

## Stat 95 SPSS Assignments General Information

**Learning Objectives:** The purposes of the SPSS assignments are: (1) To introduce you to SPSS as a tool to analyze data; (2) To teach you to create and interpret tables and graphs for these data; (3) To make decisions about data; and (4) To teach you to write clearly, correctly, and concisely about data and the results of statistical analyses performed on these data.

**General Instructions:** For all assignment, enter the data into SPSS. Then follow the instructions below to obtain the statistics and graphs required for each assignment. Use this statistical and graphical information to answer the questions for each assignment. When answering all questions, use complete sentences and proper grammar. If the question asks for a numerical value, don't just write the number. Write a sentence describing the value. For example, if asked for the mean, you should write something like this: "The mean for the Group 1 was 54.00." Be sure to include the number/letter of each question when answering. Type (and double-space) your answers on a separate sheet and staple this sheet to your SPSS printouts. Be sure to include your name, instructor, section number, and date in the right-hand corner of the answer sheet. Place the answer sheet on top of the SPSS printouts. Do not include the assignment sheets (those with the data). Please note that individual instructors may schedule the assignments in different orders. Therefore, the order of the instructions below may not correspond to the order of assignments given by your instructor. Pay close attention to the assignment name. If you have questions, consult with your instructor. Unless told otherwise by your instructor, use an alpha level of .05 for all hypothesis tests.

**Analysis Instructions (apply to version 16.0; other versions may require different commands)****Assignment: Frequency Distributions and Descriptive Statistics**

1. Under **Variable View**, name the two variables, “Section1” and “Section2.” (Please note that you cannot use spaces in your variables names).
2. Go to **Data View** and enter the data into SPSS
3. To obtain descriptive statistics and frequency distribution tables and histograms:
  - a. Select **Analyze**→**Descriptive Statistics**→ **Frequencies**
  - b. Choose the two variables and move them to the **Variable(s)** box on the right side, using the → button located between the two variable boxes.
  - c. Check **Display frequency tables**
  - d. Press the **Statistics** button. Under **Central Tendency**, select **Mean**, **Median**, and **Mode**. Under **Dispersion**, select **Std. deviation**, **Range**, **Minimum**, and **Maximum**. Then select **Continue**.
  - e. Select **Charts** and choose **Histograms**: Check **With normal curve**. Check **Frequencies** under **Chart Values**. Then select **Continue**
  - f. Finally, select **OK**
4. Your results will appear in a separate screen as three tables (one Statistics table and two Frequency Tables) and two frequency distribution Histograms.

**Assignment: Correlation and Scatterplots**

1. Under **Variable View**, name the three variables, “Drinks,” “GPA,” and “Classes”
2. Go to **Data View** and enter the data into SPSS
3. To obtain descriptive statistics and frequency distribution see the first set of instructions above.
4. To obtain Pearson’s correlation coefficient:
  - a. Select **Analyze**→**Correlate**→ **Bivariate**
  - b. Choose all three variables and move them to the **Variables** box on the right side, using the → button located between the two variable boxes. The order does not matter.
  - c. Under **Correlation Coefficients**, Check **Pearson**
  - d. Under Test of Significance, Check **Two-tailed** or **One-tailed**, depending on the situation described in the assignment.
  - e. Check **Flag significant correlations**

