chart include most of the required components (e.g., titles, labels, frequencies, percentages). The selected graph may not be the most appropriate for the variable's level of measurement. Basic effort.
 - **Approaches/below expectations.** More than two major required components are missing. Lower effort demonstrated.
-

Central tendency report (5 points x3)

- **Exceeds expectations.** The measure of central tendency is the most precise choice for the variable's level of measurement and correctly described (e.g., what is a mode?). The computed value or category is correct and explained (e.g., what does a mode of "2" indicate about our data?). An explanation of why this measure of central tendency was chosen over the others is included (3-5 sentences).
 - **Meets expectations.** The measure of central tendency is not the most precise choice for the variable's level of measurement, but the description is apt (e.g., what is a mode?). The computed value or category is explained (e.g., what does a mode of "2" indicate about our data?). An explanation of why this measure of central tendency was chosen over the others is included (but it may be more or less correct).
 - **Approaches/below expectations.** The measure of central tendency is not applicable precise choice and there is insufficient discussion about the selection. The computed value is missing, incorrect, or underexplained.
-

Variability report (5 points x3)

- **Exceeds expectations.** The measure of variability is the most precise choice for the variable's level of measurement and correctly described (e.g., what is a standard deviation?). The computed value or category is correct and explained (e.g., what does a standard deviation of "2" indicate about our data?). An explanation of why this measure of variability was chosen over the others is included (3-5 sentences).
 - **Meets expectations.** The measure of variability is not the most precise choice for the variable's level of measurement, but the description is apt (e.g., what is a standard deviation?). The computed value or category is explained (e.g., what does a standard deviation of "2" indicate about our data?). An explanation of why this measure of variability was chosen over the others is included (but it may be more or less correct).
 - **Approaches/below expectations.** The measure of variability is not applicable precise choice and there is insufficient discussion about the selection. The computed value is missing, incorrect, or underexplained.
-

Presentation quality (10 points)

- **Check-plus.** Report is polished, organized, adheres to standard written English, and provides all applicable external resources. There are paragraphs with spaces, appropriate numbering, and subheadings. Contains minimal spelling and grammatical errors. *A person who is not in this class could pick up this report, read it, and understand it.*
- **Check.** Report is clear, adheres to standard written English, and provides all applicable external resources. There are paragraphs with spaces, appropriate numbering, and subheadings. There may be some formatting errors that make the project difficult to read. Minimal spelling and grammatical errors.
- **Check-Minus.** Presentation is unclear. There are formatting errors that make the project difficult to read. Contains more than three spelling and major grammatical errors.

PROJECT 2 EXAMPLE (partial)

The following example is only one-third (1/3) of what is required for a full project – it shows you how PART 3 might look. Remember, your full project will contain THREE parts (representing three different variables). **IMPORTANT: Borrowing or copying the language of my example may result in a “0” for that section or for the entire report. Be careful as my example draws from entirely different data.**

PART THREE: An interval-ratio variable analysis of education in the U.S.

Executive Summary

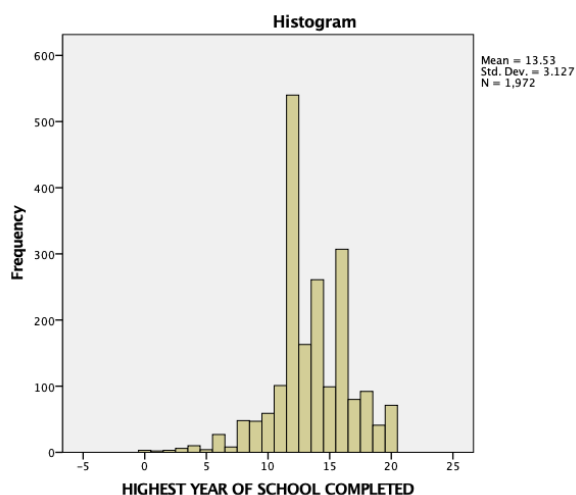
Education is a crucial part of social participation – it can impact our family trajectory, the types of jobs we can get, our health, and so much more. Below I present data on the highest level of education (measured in years) for US residents. My data is drawn from a representative sample of US residents (n=1972). The frequency table indicates that the most frequently reported year of completion of school is 12 years (which suggests high school completion). Looking at the chart, it also appears as though more people have completed more than 12 years of school than fewer than 12 years of school. Here is a sentence describing the central tendency of my data, and it looks like my data is skewed, so perhaps there is another measure to consider. This is a sentence about the variability of my data, which says something about US resident education, but also about the measure of central tendency. I that believe years of education completed may be related to parent's socioeconomic status (the higher the status may have resulted in more schooling completed by respondent) and parental education level (i.e., the more schooling completed by parent may have resulted in more schooling completed by respondent). My sample is a random sample of over 1000 US citizens and my results likely reflect the education levels of the US population; however, it may not reflect the global population which may be different due to types of education and different norms and values across countries. HINT: Remember that your sample is drawn from Dr. Toothman's statistics class. Some portions of this report are more general to prevent copying – remember, I'm looking for detail and statistical literacy.

Frequency table and Chart

TABLE 1: Frequency table of Highest Year of School Completed, n=1972

Years	Frequency	Percent	Cumulative %
0	3	.2	.2
1	2	.1	.3
2	3	.2	.4
3	6	.3	.7
4	10	.5	1.2
5	4	.2	1.4
6	27	1.4	2.8
7	8	.4	3.2
8	48	2.4	5.6
9	47	2.4	8.0
10	59	3.0	11.0
11	101	5.1	16.1
12	540	27.4	43.5
13	163	8.3	51.8
14	261	13.2	65.0
15	99	5.0	70.0
16	307	15.6	85.6
17	80	4.1	89.7
18	92	4.7	94.3
19	41	2.1	96.4
20	71	3.6	100.0
Total	1972	99.9	

FIGURE 1: Histogram of Highest Year of School Completed, n=1972



Central tendency report

HINT: You'll be reporting ONE best measure for your variable -- I'm going to report several here. If you report all three for each of your variables, you will lose points. **Please do not copy my language, I provide these stems for your reference only. Your report should also contain several sentences of your own thought.**

- The median is the value that _____. The median is the best measure of central tendency for _____ types of variables because _____. For this sample, the median is 13, which indicates 13 years of education. A median value of 13 indicates that _____.
- The mode is the value that _____. The mode is the best measure of central tendency for _____ types of variables because _____. The mode 12 years of education. A value of 12 indicates that _____.
- The mean is a value that _____. The mean is the best measure of central tendency for _____ types of variables because _____. For this sample, the mean is 13.53 years of education A value of 13.53 indicates that _____.

Variability report

HINT: You'll be reporting ONE best measure for your variable -- I'm going to report several here. If you report all three for each of your variables, you will lose points. **Please do not copy my language, I provide these stems for your reference only. Your report should also contain several sentences of your own thought.**

- The range is a value that _____. The range is the best measure of variability for _____ types of variables because _____. The range for this data is 20 _____. A value of 20 indicates that _____.
- The IQR is a value that _____. The IQR is the best measure of variability for _____ types of variables because _____. The IQR for this data is 4 _____. A value of 4 indicates that _____.
- The standard deviation is a value that _____. The standard deviation is the best measure of variability for _____ types of variables because _____. The standard deviation for this data is 3.12 _____. A value of 3.12 indicates that _____.

