# Using SPSS, Chapter 2:
Descriptive Statistics

Chapters 2.1 & 2.2 Descriptive Statistics

SPSS can calculate all sorts of descriptive statistics on raw data. However, they are not all found in the same place. This page demonstrates the options available under the `Descriptives` menu and the next page demonstrates those found under the `Frequencies` menu.

- **Mean, Standard Deviation, Variance, Range, Minimum, Maximum**

$$\text{Analyze} \rightarrow \text{Descriptive Statistics} \rightarrow \text{Descriptives...}$$

1. The `Descriptives` window opens (below left). Send the variables you want to the `Variables` box.

2. Click options and select the statistics you want displayed from the `Options` window (below right).

3. Click Continue in the `Options` window and OK in the `Descriptives` window.

4. The requested statistics are displayed in the SPSS Viewer window (bottom figure).

Descriptive Statistics
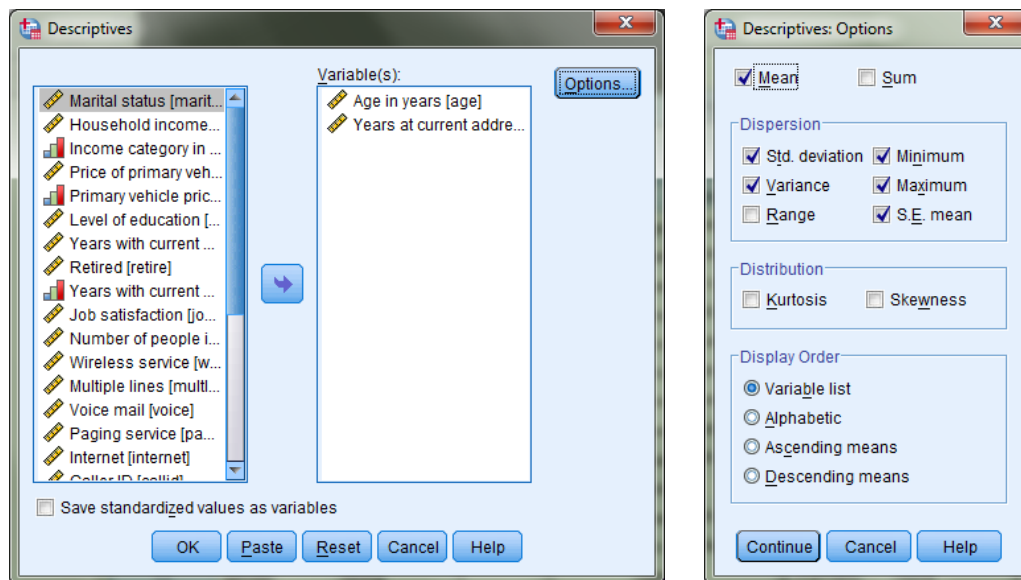
	N	Minimum	Maximum	Mean		Std. Deviation	Variance
	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Statistic
Age in years	6400	18	77	42.06	.154	12.290	151.032
Years at current address	6400	0	56	11.56	.124	9.938	98.767
Valid N (listwise)	6400						

- **Mean, Median, Mode, Standard Deviation, Variance, Range**
  Quartiles, Percentiles, and some basic chart tools (Bar, Pie, Histogram)
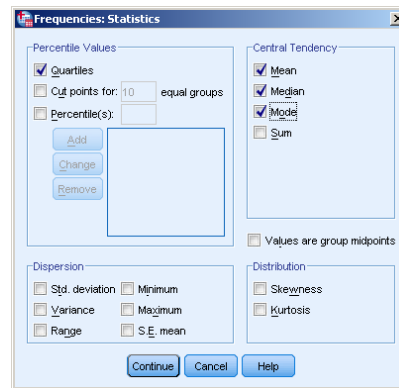
  Analyze → Descriptive Statistics → Frequencies...

1. The `Frequencies` window opens. Send the variables you want to the `Variables` box. If you uncheck `Display frequency tables` you'll get an error message stating that there is nothing to display. That's OK provided you perform the at least one of the following.