- f. Select **OK**
  - g. Press the **Statistics** button. Under **Central Tendency**, select **Mean**, **Median**, and **Mode**. Under **Dispersion**, select **Std. deviation**, **Range**, **Minimum**, and **Maximum**. Then select **Continue**.
  - h. Select **Charts** and choose **Histograms**: Check **With normal curve**. Check **Frequencies** under **Chart Values**. Then select **Continue**
  - i. Finally, select **OK**.
5. Your results will appear in a separate screen as a single table of correlation coefficients, significance values, and *N* (each value will be presented twice, once above and once below the diagonal).
6. To obtain scatterplots (There are several ways to get scatterplots. This is one just way.)
- a. Select **Graphs**→**Chart Builder**
  - b. Select **Scatter/Dot** from the **Choose From** box in the lower left-hand corner.
  - c. Select the first scatterplot and drag it to the large box in the upper-middle.
  - d. Select and drag the appropriate variable to the X axis and to the Y axis.
  - e. Select **OK**
  - f. You will see a scatterplot of the data.
7. To add a best-fit line to the scatterplot:
- a. Double-click on the scatterplot. This will open an editing window.
  - b. In second row of the editing window, select the **Add Fit Line at Total** button. It looks like this: 
  - c. Make sure that **Linear** is selected under **Fit Method**.
  - d. Select **Apply** and close the editing window. You should see a straight line running through the middle of the data. SPSS will also provide the coefficient of determination ( $r^2$  or *R sq Linear*) near the best-fit line.

**Assignment: Comparing Two Means using the Independent-samples  $t$  test**

1. Under **Variable View**, name the two variables, “Errors” and “Group”
2. For group, click on **Values**. Then click on the **Value Labels** button. It looks like this: 
3. Under **Value Labels**, enter “1” in the **Value** box. Enter “Cell-phone” in the **Label** box. Then select the **Add** button. You will see this in the center box: 1.00 = “Cell-phone”. This command tells SPSS that when it sees the number 1 in the “Group” column, the data point right next to it in the “Errors” column comes from someone in the Cell-phone group.
4. Under **Value Labels**, enter “2” in the **Value** box. Enter “Radio” in the **Label** box. Then select the **Add** button. You will see this in the center box: 2.00 = “Radio”. This command tells SPSS that when it sees the number 2 in the “Group” column, the data point right next to it in the “Errors” column comes from someone in the Talk-radio group.
5. Select **OK**
6. Go to **Data View** and enter the errors for the Cell-phone group in the “Errors” column. Then type in the number 1 in the first 20 spaces in the “Group” column. This tells SPSS that the first 20 scores are from people in the Cell-phone group.
7. Starting in space 21 in the “Errors” column, enter the errors for the Radio group (right below the first 20 data points). Make sure that you have 40 total spaces filled in the “Errors” column. Then type in the number 2 in the spaces 21-40 in the “Group” column. This tells SPSS that the next 20 scores are from people in the Radio group.
8. To compute the independent-samples  $t$  test:
  - e. Select **Analyze** → **Compare means** → **Independent-samples T Test**
  - f. Place the dependent variable in the **Test Variables** box on the right side, using the → button located between the two variable boxes.
  - g. Place the independent variable in the **Grouping Variable** box on the right side, using the → button located between the two variable boxes.
  - h. Click the **Define Groups** button and enter “1” in the **Group 1** box and “2” in the **Group 2** box.
  - i. Select **Continue**
  - j. Select **OK**

- k. Your results will appear as two tables (one with descriptive statistics and the other with the results of the independent-samples  $t$  test).
  - l. In the table of descriptive statistics, pay attention to the mean and standard deviation.
  - m. In the independent-samples  $t$  test table, only pay attention to the first row, labeled “Equal variances assumed” and ignore the data under the area labeled “Levene’s Test for Equality of Variances.”
8. To make a bar graph with error bars (There are several ways to get bar graphs. This is one just way.)
- a. Select **Graphs**→**Legacy Dialogs**→**Interactive**→**Bar**
  - b. Select and drag the appropriate variable to the X axis and to the Y axis
  - c. Be sure that under **Bars Represent**, **Means** is chosen (lower left-hand corner)
  - d. Click on the **Error Bars** tab at the top of the dialog box.
  - e. Select **Display Error Bars**
  - f. Under **Units**, select **Standard Error of Mean**. Make sure that the value is set to **1.0**. You can choose any **Shape** for the error bars.
  - g. Select **OK**
  - h. You will see a bar graph of the data.
9. To compute  $r^2$  (which measures the strength of the association between the independent and dependent variables) use this formula (the  $t$  and  $df$  are obtained from the independent-samples  $t$  test)

$$r^2 = \frac{t^2}{t^2 + df}$$

10. To compute Cohen’s  $d$  (which measures the distance between the two group means in standard deviation units – much like a  $z$  score)
- a. Take the average of the standard deviations for both groups:  $(SD_1 + SD_2)/2$ ; call this the “pooled SD” or  $SD_{pooled}$
  - b. Divide the Mean Difference obtained from the independent-samples  $t$  test by the pooled SD.