APPENDIX A: VARIABLE CODEBOOK

A codebook contains all possible variables in a given data set. Below is a list containing a subset of questions (i.e., variables) appearing in the most recent wave of the Global Attitudes Survey that I have given to your class (for sample information, read the assignment page).

For this project, you're supposed to select one variable measured at the nominal level, one at an ordinal level, and a variable measured at an interval/ratio level. In order to determine how a variable is measured, examine the response categories! If you want to check your variable selection, send an email to us!

Variable/Question/Measurement as it appears on survey	
(a1) In years, how old were you at your last birthday? • Scale 0 to 99	 Oh! A scale is our way of measuring a variable on an interval-ratio level.
(a2) Which best describes your gender? • Man (1) • Woman (2) • Trans Man (3) • Trans Woman (4) • Other (5)	 Some categories for nominal and ordinal variables may be numbered to help us keep track of them in SPSS. This doesn't change their level of measurement!
(a3) Please indicate your major ethnic group membership: • White (e.g., Caucasian, European, Irish, Italian, Arab, Middle Eastern) (a3_1) • Black or African-American (e.g., Negro, Kenyan, Nigerian, Haitian) (a3_2) • South Asian, or Asian-American (e.g., Indian, Pakistani) (a3_3) • East Asian, or Asian-American (e.g., Chinese, Filipino) (a3_4) • South Asian, or Asian-American (e.g., Arabic) (a3_5) • Native American (a3_6) • Hispanic/Latino (e.g., Mexican, Puerto Rican, Cuban) (a3_7)	
(a4) Which best describes your current marital status? (code MARITAL) • Married (1) • Living together as married (2) • Divorced (7) • Separated (3) • Widowed (4) • Single (5)	
(a5) How many children have you have? • Scale 0 to 8	

(a6) Independently of whether you attend religious services or not, would you say you are • A religious person (1) • Not a religious person (2) • An atheist (3)
(a7) In political matters, people talk of "the left" and "the right." How would you place your views on this scale, generally speaking? • Scale 1 to 10 (1 = Left; 10 = Right)
(a8) Are you employed now or not? (If more than one job: only for the main job) • Full time employee (30 or more hours per week) (1) • Part time employee (fewer than 30 hours per week) (2) • Self employed (3) • Retired/pensioned (4) • Homemaker (5) • Student (6) • Unemployed (7) • Other (8)
(a9) People sometimes describe themselves as belonging to the working class, the middle class, or the upper or lower class. Would you describe yourself as belonging to the • Upper Class (1) • Upper Middle Class (2) • Lower Middle Class (3) • Working Class (4) • Lower Class (5)
(a10) All in all, how would you describe your state of health these days? • Very Good (1) • Good (2) • Fair (3) • Poor (4)
On this list are various groups of people. Could you please mention any that you would not like to have as neighbors? _____... • Would not like to have as neighbors (1) • Would not mind having as neighbors (2)
(a11_1) Drug addicts
(a11_2) People of a different race
(a11_3) People who have AIDS
(a11_4) Immigrants/Foreign Workers
(a11_5) Homosexuals
(a11_6) People of a different religion
(a11_7) Heavy drinkers
(a11_8) Unmarried couples living together
(a11_9) People who speak a different language

Now I will briefly describe some people. Using this card, would you please indicate for each description whether that person is very much like you, like you, somewhat like you, not like you, or not at all like you? • Very much like me (1) • Like me (2) • Somewhat like me (3) • A little like me (4) • Not at all like me (5)
(a12_1) It is important to this person to think up new ideas and be creative; to do things one's own way.
(a12_2) It is important to this person to be rich; to have a lot of money and expensive things.
(a12_3) Living in secure surroundings is important to this person; to avoid anything that might be dangerous.
(a12_4) It is important to this person to have a good time; to "spoil" oneself.
(a12_5) It is important to this person to do something for the good of society.
(a12_6) It is important for this people to help the people nearby; to care for their well-being.
(a13) Please indicate which of the following problems you consider the most serious one for the world as a whole? (Interviewer: mark only one). • People living in poverty and need (1) • Discrimination against girls and women (2) • Poor sanitation and infectious diseases (3) • Inadequate education (4) • Environmental pollution (5)
(a14) Here are two statements people sometimes make when discussing the environment and economic growth. Which of them comes closer to your own point of view? • Protecting the environment should be given priority, even if it causes slower economic growth and some loss of jobs. (1) • Economic growth and creating jobs should be the top priority, even if the environment suffers to some extent. (4) • Other. (5)
I am going to name a number of organizations. For each one, could you tell me how much confidence you have in them: is it a great deal of confidence, quite a lot of confidence, not very much confidence or none at all? • A great deal (1) • Quite a lot (2) • Not very much (3) • None at all (4)
(a15_1) The churches
(a15_2) The armed forces
(a15_3) The press
(a15_4) Television
(a15_5) Labor unions

(a15_6) The police
(a15_7) The courts
(a15_8) The government (nation)
(a15_9) Political Parties
(a15_10) Parliament
(a15_11) The Civil service
(a15_12) Universities
(a15_13) Major Companies
(a15_14) Banks
(a15_15) Environmental Organizations
(a15_16) Women's organizations
(a15_17) Charitable or humanitarian organizations
In the last 12 months, how often have you or your family • Often (1) • Sometimes (2) • Rarely (3) • Never (4) • NA (5)
(a16_1) Gone without enough food to eat
(a16_2) Felt unsafe from crime in your home
(a16_3) Gone without medicine or medical treatment that you needed
(a16_4) Gone without a cash income
The questions/variables listed below (49 through 53) are measured/coded as follows: Now, I would like to read some statements and ask how much you agree or disagree with each of these statements. For these questions, responses range from a "1" meaning that you "completely disagree," to "10" meaning that you "completely agree." • Scale 1 to 10
(a17_1) Science and technology are making our lives healthier, easier, and more comfortable.
(a17_2) Because of science and technology, there will be more opportunities for the next generation.
(a17_3) We depend too much on science and not enough on faith.
(a17_4) One of the bad effects of science is that it breaks down people's ideas of right and wrong.
(a17_5) It is not important for me to know about science in my daily life.

APPENDIX B: SPSS ANALYSIS INSTRUCTIONS

This section contains the following instructions:

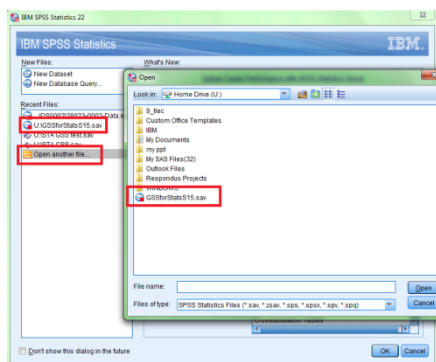
1. How to open data files in SPSS
2. How to view Variables by 'Variable Name' rather than 'Variable Label'
3. How to make frequency distributions
4. How to find measures of central tendency and variability
 - a. Mean, Standard Deviation, Variance, Range
 - b. Median and Mode
 - c. Interquartile Range
5. Data visualizations
 - a. Bar Chart
 - b. Pie Chart
 - c. Line Graphs
 - d. Histograms

Toothman STA2122 Project 2

1. Opening Your Data

After SPSS is installed on your computer, open the program.

