SPSS for Windows (Version 20)

Dr. Jann-Huei Jinn

1. Accessing SPSS and R:Drive:
　1) Double click the **Application** icon.
　2) In the **Application** Window double click on "**Statistics**" folder.
　3) Double click on SPSS 20.
　4) Click on OK (default is "Open an existing data source"). In the"Open data" dialog
　　 Window, click on the downward arrow next to "Look in"
　5) Choose "GVSU-LABDATA(\\LAB.ADS.GVSU.EDU\DFS) (R:) This is the
　　 R:Drive.
　　　 Note: This is the way to get to the files stored in folders on the R:Drive.

2. How to download the files of the Agresti and Franklin's text:
　1) Double click on the R:drive, then select "**STAT**" folder.
　2) Double click on **Agresti and Franklin Materials**
　3) Double click on the fold "Data Sets for Agresti and Franklin 2nd Edition" then you
　　 will see all the data files are arranged alphabetically.
　4) Double click on the data file which you want to open

3. How to key in the data:
　1) In the "IBM SPSS Statistics 20"dialog window select "Type in data" then click **OK**.
　2) In the IBM SPSS Statistics Data Editor window click **Variable View** then type a
　　 variable's name in the first row and hit <Enter> key, continue to do so until you
　　 typed all the variables. Click **Data View** and you will see all the variables appear on
　　 top of each column.
　3) To enter a value, use the cusor keys or the moose to highlight a cell, and then type a
　　 value. <Enter> accepts the value and moves down one row. <Tab> accepts the value
　　 and moves right one column.
　4) For qualitative (categorical, attribute) variables, in the IBM SPSS Statistics Data
　　 Editor window click **Variable View** then type a variable's name and hit <Enter> key.
　　 Then click the "Numeric" cell (Under the **"Type"** column) then click the small square
　　 box (with 3 dots) and you will see a "Variable Type" dialog box. In the Variable Type
　　 dialog box, click **String** then click **OK**.
　5) In the "**Width**" column, you can change the # of spaces for the categories.
　6) If you type numerical values as symbols for categorical variables then click **"Values"**
　　 column to label them.
　7) To delete a case or variable, click the row or column heading and *Clear* it, by using
　　 the Edit menu or pressing <Delete>.
　8) To replace a value, select the cell and enter the new value. The new value replaces the
　　 old value in the cell editor.

4. How to exit SPSS: 1) Click **File** from menu bar
                         2) Click **Exit**

5. How to select Cases Based on Conditional Expressions: Click **Data** in the menu bar
    then select **Cases**.
     1) Select <u>If Condition is Satisfied</u> then click "If" button.
     2) Enter the conditional expression. String constants must be enclosed in quotation
        marks or apostrophes. Numeric constants must be typed in American format, with
        the period (.) as the decimal indicator.

6. How to open a file from a disk or flash drive:
     1) Click **File** from menu bar then click **Open**
     2) In the "Open File" window select an appropriate drive
     3) Click on the name of the file desired then click **Open**

7. How to save a file on a disk or flash drive:
     1) Click **File** from menu bar
     2) Click **Save as**
     3) Select an appropriate drive you want to save your file(s)
     4) Type in filename under "File Name" with appropriate extensions. To save data, type
        <u>filename</u>.**sav** To save chart, type <u>filename</u>.**spo**
     5) Click **OK**

8. How to print SPSS output: 1) Click **File** from menu bar
                         2) Click **Print**
                         3) Click **OK**

        Note: Printing can be done from any window. You can change the printing
              defaults via **File Print**.

9. Procedures to create some of the charts or graphs:
    1) Click **Graphs** then **Legacy Dialogs** then choose any options of interest by clicking
       on the desired graph (e.g. **Bar, Line, Pie Chart, Box Plot, Scatterplot, Histogram,**
       etc.)
    2) When create a **Pie chart** or **Bar chart**, in the "**Data in Chart"** box select "Values of
                     Individual cases." Click **Define**
    3) To create a **Pie** chart , click **Variable** in the "Slice Labels" box then click the
       variable's name and click ► to move it into the "variable" box. Click the variable
       which contains the counts in the data set and click ► to move it into the "Slices
       Represent" box. Click **OK**.

4) To get the percentages in the Pie chart, double click the graph in the "Output- PASW viewer" window. In the <u>Chart Editor</u> dialogbox, right click the mouse and you will get a "Properties Window". Select <u>show data labels</u> then the percentages will be displayed.
You can move "count" into the "Disaplyed" box then click "Apply" button and close all the dialog boxes. Click the pie chart to get rid of the frame.
5) When create a single **boxplot** in the "**Data in Chart are"** box select " Summaries of Separate Variables". When create side-by- side **boxplot,** you have to define a grouping variable. In the "**Data in Chart are"** box select "Summaries for Group of Cases." Click the grouping variable into the **category axis** box
6) Click on **OK**
Note: More details will be demonstrated in the class.