2. Click Statistics... and select the statistics you want displayed. Click Continue

3. If you want a chart, click Charts... to select from three basic chart options. Click Continue.

4. Click OK in the `Frequencies Window` and the requested statistics and/or charts are displayed in the SPSS Viewer window.
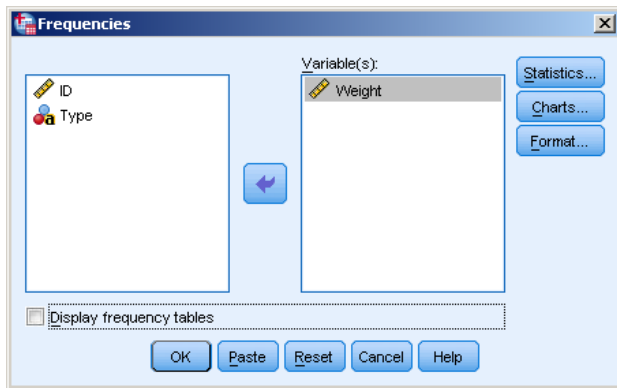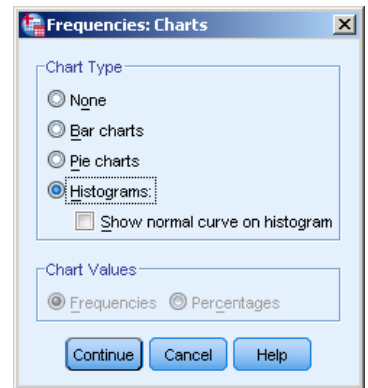
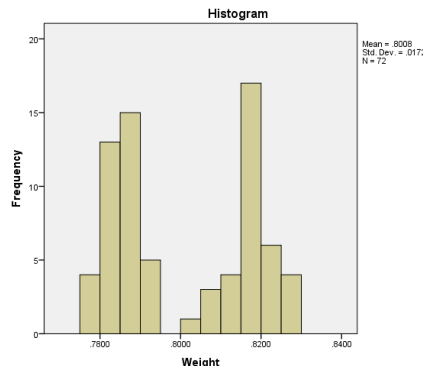Frequencies Window   Statistics Window   Charts Window

Statistics Output   Charts Output

Chapter 2.4 - Creating Box Plots

The video demonstrates how to create single and side-by-side box plots in the order outlined below. It is worth mentioning that SPSS defines outliers as those variables that are more than 1.5 IQR's below Q_1 or above Q_3. These outliers are depicted with points on the graph beyond the *whiskers*.

- **Create a single box plot. This can be done if the data is in one or two columns.**

Graphs → Chart Builder

1. Select the Simple Boxplot option from the Gallery and drag it to the Chart Preview area.
2. Drag your quantitative variable to the Y-Axis? box.
3. Click OK.

- **Create side-by-side box plots when the variables are in two columns.**

Graphs → Legacy Dialogs → Boxplot ...

1. Highlight the Simple boxplot.
2. In the Data in Chart Are box, select Summaries of separate variables.
3. Click Define.
4. Highlight and move the quantitative variables you desire into the the Boxes Represent: area.
5. Click OK.

- **Create side-by-side box plots when the variable is in one column.**

Graphs → Chart Builder

1. Select the Simple Boxplot option from the Gallery and drag it to the Chart Preview area.
2. Drag your quantitative variable to the Y-Axis? box.
3. Drag your variable that defines the different box plots to the X-Axis? box.
4. Click OK.

Getting Started

When you start SPSS there are two windows that are displayed.

- **SPSS Statistics Data Editor:** This is the window contains the actual data and data properties. There are two tabs at the lower left corner of this window.