Next, open the `examplename.sav` file (this file is posted available on Canvas). SPSS should open a dialogue that shows your “Recent Files” and the option to “Open another file....” If you have recently opened up the data set, you can just click on it and then say OK. If not, click “Open another file...”, find the file, and then click “Open.”

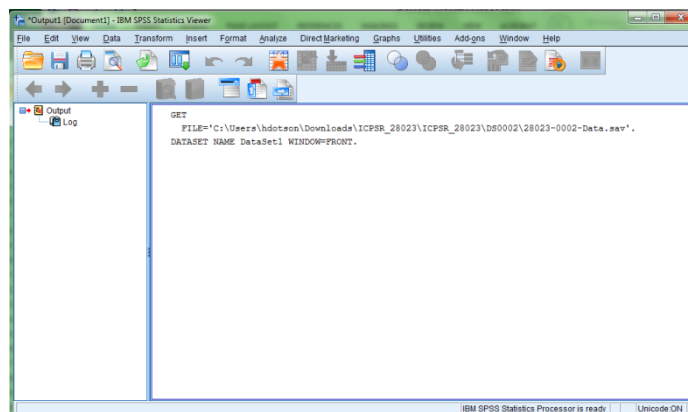


The data in this example is labeled `GSSforStats.sav` – however your data could have a different name!!!

Once the file is open, you should be able to see all of your data. This is called the “Data Editor” view. I’ve included a screenshot for what this should look like, using a different data set.

[illegible]

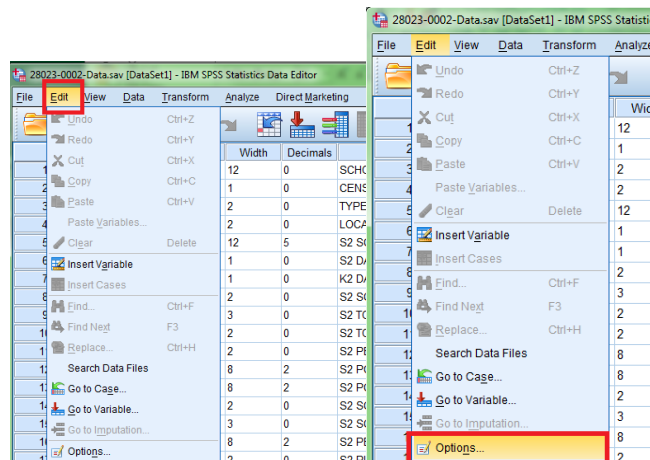
SPSS will also open up another window, called the “Output Viewer.” When you run your analyses, they will appear here.



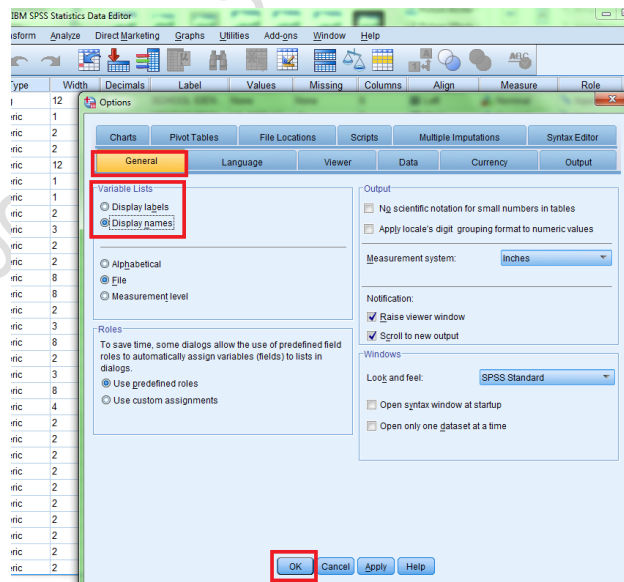
2. How to view Variables by 'Variable Name' rather than 'Variable Label'

To make your life easier, I strongly recommend setting up SPSS to show you Variable Names rather than Variable Labels. When conducting your analyses, it will be much easier to find your variables this way.

To do this, first click on “Edit” under the Data Editor viewer. Next, click “Options.”

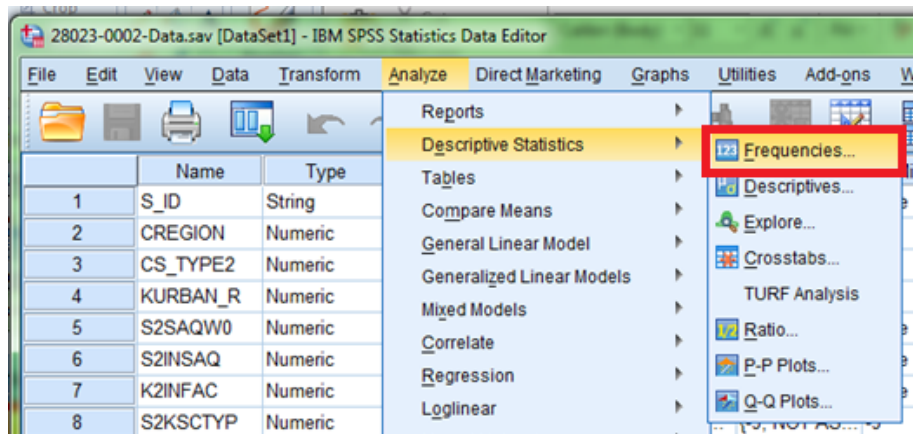


SPSS will bring up a new dialogue window. From here, click on the “General” tab at the top, then look for the “Variable Lists” box. Click on “Display names.” Then “OK.”

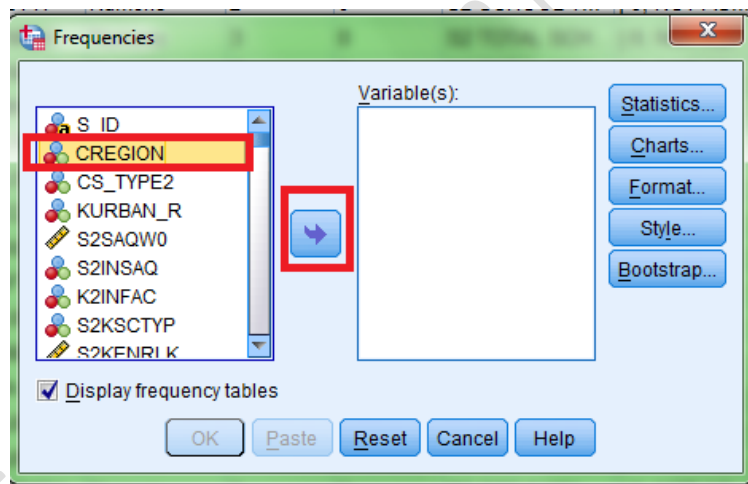


3. How to make frequency distributions

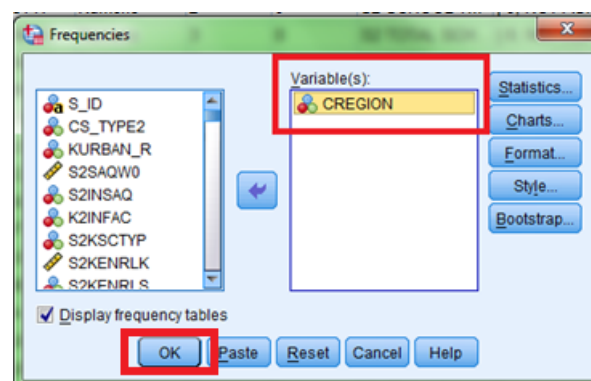
To run a frequency distribution, click on “Analyze”, then “Descriptive Statistics”, and then “Frequencies....”



