Fletcher Jones
Student Peer to Peer
Technology Training Program

Basic Statistics using SPSS
Intro to SPSS

Intro to SPSS...................................................................................................................................................... 2
Opening SPSS.......................................................................................................................................................... 3
Types of SPSS files and editors ............................................................................................................................ 4
Creating a New Data File (File>New>Data)........................................................................................................... 4
Opening an Existing Data File (File>Open>Data).................................................................................................... 5
  1 Opening an existing SPSS file.......................................................................................................................... 5
  2 Opening a non-SPSS (*.sav) file type............................................................................................................... 6
Reading a Text File (File>Read Text Data).......................................................................................................... 8
Data Editor ............................................................................................................................................................. 9
  Data View - contains data ................................................................................................................................. 10
  Variable View – information about the data...................................................................................................... 13
Defining Variables ................................................................................................................................................ 14
  Defining Value Labels and Values ................................................................................................................... 15
Entering Data ....................................................................................................................................................... 17
Analyzing Data ..................................................................................................................................................... 18
  Descriptive Statistics ....................................................................................................................................... 18
    Frequencies ................................................................................................................................................... 18
    Cross Tabs ..................................................................................................................................................... 19
  Reports ............................................................................................................................................................... 19
    Case Summaries ............................................................................................................................................ 19
Statistics Coach .................................................................................................................................................. 19
Viewer window and Output files (*.spo) ................................................................................................................ 19
  Charts and Graphs ............................................................................................................................................ 20
Saving Output ...................................................................................................................................................... 20
Export Output ...................................................................................................................................................... 21
SPSS for Windows is a tool used for managing data and conducting statistical analyses. The analyses range from descriptive statistics, such as summaries and frequencies, to T-tests, regression models and life tables.

**Opening SPSS**

You can open SPSS for Windows from the Start menu.

1. **Start>Programs>Courseware>SPSS for Windows>SPSS 11.5 for Windows**

The session will either open a dialog box with options or it will default to an empty data sheet in the Data Editor window.

- If your session defaults to the dialog box with options, then choose the **Type in Data** radio button so that a blank Data Editor Window opens.

The Data Editor is the window that is opened at start-up and is used to enter and store data in a spreadsheet format.
Types of SPSS files and editors

SPSS utilizes a graphical environment with descriptive menus and simple dialog boxes, similar to the Microsoft products with which you are most likely familiar.

1. **Browse File >New and File >Open**

There are 5 different windows available to create or open, and each of these is associated with a particular type of SPSS file.

- **Data** Opens the Data Editor with a blank data sheet
- **Syntax** Opens the Syntax Editor window for pasting dialog box choices into a syntax window, where your selections appear in the form of command syntax. You can then edit the command syntax to use special features of SPSS not available through dialog boxes.
- **Output** Opens the Viewer window where all statistical results from analyses are displayed
- **Script** Opens the Script Editor for writing and modifying basic scripts
- **Other**

For this workshop, we will focus on the Data and Output files.

Creating a New Data File (File>New>Data)

Two quick ways to open a new data file:

1. Opening a new session of SPSS automatically opens a new blank datasheet in the Data Editor window.

2. Or explicitly choose **File>New>Data**.
Opening an Existing Data File (File>Open>Data)

1. Opening an existing SPSS file
   a. Choose File>Open>Data
      By default, the SPSS-format data files (.sav) are displayed.
   b. Navigate to the demo.sav file and open it.
2 Opening a non-SPSS (*.sav) file type

a Choose File > Open > Data

b In the Open File dialog box, navigate to the Files of Type drop-down list to display other file types.

For example, to open an EXCEL file:

1 File > Open > Data

2 Navigate to the Files of Type drop-down list, change the file type to Excel (*.xls), locate the file, and choose Open.
3 The **Opening Excel Data Source** dialog box opens.

![Opening Excel Data Source dialog box]

- You are given the option to specify that the variable names be taken from the first row of data in the spreadsheet.
- You can also choose to import a specific range of cells.

4 To read the first row of data in your spreadsheet into the column headings (or variable names) in the SPSS data sheet, check the box labeled “**Read variable names from the first row of data**” and then click **OK**.