  - **Data View:** When you click this tab you actually see the data arranged by column. The data should be in standard format where rows = cases and columns = variables.
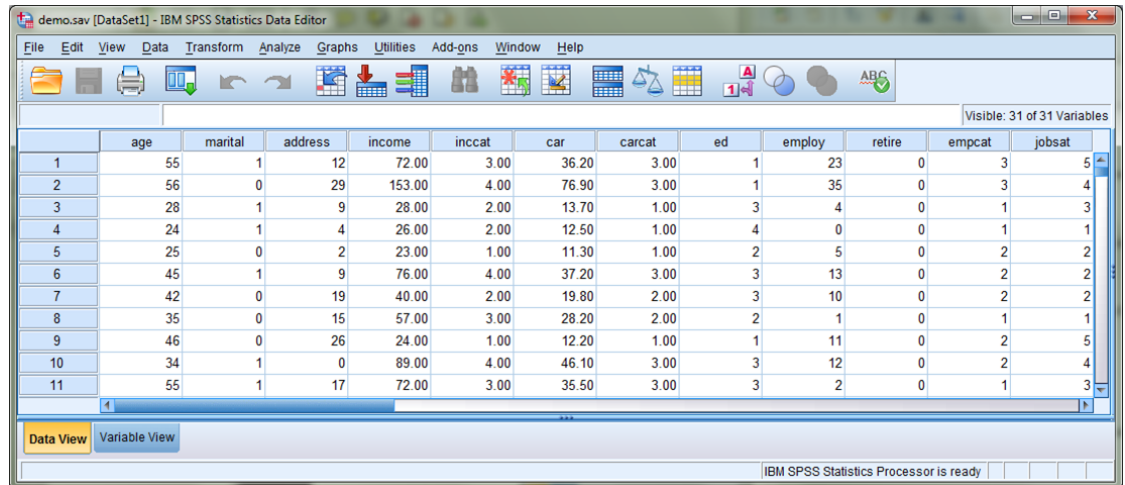
  - **Variable View:** This is where you define various attributes for each variable.
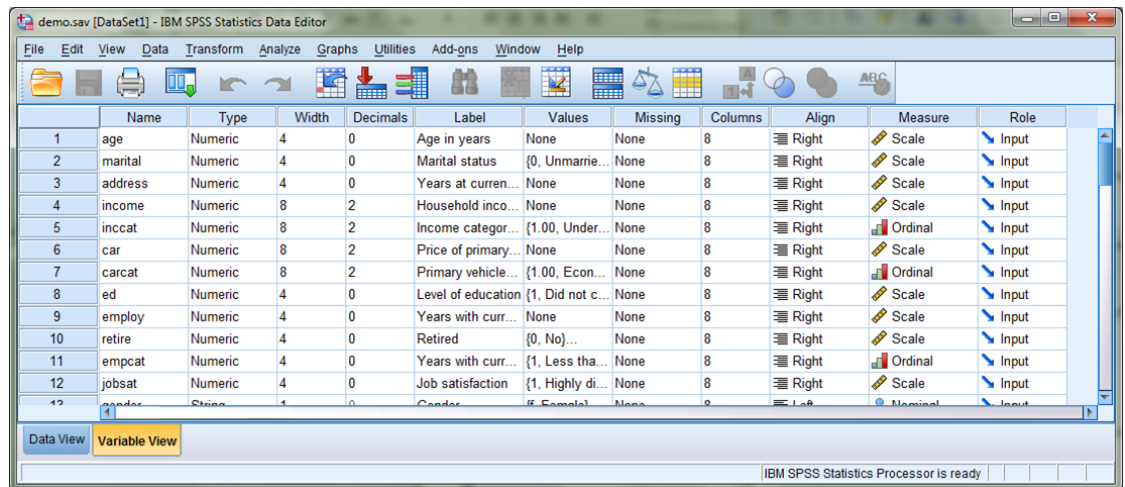
- **SPSS Statistics Viewer:** This is the window that depicts that actual commands used to perform tasks such as opening data, displaying graphs, conducting tests, etc. This is also the window where the results are displayed such as tables and graphs.
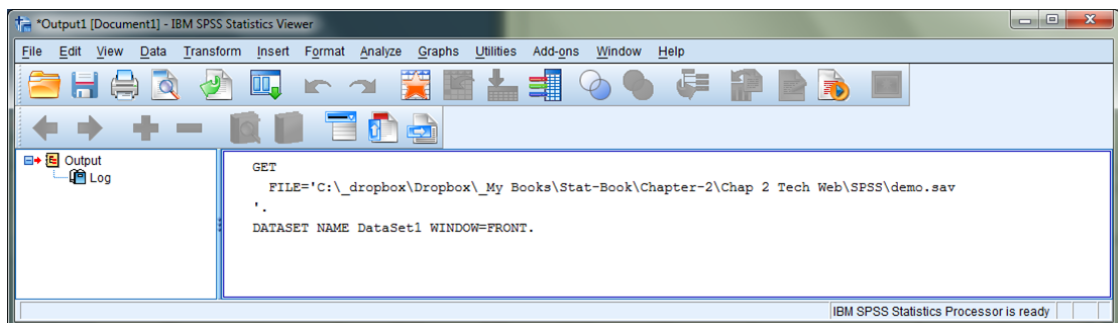
Creating and Importing Data

- There are two ways to get data into SPSS.

  - You can enter the data by typing it directly into the data editor.
  - You can open an existing data file by selecting the $\boxed{\texttt{File}}$ tab, then $\boxed{\texttt{Open}}$, then $\boxed{\texttt{Data...}}$.
    Then select the type of file from the list of options. If it is not already an SPSS (`.sav`) data file,
    you will be prompted to answer some questions. For example, if you open an Excel file it may ask
    which worksheet and whether or not the first row contains labels.

- Make sure your data is formatted as described below.

  - <u>Rows = Cases</u>
    Each row represents a case such as each respondent to a questionnaire.

  - <u>Columns = Variables</u>
    Each column represents a variable being tracked or measured. For example, the answers to a specific
    question on a questionnaire defines it's own variable (column). As such, each row represents an
    individual case for all variables.

  - <u>Cells contain values</u>
    Each cell contains a single value of a variable for a case.

  It is possible to enter data in the form of a frequency table but then you must do some alterations
  before analyzing such data.

- Once you have the data opened in the data editor, click the $\boxed{\texttt{Variable View}}$ tab at the bottom of the
  data editor. In this view, each variable is now a row and you must make sure all your variables are
  defined appropriately. The most important distinctions are

  - $\boxed{\texttt{TYPE}}$: The most common types are
    * `Numeric`: Used for quantitative data. These are numbers with no commas and a period
      delimiting the decimal places. SPSS will not allow you to enter non-numeric characters into a
      cell of numeric type.
    * `Date`: Used for dates or times from a menu of formats.
    * `String`: Used for qualitative data. Avoid symbols such as *, -, +, ?, etc.
  - $\boxed{\texttt{Measure}}$: There are three levels of measurement.
    * `Scale` is for ratio or interval levels of measurement.
    * `Ordinal` is for ordinal or ranked data.
    * `Nominal` is for qualitative data.
  - $\boxed{\texttt{Values}}$: If you have numeric values representing qualitative data such a 1=male and 0=female,
    you will probably want this to be labelled accordingly in graphs and outputs. Click on the cell in
    the `Values` column for that variable and assign labels for each value.