Next, find the variable you want to analyze. It will show the ‘variable name’ in your list. For this example, I will analyze the variable CREGION. Click on the variable name you would like to analyze (it will highlight the variable), and then click on the → arrow.



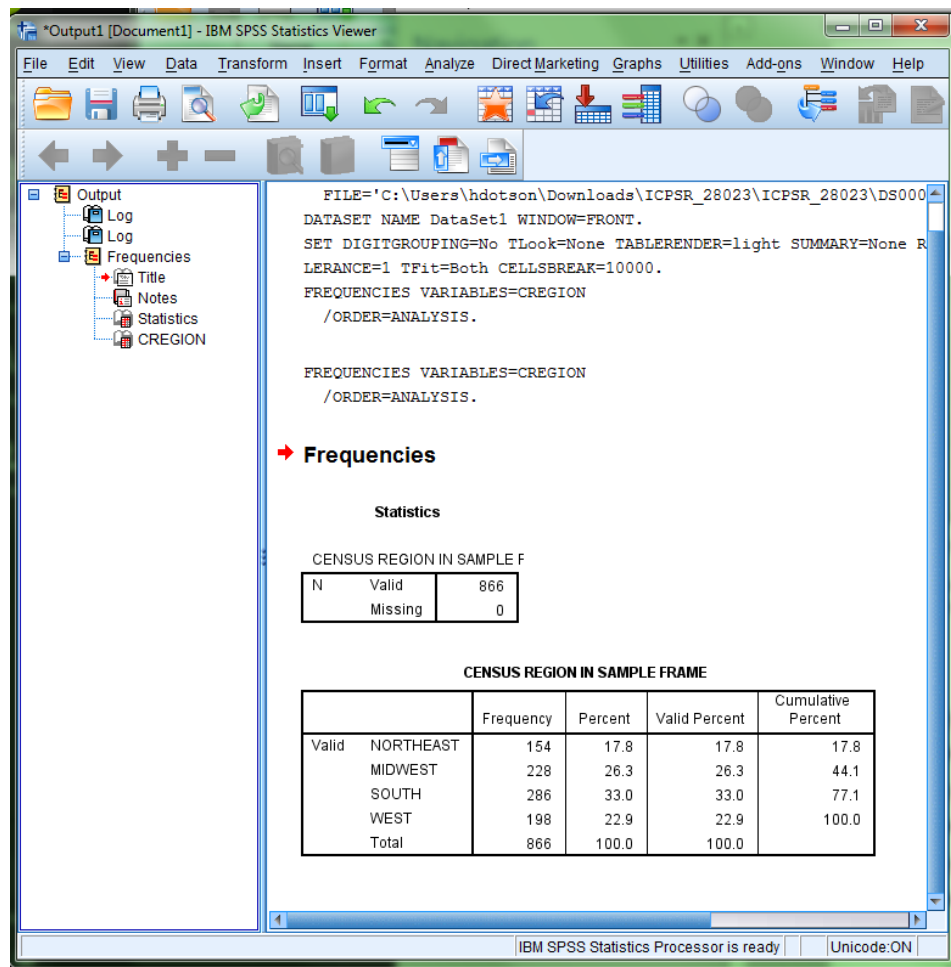
Notice what happens after you click the → arrow:



Then click “OK” to run the frequency distribution.

Helpful tip: you can run more than one frequency distribution if you would like, just select more variables. For simplicity, and until you are comfortable, I recommend doing one at a time.

The frequency distribution will appear in the Output Viewer window.



FILE='C:\Users\hdotson\Downloads\ICPSR_28023\ICPSR_28023\DS000
 DATASET NAME DataSet1 WINDOW=FRONT.
 SET DIGITGROUPING=No Tlook=None TABLERENDER=light SUMMARY=None R
 LERANCE=1 TFit=Both CELLSBREAK=10000.
 FREQUENCIES VARIABLES=CREGION
 /ORDER=ANALYSIS.

 FREQUENCIES VARIABLES=CREGION
 /ORDER=ANALYSIS.

➔ **Frequencies**

Statistics

CENSUS REGION IN SAMPLE F

N	Valid	866
	Missing	0

CENSUS REGION IN SAMPLE FRAME

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid NORTHEAST	154	17.8	17.8	17.8
MIDWEST	228	26.3	26.3	44.1
SOUTH	286	33.0	33.0	77.1
WEST	198	22.9	22.9	100.0
Total	866	100.0	100.0	

IBM SPSS Statistics Processor is ready | Unicode:ON

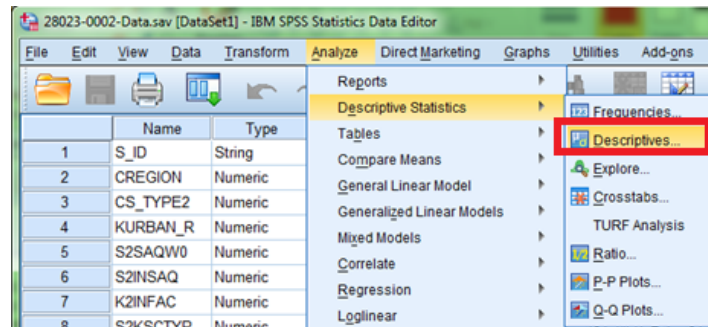
Look at all that has been presented. We can see that N=866. We can also see how many cases were in the Northeast, Midwest, South, and West. We can also see the percent of cases and the cumulative percent. “Valid Percent” should be the same as “Percent” in your data set, since I have manually removed all missing cases (that’s also why ‘Missing’ will equal zero in the GSS data set I have given you – in real life, we have to figure out what to do with missing cases, but that is beyond the scope of this course).

4. How to find measures of central tendency and variability

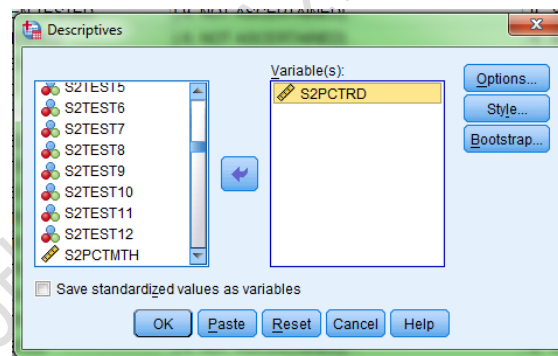
SPSS will also run all of the measures of central tendency and variation you need. However, that does not mean that a particular test is the 'right' one. It is your job to make sure you use the correct measure of central tendency and measure of variation, depending on the type of variable you have. SPSS will typically run things because you tell it to, even if it is not a statistically sound test.

4a. Mean, Standard Deviation, Variance, Range

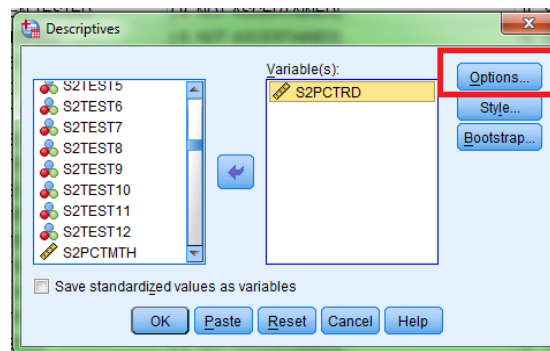
To have SPSS tell you the mean value of a variable, go to the Data Editor view. Then, click on "Analyze", "Descriptive Statistics", then "Descriptives...."



