# Chapter One: Getting Started With IBM SPSS for Windows

Edward E. Nelson

## Using Windows

The Windows start-up screen should look something like Figure 1-1. Several standard desktop icons will always appear on start up. Note the Taskbar along the bottom, with the Start icon at the far left. If you're using your own your own computer, you can do quite a bit of customizing of your desktop by choosing your favorite colors and scenes, rearranging and adding icons, moving the Taskbar to a different location, hiding it from view, and so on.



**Figure 1-1**

Although Windows makes use of the right-button key on the mouse, we will only use the left button for now. When we say to click on the mouse button, it will always mean the left one. A single click will often take you where you want to go, but if one click doesn't do anything, try a double click. (Double clicking means to press the left mouse button twice in rapid succession. If nothing seems to happen, you probably need to double click more rapidly.)

Now click on the Start icon and the Start menu will appear. Your screen should look something like Figure 1-2. This is the basic Start menu. It can also be customized by adding your most often used programs to it. Without clicking anything yet, move the mouse pointer up and down the Start menu. As you encounter each item, it will become highlighted. Some of the items on the Start menu have little downward pointing arrows on the right. If one is present, that means there are more menus for that item.
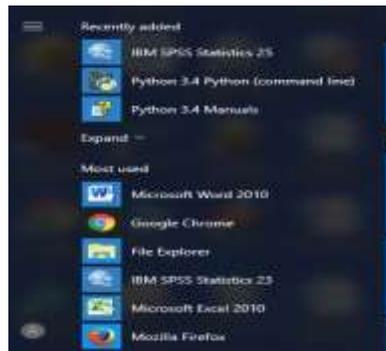


**Figure 1-2**

## Starting IBM SPSS for Windows

The IBM SPSS 25 icon should be on the Start menu.  If you are using a computer in a lab, it is common for the icon to be placed in a folder.  If you customize your computer, all you have to do to start IBM SPSS is to point to the IBM SPSS 25 icon on the desktop and double click.  Then wait while IBM SPSS loads.

After IBM SPSS loads, you may, depending on how IBM SPSS is set up, get a menu that asks, "What's new?" at the top.  For this tutorial, click Close to get rid of this.  Now the screen should look like Figure 1-3.

The Taskbar is probably visible at the bottom of the screen.  It shows an IBM SPSS icon. You can also click on this icon to open IBM SPSS.
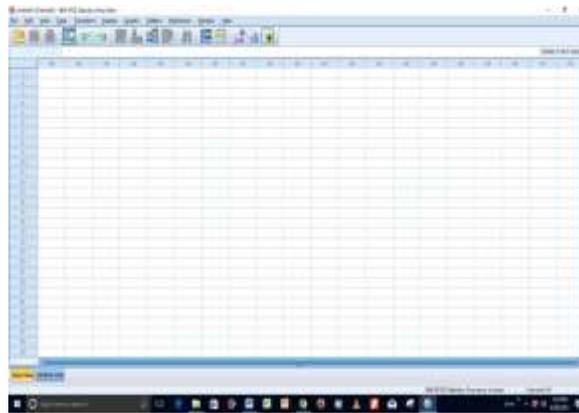
Next, observe the three small symbols in the uppermost right-hand corner of the IBM SPSS window.  The one furthest to the right, with an X in it, is used when you want to close any program you are using.  Don't worry if you click on it accidentally, a dialog box will pop up asking if you want to save anything that changed since the last time you saved your work, before it actually lets you exit the program.



**Figure 1-3**

The middle of the three small symbols allows you to expand the window in which you are working so it fills up the whole screen, or to shrink it down to a smaller size.  If the middle square shows two cascading rectangles in it, the window is already as big as it can get—clicking on this square will reduce the window in size.  Try this now.  In this shrunken window, the middle button now shows only one rectangle.  Click on it to get back to the full screen view.

The last symbol, to the left of the other two, has what looks like a minus sign on it.  Click this and watch what happens.  You haven't actually closed IBM SPSS.  Click on the icon in the task bar and IBM SPSS will reappear on your screen.