$$d = \frac{\bar{X}_1 - \bar{X}_2}{SD_{pooled}} \quad \text{where} \quad SD_{pooled} = \frac{SD_1 + SD_2}{2}$$

c. Here are the formulas:

d. Note that this formula for the pooled SD only works when the groups have equal  $n$

11. A note on significance values (often called  $p$ -values):

a. By default, SPSS only reports  $p$ -values to the third decimal place. This means that for  $p$ -values lower than .0009, SPSS will give you a value of .000. Please note that it is impossible to have a  $p$ -value equal to 0. There is always going to be a non-zero number somewhere in the  $p$ -value, even if SPSS doesn't show it. You have two options in this case (*\*Ask your instructor which option he or she would prefer*):

i. Get the exact  $p$ -value

1. Double-click on the  $p$ -value itself. This will open an editing window.

2. Double-click on the  $p$ -value again. The exact  $p$ -value will appear.

ii. Provide an approximation of the actual  $p$ -value

1. Report the  $p$ -value as  $p < .001$ ,  $p < .01$ , or  $p < .05$

b. For the independent-samples  $t$  test, SPSS will only compute a two-tailed (or *non-directional*) test and will report the associated two-tailed  $p$ -value (under the *Sig.* column). To obtain the one-tailed (*directional*)  $p$ -value simply divide the given  $p$ -value by 2. Report the resulting value.

### Assignment: Comparing Three Means using One-Way ANOVA and Post-hoc Tests

1. Under **Variable View**, name the two variables, "Pounds" and "Group."
2. To define group membership, we will use a similar procedure to what we used for the independent-samples  $t$  test: we will use numbers to indicate group membership. The only difference is that we now have three groups, instead of two. Enter 1 for the Diet group, 2 for the Exercise group, and 3 for the Diet+Exercise group.
3. Select **OK**
4. Go to **Data View** and enter the data in the same way that you did for the independent-samples  $t$  test, except that you will have three groups, instead of two.
5. To compute the one-way ANOVA:
  - c. Select **Analyze** → **Compare means** → **One-Way ANOVA**

- d. Place the dependent variable in the **Dependent List** box on the right side, using the → button located between the two variable boxes.
  - e. Place the independent variable in the **Factor** box on the right side, using the → button located between the two variable boxes. (Note: *Factor* means independent variable)
  - f. Select **Post Hoc . . .** and choose the appropriate post-hoc test (ask your instructor what post-hoc test(s) he or she wants you to compute). A reasonable choice is **Tukey**. After you made your selection, click **Continue**
  - g. Select **Options**. Under **Statistics**, select **Descriptive** then select **Continue**.
  - h. Select **OK**.
  - i. Your results will appear as three tables (one with descriptive statistics, one with the results of the one-way ANOVA, and one with the results of the post-hoc tests).
12. To make a bar graph with error bars, see the instructions above under the independent-samples *t* test.
13. To compute  $\eta^2$  (which measures the strength of the association between the independent and dependent variables), simply divide the *SS* for your factor (*SS* between groups) by *SS Total*.
14. A note on the post-hoc test results: SPSS gives you each post-hoc test twice and switches the order of the groups. Only pay attention to the first, second, and fourth lines of the post-hoc test results. The other lines give you the same results but with the groups reversed (so only the sign on the mean difference will change).