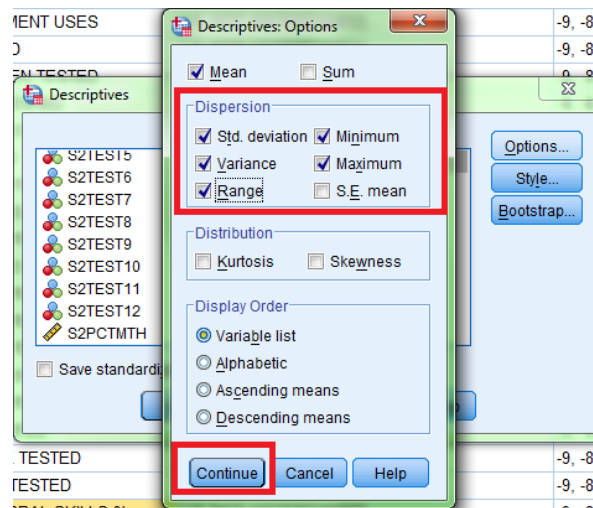
Then, SPSS will bring up the "Descriptives" dialogue. Find the variable you want to analyze, then click the → arrow to bring it to the 'Variables' column. In this case, I will be finding the mean of then S2PCTRD variable.



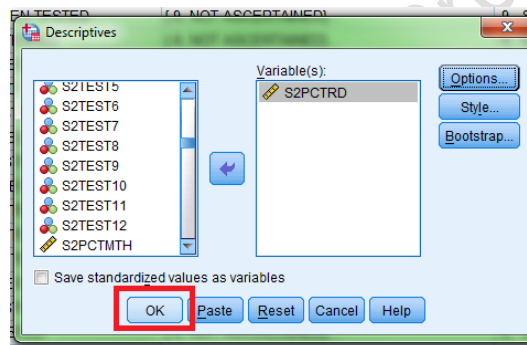
Before analyzing the data, click on "Options." This will allow you to also find out the standard deviation, variance, and range.



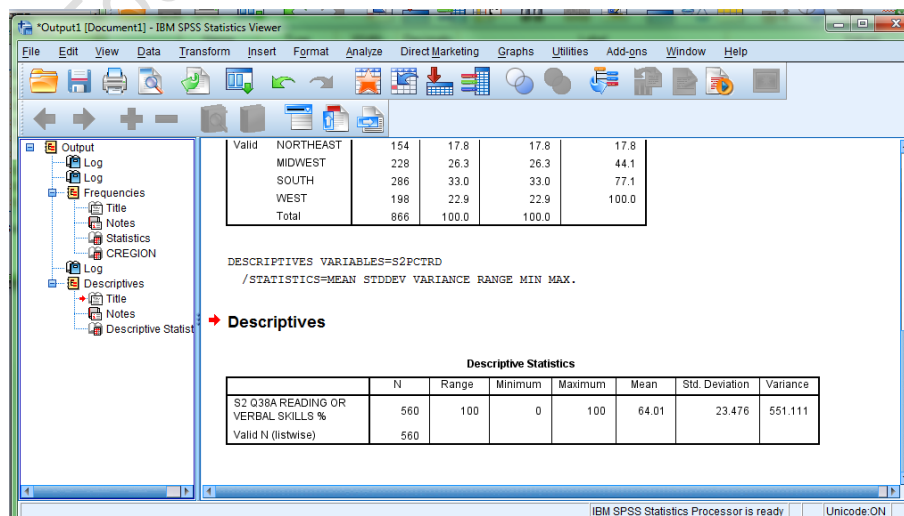
SPSS will bring up a new dialogue. Make sure to check the boxes of everything you would like to know. Once you are finished, click “Continue.”



SPSS will bring up the “Descriptives” dialogue from a moment ago. Click “OK” to finish.

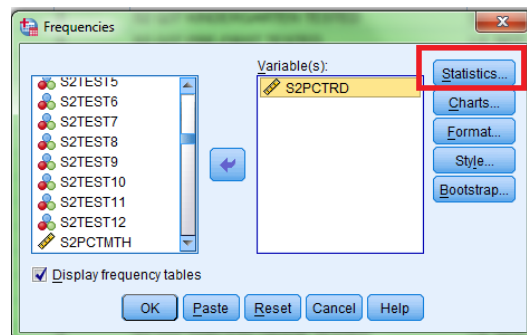


The output will appear in the Output Viewer. Notice that we see the N (560), range = 100, minimum (which is the smallest score) = 0, maximum (which is the greatest score) = 100, the mean = 64.01, the standard deviation (labeled as Std. Deviation) = 23.476, and the variance = 551.111.

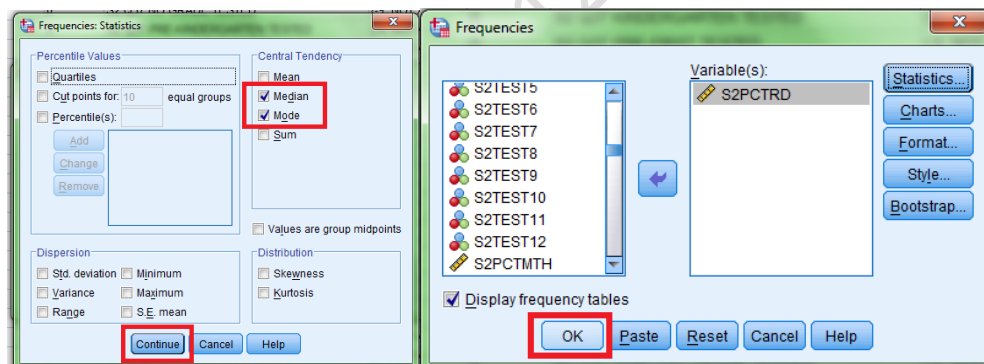


4b. Median and Mode

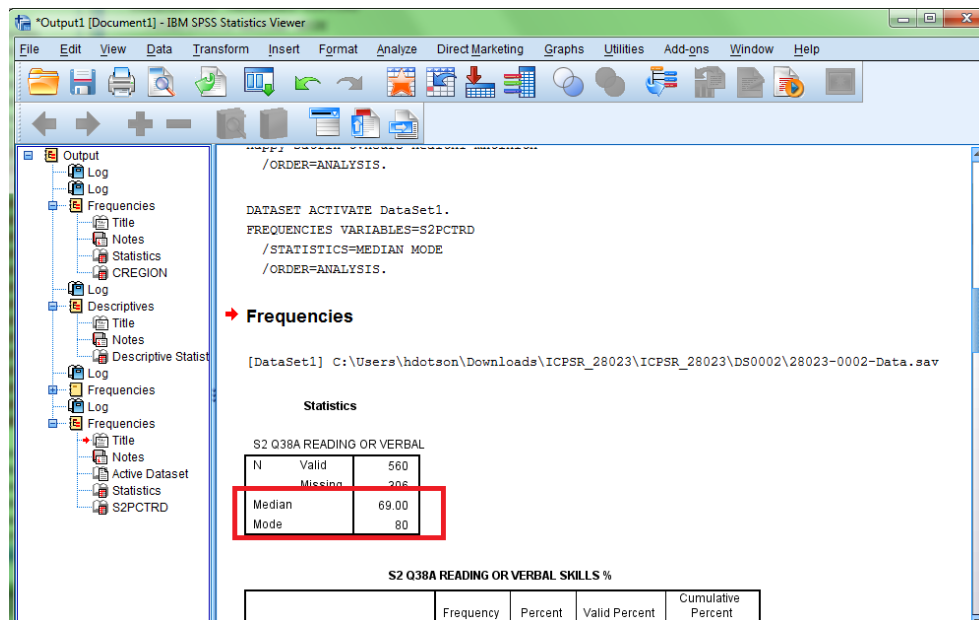
To find the median and the mode, we have to use “Frequencies” just like we had to do to find the frequency distribution. I am going to use the S2PCTRD variable here. Then, click on “Statistics...”