5 You can choose to put in smaller range of cells: A1:AB100.

6 The data then appears in the Data Editor, and the column headings, which will be used as variable names, match the first row of spreadsheet data.

![Untitled - SPSS Data Editor]

- AGE
- MARITAL
- ADDRESS
- INCOME
- INCCAT
- CAR
- C
Reading a Text File (File>Read Text Data)

You can read text data by choosing File>Open and changing the file type in the Files of Type drop-down list to *.txt.

However, the File> Read Text Data command automatically defaults to files of Text type, so it’s more convenient.

1. Choose File> Read Text Data,
2. The Open Dialog Box opens. Locate the demo.txt text file, and open it.
3 The **Text Import Wizard** opens. This is similar to the Text Import Wizard in Excel.

4 Follow the 6 steps in the Wizard.

   **Note:** In Step 2, choose Yes to the radio button for the question “Are variable names included at the top of your file?” if you want to read the first row of data as the variable names in your SPSS data file.

**General File Notes:**

- Only one data file can be open in the Data Editor at a time.
- When a new data file is opened, the current data file is automatically closed.
- If you want to have multiple data files open at the same time, you have to start multiple sessions.

**Data Editor**

Once a file is opened by any of the above means, the data is displayed in the **Data Editor** window. This window allows you to create and edit the data files.
1. Choose File>Open>Data

2. Navigate to the demo.sav file and open it.

In the Data Editor, there are two views to choose from within the active data file, the Data View and the Variable View. They are listed in the tabs on the bottom of the data sheets.

**Data View - contains data**

- rows represent cases (observations)
- columns represent variables

For example, in the below demo file, a row represents a case (a respondent’s answer to a survey):
The columns (variables) represent the questions (or data points) asked in the survey:

<table>
<thead>
<tr>
<th></th>
<th>age</th>
<th>marital</th>
<th>address</th>
<th>income</th>
<th>inccat</th>
<th>car</th>
<th>carcat</th>
<th>ed</th>
<th>employ</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>55</td>
<td>Married</td>
<td>12</td>
<td>72.00</td>
<td>$60 - $74</td>
<td>36.20</td>
<td>Luxury</td>
<td>Did not co</td>
<td>23</td>
</tr>
<tr>
<td>2</td>
<td>56</td>
<td>Unmarried</td>
<td>29</td>
<td>153.00</td>
<td>$75+</td>
<td>76.90</td>
<td>Luxury</td>
<td>Did not co</td>
<td>35</td>
</tr>
<tr>
<td>3</td>
<td>28</td>
<td>Married</td>
<td>9</td>
<td>23.00</td>
<td>$25 - $49</td>
<td>13.70</td>
<td>Economy</td>
<td>Some coll</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>24</td>
<td>Married</td>
<td>4</td>
<td>26.00</td>
<td>$25 - $49</td>
<td>12.50</td>
<td>Economy</td>
<td>College de</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>25</td>
<td>Unmarried</td>
<td>2</td>
<td>23.00</td>
<td>Under $25</td>
<td>11.30</td>
<td>Economy</td>
<td>High sch</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>45</td>
<td>Married</td>
<td>9</td>
<td>76.00</td>
<td>$75+</td>
<td>37.20</td>
<td>Luxury</td>
<td>Some coll</td>
<td>13</td>
</tr>
<tr>
<td>7</td>
<td>42</td>
<td>Unmarried</td>
<td>19</td>
<td>40.00</td>
<td>$25 - $49</td>
<td>19.80</td>
<td>Standard</td>
<td>Some coll</td>
<td>10</td>
</tr>
<tr>
<td>8</td>
<td>35</td>
<td>Unmarried</td>
<td>15</td>
<td>57.00</td>
<td>$50 - $74</td>
<td>28.20</td>
<td>Standard</td>
<td>High sch</td>
<td>1</td>
</tr>
</tbody>
</table>
Variable Names vs. Variable Labels

- By default, the column headings contain the “short” name of the variable. To display the more descriptive variable label, place the cursor on the column heading/variable name.