### **Assignment: Comparing Proportions using the Two-Way Chi-Square Test of Independence**

1. Under **Variable View**, name the two variables, “Parent” and “Diagnosis.”
2. To define group membership, use the **Value Labels** procedure like you did for the *t* test.
  - a. For the “Parent” variable, use 1 for “Yes” and 0 for “No”
  - b. For the “Diagnosis” variable, use 1 for “Yes” and 0 for “No”
3. Go to **Data View**
4. Enter the data.
  - a. Under “Parent\_AD” we want to indicate whether or not the person had a parent diagnosed with Alzheimer’s Disease (AD). To do so, use 1 to indicate “Yes” and 0 to indicate “No.” For people in the first column of the data sheet, enter 1. For people in the second column enter 0.

- b. Under “Participant\_AD” we want to indicate whether or not the person has received a diagnosis of AD. To do so, use 1 to indicate “Yes” and 0 to indicate “No.”
4. To compute the two-way chi-square analysis:
  - a. Select Analyze→Descriptive Statistics→Crosstabs
  - b. Select “Parents” and move it into the box labeled **Row(s)** using the → button.
  - c. Select “Diagnosis” and move it into the box labeled **Column(s)** using the → button.
  - d. Select **Display clustered bar charts** in the bottom left-hand corner.
  - e. Select **Statistics**.
    - i. Select **Chi-square** (the test statistic)
    - ii. Select **Phi and Cramer’s V** (the strength-of-association measure)
    - iii. Select **Continue**
  - f. Select **Cells**
    - i. Under **Counts**, select **Observed**
    - ii. Under **Percentages**, select **Row, Column, and Total**
    - iii. Select **Continue**
    - iv. Select **OK**
  - g. The results will appear as three tables (one for counts and percentages, one for the chi-square test, and one for the Phi and Cramer’s V) and one bar graph.
  - h. Under the table labeled **Chi-Square Tests**, only pay attention to the line labeled **Pearson Chi-Square**.
  - i. To obtain  $r^2$ , simply square the value of Phi (not its significance value!). The result has the same interpretation as any other coefficient of determination.

**A note on writing statistical results in APA style:**

1. Be clear. Use straightforward language. Don’t try to sound “scientific.” Just state what happened. Clearly identify all values that you provide. Don’t make the reader guess what you’re trying to say.
2. Be concise. Don’t use unnecessary words. Don’t be too “wordy.” Get to the point.
3. Be correct. Report values accurately (round to the nearest two decimals). Always report two decimal places, even if they are both zeroes. Accurately report the descriptive and inferential results. Report all required values.

## 4. Other stylistic remarks:

- a. Italicize all symbols.
- b. Place spaces around equal signs. Like this: “ $M = 1.00$ ” **not** like this: “ $M=1.00$ ”
- c. Use  $M$  and  $SD$  for the mean and standard deviation, respectively. Always include both when reporting means.
- d. You don’t need a zero in front of the decimal point for quantities that can’t go above 1.00 (e.g.,  $p$ -values, Pearson’s  $r$ ).
- e. When reporting the results of a hypothesis test
  - i. Identify the test being used (e.g., independent-samples  $t$  test, one-way ANOVA)
  - ii. Report the following values: the obtained test statistic,  $df$ , *Sig.* ( $p$ -) value, direction of test (if appropriate), and effect size and/or strength-of-association measures.
  - iii. Independent-samples  $t$  test example (where the obtained  $t = 4.20$ ;  $df = 23$ ;  $p = .0003$ , and it’s a two-tailed test, with Cohen’s  $d = 1.20$  and  $r^2 = 0.43$ )  
 $t(23) = 4.20, p = .0003, two-tailed, d = 1.20, r^2 = 0.43$   
Alternatively, you could write an approximate  $p$ -value:  
 $t(23) = 4.20, p < .001, two-tailed, d = 1.20, r^2 = 0.43$   
*\*Ask your instructor which format he or she prefers – exact or approximate  $p$ -values*