SPSS will bring up a new dialogue window. Look at all the options! If you want to know the Median and Mode, click on those boxes. You can also use this dialogue to find out quartiles, percentiles, standard deviation, variance, range, minimum, maximum, and the mean. Once you're done, click “Continue.” Then, click “OK”

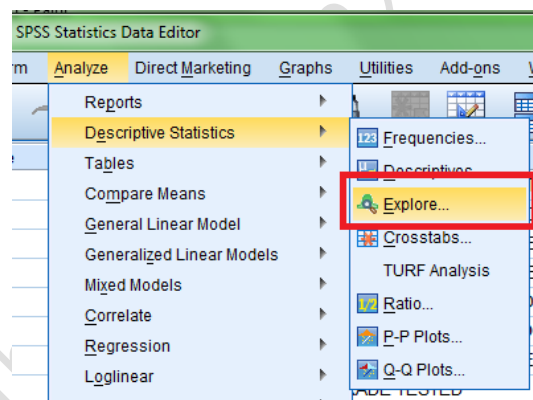


The median and mode will appear in the Output Viewer, in the “Statistics” box. The median is 69 and the mode is 80.

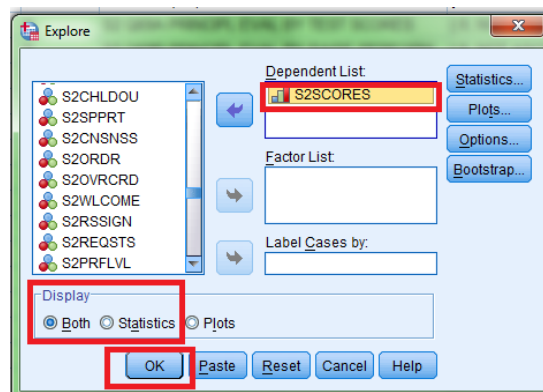


4c. Interquartile Range

To find the interquartile range, we go back to the Data Editor window, click on “Analyze”, then “Descriptive Statistics”, and finally, “Explore....”



Select the variable you want to use here. I will be using the variable S2SCORES (this variable has 3 categories). Then, make sure under “Display” you have either “Both” or “Statistics” selected. Next, click OK.



Go back to the Output viewer. In a list of statistics, the interquartile range will be reported. In this case, the IQR is 1.

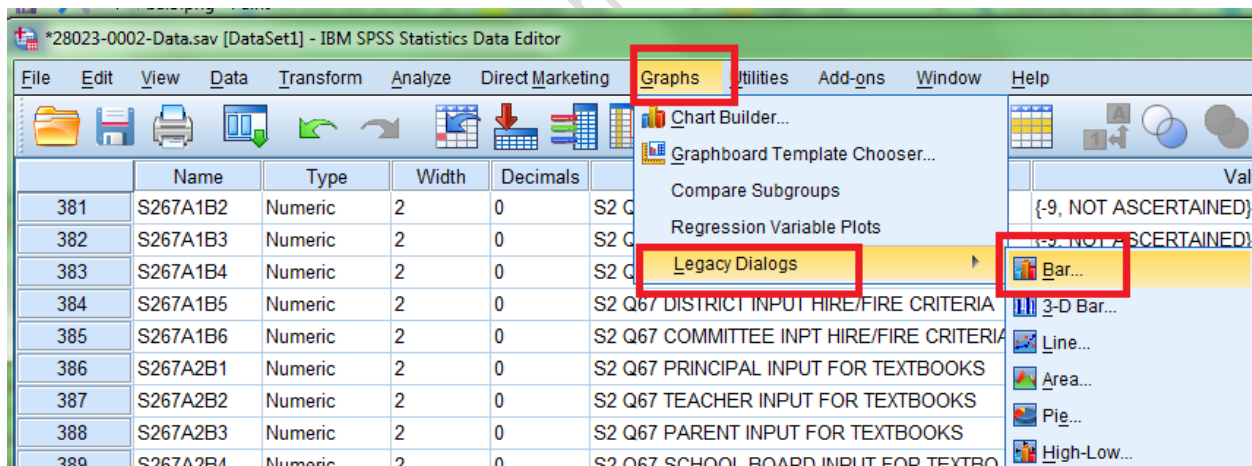
Descriptives				Statistic	Std. Error
S2 Q69A PRINCIPAL EVAL BY TEST SCORES	Mean			2.02	.025
	95% Confidence Interval for Mean	Lower Bound		1.97	
		Upper Bound		2.07	
	5% Trimmed Mean			2.02	
	Median			2.00	
	Variance			.502	
	Std. Deviation			.708	
	Minimum			1	
	Maximum			3	
	Range			2	
	Interquartile Range			1	
	Skewness			-.024	.085
	Kurtosis			-1.002	.171

5. Data Visualizations

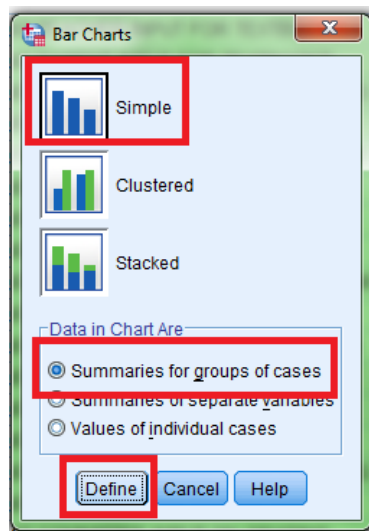
SPSS will also create professional graphics for you, using your own data. The first steps to create graphs in SPSS are the same, regardless of whether you are trying to create a bar chart, pie chart, or a histogram.

5a. Bar Chart

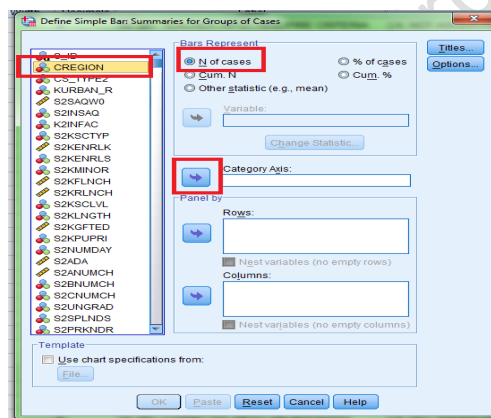
First, return to the Data Editor, click on “Graphs”, then “Legacy Dialogs”, and then “Bar....”