Variable Data Values vs. Variable Value Labels

- By default, the actual data in the cells defaults to data values, not descriptive data labels. (1, 2 vs. married, unmarried). To switch between data values and descriptive data Value Labels, choose View>Value Labels. Note: When in the Value Labels mode, the data can be chosen from a drop down list.
The Variable view contains the attributes of each of the variables in the data file. In the Vehicle view:

- Rows are variables
- Columns are variable attributes

You can add or delete variables and modify attributes of variables, including:

- Variable name
- Data type
- Number of digits or characters
- Number of decimal places
- Descriptive variable and value labels
- User-defined missing values
- Column width
- Measurement level

You can switch between the two views by:

1. Choosing the tab at the bottom of the Data Editor window
2. Clicking on the column heading / variable name while in Data View OR Clicking on the row number while in Variable View
Choosing View>Variables (or View>Data) from the View Menu

Defining Variables

1. Definition of the attributes is done in Variable View.
2. Choose File>New>Data and choose the Variable View tab.
3. Begin by entering the name attributes of the variables.

   - Type the names in the Name column.
     - For this case, type `age`, `gender` and `income`.
   - Note: Only short names (no spaces) can be entered. The Label column is where you can add more descriptive information.

4. New variables are automatically given a numeric data type.
   - To change the variable type, choose the data cell in the Type column and choose the ellipse. For example, choose income.

4. The Variable Type dialog box will appear.
5 Select Dollar as the variable type for income, select the format ($#,###), and click ok to save your changes.

6 In addition to changing the data type, you can make changes to the other attributes (Width, Decimals, Label, Missing Values, etc.).

- For example, to change the decimals for age and gender, choose the age and gender data cells in the Decimals column and use the arrows to choose the decimal points you desire - 0 for this example.

![Untitled - SPSS Data Editor](image)

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Width</th>
<th>Decimals</th>
<th>Label</th>
<th>Values</th>
<th>Mi</th>
</tr>
</thead>
<tbody>
<tr>
<td>age</td>
<td>Numeric</td>
<td>8</td>
<td>0</td>
<td>None</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>gender</td>
<td>Numeric</td>
<td>8</td>
<td>0</td>
<td>None</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>income</td>
<td>Dollar</td>
<td>6</td>
<td>0</td>
<td>None</td>
<td>None</td>
<td></td>
</tr>
</tbody>
</table>

### Defining Value Labels and Values

Labels are the more descriptive variable labels associated with the Name attribute. They are most beneficial in charts and reports.

Values are the descriptive labels for values associated with the actual data value (1, 2 vs. married, unmarried). In other words, they provide a method for mapping the variable values to a string label.

1 While in Variable View, choose the cell in the Label column under the age row, and type in “Age of Respondent”
   - In the Label column under the gender row, type in “Gender of Respondent”
   - In the Label column under the income row, type in “Annual Income of Respondent”

2 Add Values to these Labels.
   - Click the Values column cell for the gender row and choose the ellipse

3 The Value Label Dialog Box opens
The Value is the actual numeric value.
The Value Label is the string label attached to that numeric value.

4 Type 0 in the Value text box
5 Type Male in the Value Label text box
6 Click Add to add it to the list
7 Repeat the process for Female, using 1 as the Value
8 Click OK when finished
9 Now you see the labels in the Values column.
**Entering Data**

1. Click on the **Data View** to enter data.

2. In the first row, **enter 25 for age, 1 for gender and 100,000 for income**.
   - Note: The variable name and row number of the active cell are displayed in the top left corner of the Data Editor.
   - Note: When you select a cell and enter a data value, the value is displayed in the cell editor at the top of the Data Editor.

3. Choose **View/Value Labels** if not already chosen.

4. Now you will see the **descriptive values** that we defined earlier for the gender variable.

5. If the Value Labels are being shown, you will see that there is a drop-down list for data entry. For data entry purposes, this has the added benefit of being able to use the convenience of a drop down list and thus allows suggestion of an appropriate value.

6. **Enter more data to see how drop down list is used.**
Analyzing Data

The Analyze Menu contains the list of statistical analysis categories.

Most of the categories have several procedures which are available within the category. This is indicated by an arrow to the right of the category within the Analyze menu.

We will go over a few common analyses within the **Descriptive Statistics category**.

The help section, particularly the **Statistics Coach** section, is very useful for detailed information regarding statistical analyses.

**Descriptive Statistics**

These are summary statistics. Any statistic that is selected will be summarized in the output window.