10. Procedures to perform some descriptive statistics:

1) Cilck **Analyze** then **Descriptive Statistics** then **Frequencies**.
2) Click on the desired variable(s) then click ► to move it into the "Variable(s)" box.
3) Cilck **Statistics** then click on the desired statistic(s), e.g. quartiles, percentiles, mean, median, mode, sum, variance, standard deviation, minimum, maximum, range, etc.
4) When click "Percentiles", you need to key in a number, e.g., (25 means 25th percentile,) then click "**Add**."
5) Click **Continue** then **OK.**

Note: If you click **Analyze** then **Descriptive Statistics** then **Descriptives**, you will be able to get many useful statistics except the quartiles and percentiles.

11. To obtain descriptive statistics and graphs for quantitative (dependent) variables broken down by a qualitative (independent) variable:

1) Use Analyze > Descriptive Statistics > Explore in the menu
2) Move the quantitative variable to the Dependent List box
3) Move the qualitative variable to the Factor List box
4) Click OK

12. Procedures to perform stem-and-leaf plot:

1) Click **Analyze** then **Descriptive Statistics** then **Explore.**
2) In the "Explore" dialog box click on the desired variable's name in the left box then click on the **right arrow** to move the variable to the **Dependent List** box. Click the categorical variable and move to the **Factor list** box.

3) Click **Statistics** then in the "Explore: Statistics" dialog box
check"Percentiles" then click continue and click OK

Note: You may get a stemplot and a histogram at the same output in this procedure.

13. To make a clustered bar chart:

1) Click **Graphs** then **Legacy Dialogs** then select **Bar** graphs
2) Select the **Clustered** icon and **Summaries for groups of cases** in the Bar Charts dialog box and click Define.
3) Select **N of Cases**, move appropriate categorical variables into the **Category Axis** box and **Define Clusters by** box , respectively then click OK.

Note: To obtain the clustered bar chart in Figure 6.1 (page 195 of textbook), you have to define three variables: Smoke, Divorce, and Count.The variable **Count** represents the number of observations in each cell.

   a) The data should be keyed in as follows:

Yes	Yes	238
Yes	No	247
No	Yes	374
No	No	810

   b) Click **Data** in the menu bar then click "Weight Cases.."
   c) In the "Weight Cases" dialog box select the variable with frequencies (in this example, it is Count) then click "Weight Cases by" and move that variable into the "Frequency Variable" box then click **OK**

Repeat steps 1 and 2 then

4) Select **N of Cases**, move *Smoke* into the Category Axis box, and move *Divorce* into Define Clusters by box then click OK.

14. To make a Doxplot:

1) Click on *Graphs* then Legacy Dialogs then click *Scatter/Dot*
2) Click on **Simple Dot** and then click on *Define*.
3) Highlight the variable of interest and click ► next to *X-Axis Variable*.
4) Click *OK*.

15. How to generate a random sample from any distributions:

1) Create an existing data with the number of observations equal to the size of the random sample you wish to generate. Note: The existing data file can have any numbers as the observations, for example, all "0"s.
2) Click **Transform** then click **Compute Variable**.

3) In the "Compute Variable" dialog box type the new variable's name in the "Target Variable" box.

4) In the **Function**s **group** box double click *Random Numbers* and you will see a list of special random variables in the "Functions and Special

5) Variables" box. Choose the distribution you want (e.g. RV. NORMAL) then double click it and you will see RV.Normal(?,?) in the "**Numeric Expression**" box.

6) Type in the appropriate number(s) for the parameter(s) (to replace the "?").

7) Click **OK** and you will see a random sample created under the new variable's column.

16. To select a random sample of n cases from N cases:

1) Click **Data** then click **Select Cases**.

2) Click **Random Sample of Cases** and then click the **Sample** button.

3) In the " Select Cases: Random Sample" dialog box, click **Exactly n Cases from the first N cases**.

4) Click **Continue** then click **OK**.

Note: If you are finished with analyses using only the random sample, you should " **unselect**" cases by clicking **Data**, then **Select cases**, and then **All Cases**. Then click **OK**. All cases will be included in subsequent analyses.

17. Procedures to perform Simple Linear Regression:

1) To construct a scatter plot for two quantative variables, click **Graphs** then **Legacy Dialogs** then **Scatter/Dot**

2) In the "Scatter/Dot " dialog box click on **Simple Scatter** so that the scatterplot icon is outlined. Click on **Define.**

3) In the "Simple Scatterplot" dialog box select the appropriate variables from the list of variables for the **X-Axis and Y-Axis** then click **OK.**

4) To plot the regression line (or confidence and prediction intervals) on the scatter plot, double click the scatter plot in the "Output-SPSS Viewer" window to get the "Chart Editor" window.