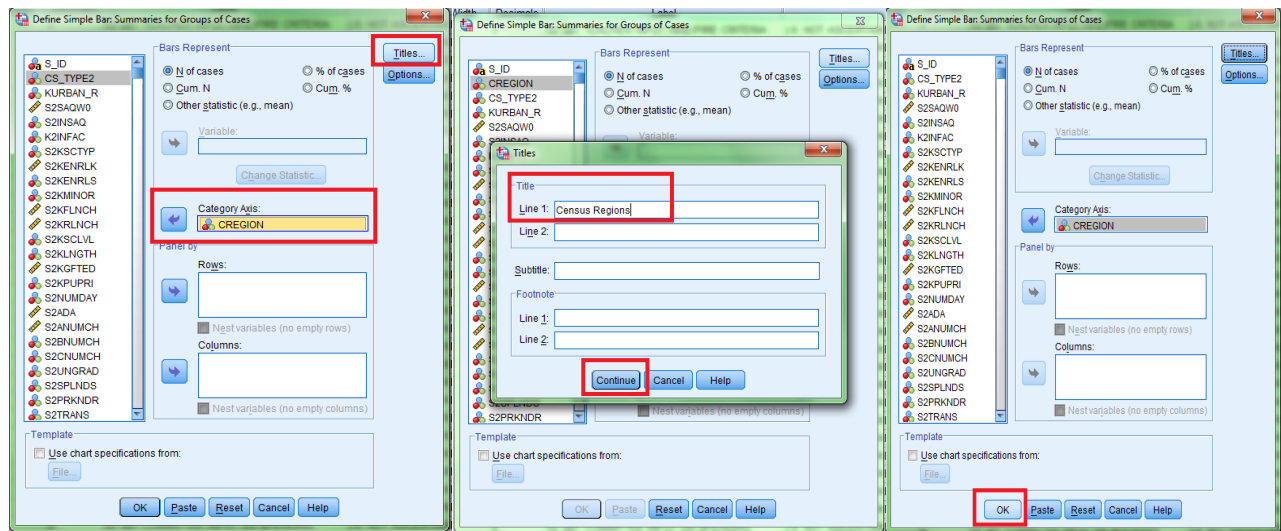
SPSS will bring up the following window. For our purposes, “Simple” is fine. Tell SPSS that you want the data in chart to be “Summaries for groups of cases” and then click “Define.”



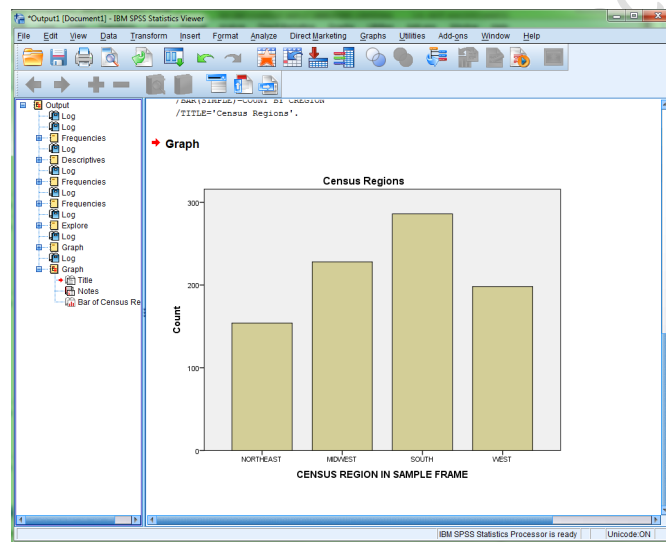
Next, find the variable that you want to use to create a bar chart. I will use CREGION. Click on the → arrow for “Category Axis”, after selecting your variable. Make sure you also have “N of Cases” selected for “Bars Represent.”



Next, we need to create Titles for our bar chart. Click on “Titles.” SPSS will bring up the following dialogue. Since I am creating a bar chart for census regions, I’m going to name the chart “Census Regions.” Click “Continue” when you are finished. This will return you back to the Simple Bar dialogue. Click “OK” to create your bar chart.

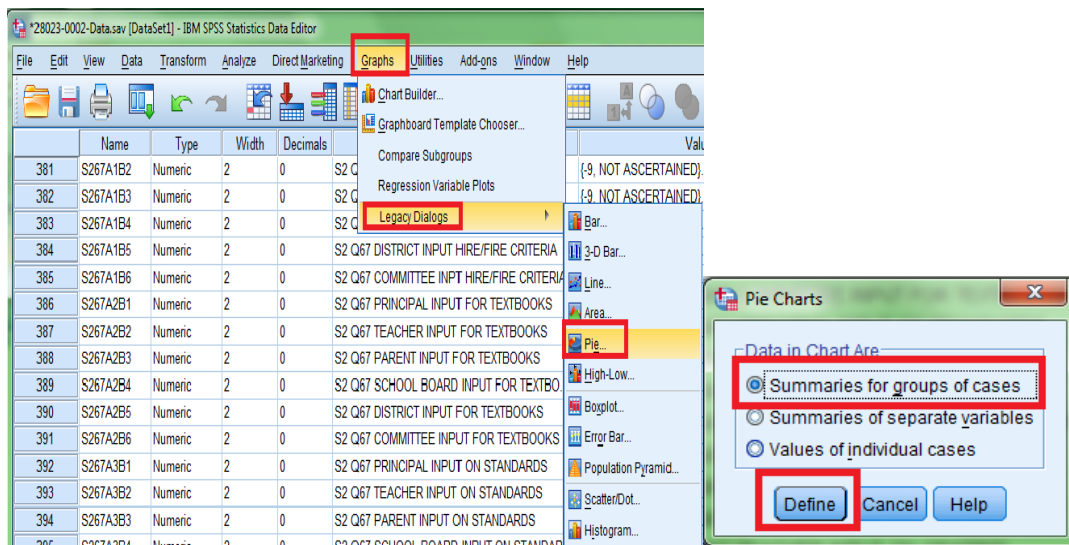


SPSS will display you bar chart in the Output viewer.

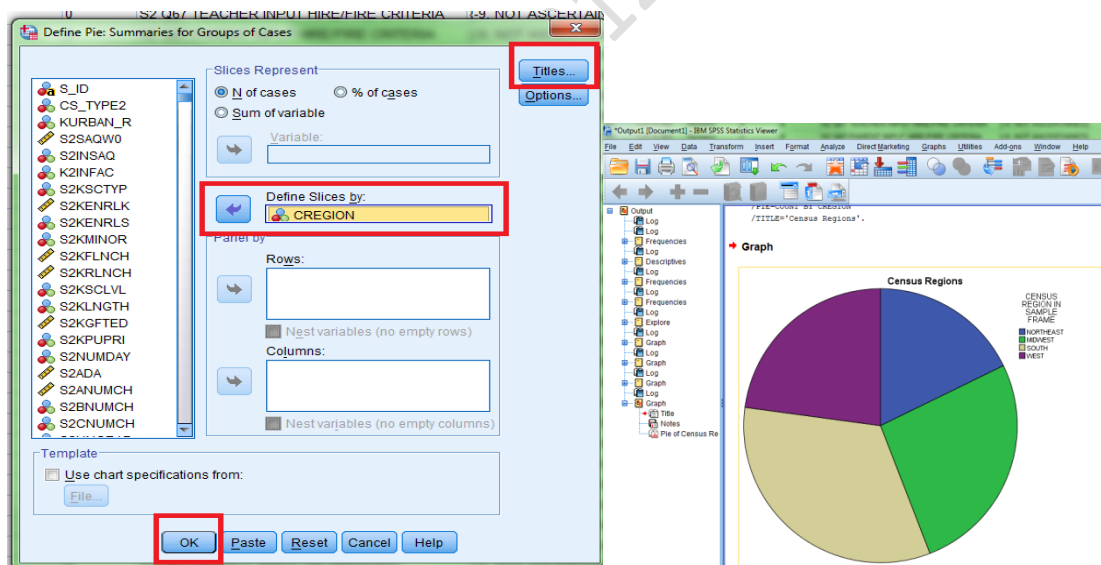


5b. Pie Charts

To create a pie chart, first click on “Graphs”, then “Legacy Dialogs” and finally, “Pie...” SPSS will bring up the following dialogue. Select “Summaries for groups of cases” and then “Define.”