**FREQUENCIES**

Displays frequency tables that contain both the number and percentage of cases for each observed value of a variable.

1. Open the demo.sav file.
2. Choose Analyze>Descriptive Statistics>Frequencies
3. The Frequencies Dialog Box opens up
4. Select your variables
   - Marital status
   - Age in years
5. Then select your statistics:
   - Click Statistics for descriptive statistics for quantitative variables.
   - Click Charts for bar charts, pie charts, and histograms.
   - Click Format for the order in which results are displayed
6. The Viewer window opens up with your output file. The output files are covered below.
7. To make changes to the analysis, choose the **Dialog Recall** button on the toolbar:
   - For instance, choose frequencies, and add Retired as a variable

Then choose the analysis and make the changes.
CROSS TABS

Displays two-way tables and their measures

1. Open the demo.sav file.
2. Choose Analyze>Descriptive Statistics>Crosstabs
3. The Crosstabs Dialog Box opens up
4. Select your variables
   - Choose Years with current employer as the Row item
   - Choose Job satisfaction as the Column item
5. Click Statistics to see options

Reports

This can be used for summary statistics and/or individual case listings for one or more variables.

CASE SUMMARIES

1. We will use Case Summaries to report a summary of wireless services by gender.
2. Open the demo.sav file
3. Choose Analyze>Reports>Case Summaries
4. The Summarize Cases Dialog Box opens up
5. Select your variables.
   - Choose wireless service as the variable to summarize and the Age in years as the grouping variable
6. Uncheck the Display Cases box. We want a summary, not individual cases.
7. Then select your statistics.
8. Choose continue
9. Choose OK

Statistics Coach

A tool under the help section which prompts you with graphics and simple questions.

Viewer window and Output files (*.spo)

The results from data analysis are displayed in the Viewer, or Output Window.

The Viewer is divided into two panes:

- The outline pane is the pane of the Viewer which contains an outline view of the objects
  - For example, the Title object represents the bold text title of the statistics.
  - Everything under Frequencies icon refers to objects associated with the frequency statistics.
- The **contents pane** is the right pane which contains statistical tables, charts, and text output.

You can use the Viewer to:

- Navigate in the outline when you have large datasets to see the corresponding item in the contents panes.
- Collapse/Expand the outline view to hide/show the results- Click the plus/minus or select **View/ Expand** or **View/Collapse**.
- Move an item in the outline pane, which, in turn, moves the corresponding item in the contents pane. Change the display order of results by moving selected items (select and drag)
- Show or hide selected tables and charts - double click or select **View / Hide**
- Edit tiles, change table properties etc.

**Edit/Options** allows you to control what is displayed in the output.

**Charts and Graphs** are also displayed in the Output Viewer:

1. Choose **Graphs>Bar**
2. The **Bar Charts Dialog Box** opens
3. Click **Clustered** and **Define**
4. Choose variables and click OK
   - Years at Current Employer as **Category Axis**
   - Job satisfaction as **Define Clusters by**

**Note:** If several analyses are run during one session, the results will be appended to the same window. A separate output window for each analysis can be created by opening new Output windows before running each procedure. Choose **File>New>Output**.

**Saving Output**

The contents of the Viewer can be saved by choosing **File>Save**.

The default saved file type is a **Viewer** document. The saved document includes both panes of the Viewer window (the outline and the contents).
Export Output

File>Export Output saves output in HTML, text, Word/rtf and Excel formats. It also saves charts in a variety of common formats used by other applications.

You can export all objects in the Viewer, all visible objects, or only selected objects.

Three export options

   - For HTML and text format, charts are exported in the currently selected chart export format. For HTML document format, charts are embedded by reference, and you should export charts in a suitable format for inclusion in HTML documents. For text document format, a line is inserted in the text file for each chart, indicating the filename of the exported chart.
   - For Word-RTF format, charts are exported in Windows metafile format and embedded in the Word document.
   - Charts are not included in Excel documents.

2. Output Document (No Charts). Exports pivot tables and text output. Any charts in the Viewer are ignored.

3. Charts Only. Available export formats include: Windows metafile (WMF), Windows bitmap (BMP), encapsulated PostScript (EPS), JPEG, TIFF, PNG, and Macintosh PICT.