5) In the "Chart Editor" window click **a point** in the scatterplot then click the "Add fit line" tool then close the "Chart Editor".

6) To obtain the least squares estimates for the regression coefficients, click **Analyze** then **Regression** then **Linear**.

7) In the "Linear Regression" dialog box, select the appropriate quantitative variable for the **Dependent** box and **Independent** box, respectively.

8) Click **OK**.

Note: The number under the heading **Beta** is your correlation coefficient, r.

18. To make a residual plot:

       1) Click on **Analyze** then **Regression** then **Linear**.
       2) In the "Linear Regression" dialog box, click on **Save** to get the
          "Linear Regression: Save" dialog box.
       3) In the "Predicted Values" box click on **Unstandardized**.
       4) In the "Residuals" box click on **Unstandardized** then click **Continue**
          then click **OK**.
       5) Click on **Graph** then **Scatter** and follow steps 1,2,3, and 4 in the
          procedures of performing simple linear regression to get the residual
          plot. More details will be discussed in the class.

   Note: For each option you selected, one or more variables will be added to the
         working data file.

19. To perform one-sample t test:

       1) Click on **Analyze** then **Compare Means** then **One-Sample T Tests.**
       2) Select the desired variable for the **Test Variable(s)** box.
       3) To test a hypothetical mean, input the value of the mean in the null
          hypothesis to the **Test Value** box.
       4) To construct a confidence interval for the population mean, input the
          value 0 to the **Test Value** box and change the level of confidence by
          clicking the **Options**. (Note: The default confidence level is 95%)
       5) Click on **Continue** then **OK.**

20. To perform paired-sample t test:

       1) Click on **Analyze** then **Compare Means** then **Paired-Samples T Tests**.
       2) Select the two desired variables for the **Paired Variables** box.
       3) To change the confidence level, click on the **Options** and input the
          desired level of confidence then click on **Continue**.
       4) Click on **OK**.
   Note: SPSS always use the first variable to minus the second variable.

21. To perform two-sample t test:
       1) Create a variable contains **all** the observations from the two
          independent samples. Create a **Grouping Variable** to correspond to the
          two independent samples.

         Note: If there is an existing categorical variable to indicate the two
            independent groups then you don't need to do this step.
       2) Click on **Analyze** then **Compare Means** then **Independent-Samples T
         Tests**.

3) Select the variable contains **all** the observations for the **Test Variable** box.
4) Select the desired variable for the **Grouping Variable** box, then click on **Define Groups**. Input the two numerical values you defined for the two groups.
5) Click on **Options** to change the desired confidence level, then click on **Continue** and click OK.

22. To perform Chi-Square Test:

Note: It is important to learn how to enter data from a contingency table into SPSS to perform the Chi-square test. The following is an example:

Olive Oil

	Low	Midium	High	Total
Colon cancer	398	397	430	1225
Rectal cancer	250	241	237	728
Controls	1368	1377	1409	4154

   a) Define three variables: Cancer, Olive oil, and Count. The variable **Count** represents the number of observations in each cell.
   b) The data should be keyed in as follows:

Colon	low	398
Colon	Medium	397
Colon	High	430
Rectal	low	250
Rectal	Medium	241
Rectal	High	237
Controls	low	1368
Controls	Medium	1377
Controls	High	1409

c) Click **Data** in the menu bar then click "Weight Cases.."
d) In the "Weight Cases" dialog box select the variable with frequencies (in this example, it is Count) then click "Weight Cases by" and move that variable into the "Frequency Variable" box then click **OK**

After you did a), b), c), d) then do the following:

1) To create a contingency table for counts or percentages for two qualitative variables, click on **Analyze** then **Descriptive Statistics** then **Crosstabs**.
2) Select the desired variable(s) for the **Rows** and **Colums** then click on **Cells**.

3) To show the desired counts in each cell, select Observed and Expected in the **Count** option box.
4) To show the desired percentages in each cell, select Row, Column, and Total in the **Percentages** option box then click **Continue**.
5) To obtain the Chi-Square statistic value, click on the **Statistics** option box then select "Chi-Square" then click on **Continue**.
6) Cilck **OK**

23. To perform One-Way ANOVA:

1) Create a variable contains all the observations from all the independent groups. Create another Factor variable and use numerical values to correspond to those independent groups.

Note: If there is an existing categorical variable to indicate different independent groups then you don't need to do this step.

2) Click on **Analyze** then **Compare Means** then **One-Way ANOVA**.
3) Select the variable contains all the observations for the **Dependent List** box.
4) Select the desired variable for the **Factor** box.
5) To perform multiple comparisons one can click on **Post Hoc**, choose the desired multiple comparison test procedure, and click on **Continue**.
6) Click on **OK**.