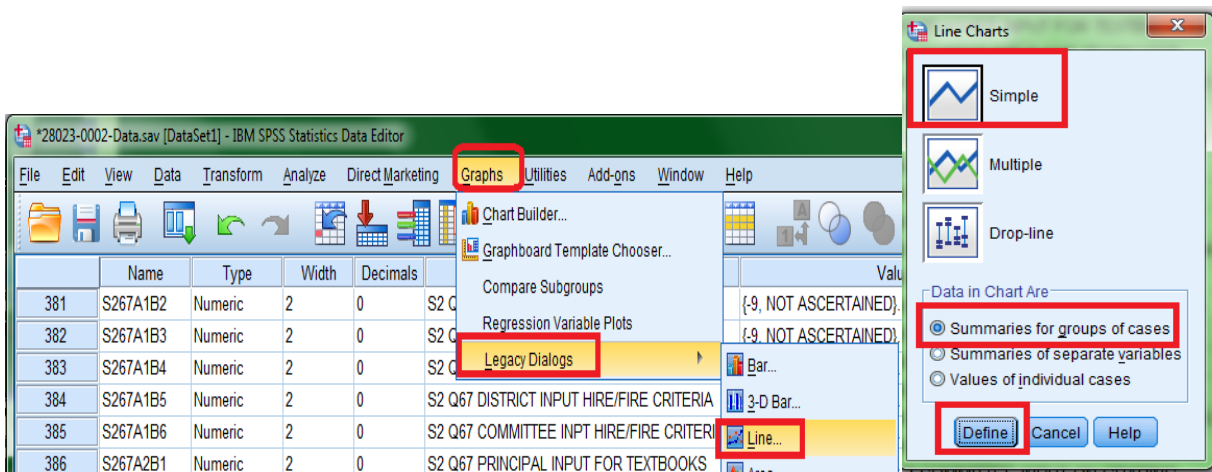


For the most part, the steps are identical to creating a bar chart. Select the variable you want to create a pie chart for, then click the → arrow below “Define Slices by:.” Create a title (as you did with a bar chart) and then click “OK.” Your pie chart will appear in the Output Viewer.



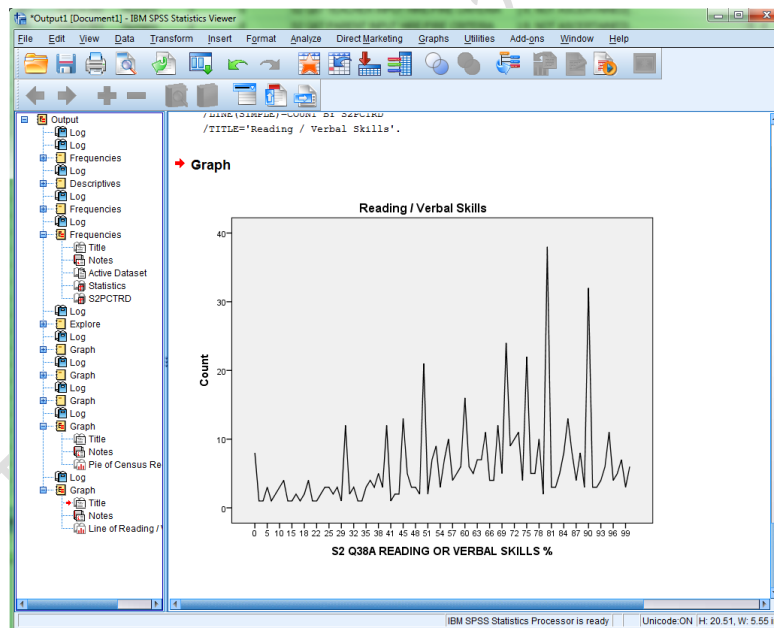
5c. Line Graphs

To create a line graph, first click “Graphs”, then “Legacy Dialogs” and finally “Line...” SPSS will bring up the following dialogue. For our purposes, a “simple” line graph is sufficient. Make sure “summaries for groups of cases” is selected.



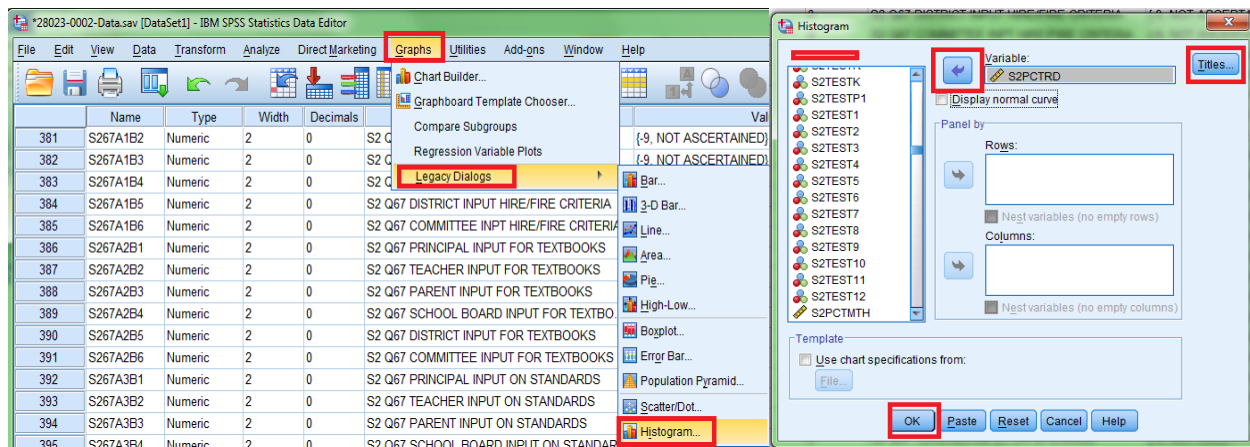
SPSS will bring up a familiar dialogue. For this, select the variable you want to create a line graph for, then click the → arrow next to “Category Axis.” In this example, I am going to use the S2PCTRD variable from earlier. Make sure you create a “Title”, and then when you are finished, click OK.

SPSS will create your line graph and it will appear in the Output Viewer.



5d. Histograms

To create a histogram, from the Data Editor window, click on “Graphs”, then “Legacy Dialogs” and finally “Histogram....” SPSS will bring up the following dialogue. Select the variable you want to use to create your histogram, and click the arrow next to “Variable.” Click “Title” to create a title for your histogram. When you are finished, click OK.



SPSS will create your histogram in the Output Viewer window.

