

SPSS 1 Assignment Instructions

***SPSS 1 is due **Wednesday, March 19th, 2020** in class (**F2F**) or by 11:59PM (**DE via SafeAssign**)

This assignment is designed to help teach you to describe a single variable – its central location, its dispersion, to create an appropriate graphic to illustrate the variable, and to discuss the way in which you variables distribute.

Use the following format:

- I. A title page with “SPSS 1: Describing a single variable” as the title and your name, section, TA’s name, professor’s name, date, G# in the upper, right hand corner.
- II. Start a new page for each variable.
- III. The variable name in **bold** and underlined at the top of the page.
- IV. Your answers should have the following sections:
 - a) A properly formatted frequency table for the variable you are describing.
 - b) A table of appropriate summary statistics.
 - c) An appropriate graphic.
 - d) A paragraph describing your variable.

For each of the variables (on the next page):

- I. Identify the level of measurement for each variable.
- II. Build a table that shows the cumulative % and frequencies.
 - a. The table must be in **APA format** (See sample below)
 - b. Some variables will have to be recoded to effectively display in a table. **RULE OF THUMB** – no more than **11** categories should appear in ANY table.
- III. Report the summary statistics that describes the variable in terms of **all the appropriate measures of central location and dispersion**. Points will be deducted for inappropriate measures. **NOTE: If recoding your variable, report summary statistics on the original, un-recoded variable.**
- IV. Create **1 appropriate** graphic to display the distribution (See sample bar chart below for an **ordinal** variable).
- V. Write a paragraph that describes the distribution in terms of central location, dispersion, outliers, and skew (if any).

SPSS 1 Assignment Instructions

These are your variables:

1. From the **WORLD 2012** dataset use the variable named “**polity**” with the label “**Higher scores more democratic (Polity)**”.
2. From the **NES 2012** dataset use the variable named “**dem_marital**” with the label “**Marital Status**”.
3. From the **NES 2012** dataset use the variable named “**relig_attend**” with the label “**Attendance: Religious Services**”.
4. From the **GSS 2012** dataset use the variable named “**science_quiz**” with the label “**Number of answers correct on a science quiz**”.
5. From the **GSS 2012** dataset use the variable named “**educ_4**” with the label “**Education in 4 Categories**”.

Sample Problem

Variable: age5

- I. Level of measurement for this variable is **ordinal**.
- II. Cumulative % frequency table:

Age in 5 Categories ^a

X	F	%	Cum %
18-30	437	21.7	21.7
31-40	384	19.1	40.8
41-50	403	20.0	60.8
51-60	369	18.3	79.1
61+	421	20.9	100.0
Total	2015	100.0	

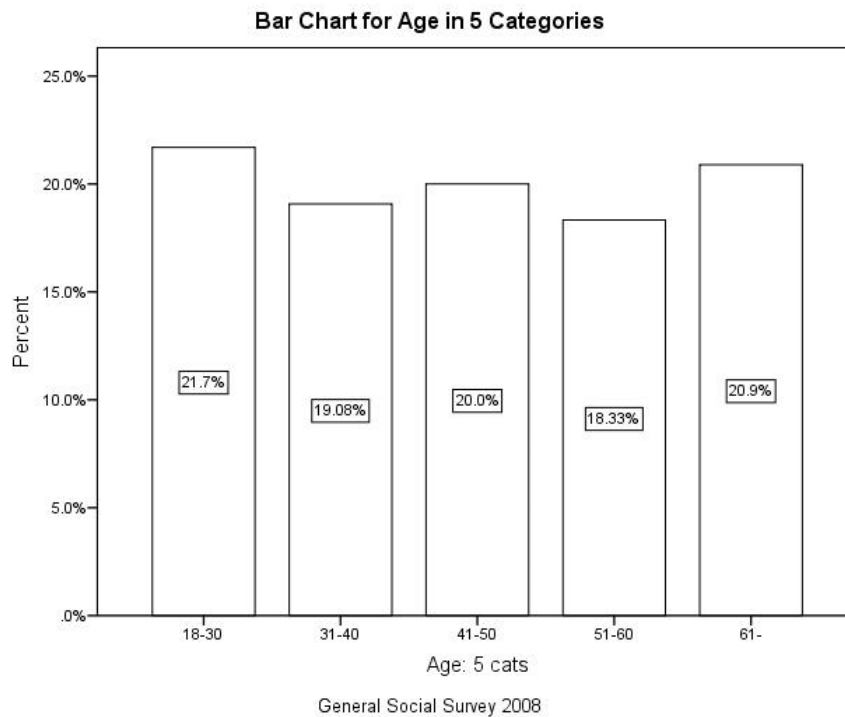
a. General Social Survey 2008

- III. Table of summary statistics:

Summary Statistics

N	2015
Median	41-50
Mode	18-30
Range	(61+) - (18-30)
Minimum	18-30
Maximum	61+
Q1	31-40
Q3	51-60
IQR	(51-60) - (31-40)
V-ratio	0.783

IV. Bar chart:



V. Descriptive paragraph requirements:

- What does your variable measure?
- Describe this variable in terms of all appropriate measurements of central location.
- Describe this variable in terms of the appropriate measurements of dispersion.
- If appropriate, is this distribution skewed negative or positive?
- Discuss any other interesting and relevant details about this distribution.

SPSS 1 Assignment Instructions

SPSS1 Frequently Asked Questions and Point Breakdown

5 questions – total of 20 points possible for each variable (100 points altogether):

1. Level of measurement (2/20 points):

- a. You must have the correct level of measurement
- b. You should not put interval/ratio. You must clearly identify if the data is interval or ratio for full points.
- c. You need to report the level of measurement on the original data, not the recoded data.

2. Cumulative Frequency Table (4/20 points):

- a. You must have a title, a source and appropriate columns (i.e. you should have the valid percent column, not the percent column)
- b. Your table must be formatted according to APA guidelines. Week 6 under course content on our Blackboard Course website has both the template and an instructional video if you would like to refresh your memory from lab.
- c. If you have more than 10 rows you should recode the data. Make sure that the valid total of your recoded cumulative frequency table matches the valid total of the cumulative frequency table on the original data.

3. Descriptive Statistics (5/20 points):

- a. Do not directly copy and paste from SPSS. You must format the tables according to the APA Guidelines.
- b. Measures of central tendency must be correct for the level of measurement of the data. (For example, if you inappropriately report the mean on categorical data you will lose points)
- c. Measures of dispersion must be appropriate for the level of measurement of the data. (For example, if you inappropriately report variance on categorical data you will lose points.)
- d. Always report the category labels for categorical data.
- e. You will need to perform some calculations. Remember, SPSS does not give you the IQR or the V ratio. You will need to calculate these out correctly for full points.
- f. Remember, if you have ordinal data the numbers associated with the data are just the coding scheme. Therefore, you need to simplify the IQR and the Range as much as possible. (For example, if you were using a Likert Scale, you could end up with Hate-Love for the range and Somewhat dislike-Somewhat love for the IQR). Remember for categorical data you must work with the category labels.
- g. You must run statistics on the original data, not the recoded data. As such, your descriptive statistics should be appropriate to the level of measurement of your original data.

4. Graphics (4/20 points):

- a. These can be copied and pasted from SPSS. You need to make sure that your graphs have titles and sources. You may use Excel but we strongly encourage you to use SPSS
- b. Your graph must be appropriate to the level of measurement of your data.
 - i. (nominal = pie chart; ordinal = bar graph; interval/ratio = histogram)

SPSS 1 Assignment Instructions

- ii. If using recoded variable, should use the level of measurement AFTER the recoding (interval recoded to ordinal should use bar graph)
- iii. However, if you use a histogram with the interval/ratio original data, you will not be penalized
- c. Have value labels. Remember to use the crosshairs in the Graph Editor on SPSS.
- d. Bar chart should use % of cases (“Percent”) not count.

5. Paragraph (5/20 points):

- a. You need to describe both central tendency and dispersion. Do not just laundry list all of the statistics. Focus on the most salient univariate statistics for your level of measurements (i.e. if you have ratio data, you should discuss the mean for central tendency and the standard deviation and the variance for measures of dispersion). You need to interpret theses as well. What do they tell you about the variable being measured? (For example, if you have a variable on age and the average age was 25, what could you infer about the population?). Report interesting trends from the statistics.
- b. Discuss the shape of the distribution curve (positive or negative skew) if it is appropriate for the level of measurement. (Hint: one good way to do this is to look at the difference of the median and the mean.)
- c. Make sure that you report outliers if it is appropriate to the level of measurement for your data. (Hint: 99% of the data is + or – 2 standard deviations from the mean)