## Leaving IBM SPSS

We're not ready to actually use IBM SPSS yet, so let's close it.  There are at least four ways to do this.  Move your mouse until the arrow is pointing at the word File in the

upper-left hand corner of the screen and press the left mouse button once.  A menu will appear.  Move the arrow so it is pointing at the word Exit and press the left mouse button. This should close IBM SPSS.

There is a second technique that can do the same thing.  (Computers usually have more than one way to do everything.)  Go back into IBM SPSS and move your mouse until the arrow is pointing at the word File in the upper-left corner of the screen and press the left mouse button once, but this time, don't release the mouse button.  Hold the mouse button down and move the mouse down until the word Exit is highlighted.  Now release the mouse button and IBM SPSS should close.  This is called click and drag and is another way to use your mouse.

The third way to close IBM SPSS is to point your mouse at the IBM SPSS icon in the upper-left corner of the screen.  The icon will be just to the left of the words Untitled: IBM SPSS Statistics Data Editor.  Move your mouse to the icon and double click on it. This has the same effect as the first two procedures; it closes IBM SPSS.

The final and fourth way to close IBM SPSS is to point your mouse at the X in the upper-right corner of the screen and click.  IBM SPSS will close.

Now you know how to move your mouse around and how to start and close IBM SPSS. We'll show you more about Windows, but not much more.  If you want to learn more about Windows, there are a lot of books available.  The nice thing about Windows is that you don't have to know much about it to use it.


**Looking at Data**

There are two data files to be used with this tutorial:
- GSS16A.  This is a subset of the 2016 General Social Survey.  The General Social Survey is a large, national probability sample done every two years.
- COUNTRIES.  This file includes data on almost all of the world's countries.

You can download these data files from the web by going to **IBM SPSS Statistics for Windows 25: A Basic Tutorial**. Also, if you are in a computer lab, someone may have copied one or both of the data files onto your hard drive.

Let's begin by starting IBM SPSS just as you did above.  Your screen should look like Figure 1-3.  (If you see a box asking, "What's new?" click on Close to close this box.) At the very top of the screen, you will see the words IBM SPSS Statistics Data Editor. Just below that line will be the menu bar with the following options:  File, Edit, View, Data, Transform, Analyze, Graphs, Utilities, Extensions, Window, and Help.  Point your mouse at File and press the left mouse button.

A box will open that is the **File** menu.  Point your mouse at **Open** and then at **Data** and click.  (Also, you could have gotten to this point by clicking on the **Open Data** icon just below **File** on the menu bar.)

This opens a larger box called the Open Data box.  Here you need to tell IBM SPSS where to find the data file to open.  In the upper part of the box, you'll see **Look in.**  Find the folder that contains the data file and click on the file name, GSS16A, to highlight it and then click on **Open**.  In a few seconds, your data matrix will appear.

A data matrix is a very important concept.  The rows contain the cases and the columns contain the variables.  (If you're familiar with spreadsheets, that's what this is.)  Row 1 is case 1; row 2 is case 2, and so on.  The top of each column contains the variable name.  In this data set, the variable names are abbreviations like *abany* and *abdefect*.  Unfortunately, the abbreviations for the variable names don't tell you very much.  We need some way to find out what these variables are.  So try this.  On the menu bar at the top of your screen, you'll see the word **Utilities**.  Point your mouse at **Utilities** and click the left button.  This will open the **Utilities** menu.  Point your mouse at **Variables** and click again.  Your screen should look like Figure 1-4.

You'll see a list of all the variables in your data on the left side of the little window. (Also, see Appendix A for a list of variables.)  Point your mouse at any of these variables and click.  To the right of the Variable List you'll see a short description of this variable. For example, point your mouse at the variable *abany* and press the left mouse button.  This


**Figure 1-4**

question asked if respondents thought that obtaining a legal abortion should be legal for a woman if she wants it for any reason.  The possible answers are YES (value 1), NO (value 2), DK or don't know (value 8), NA or no answer (value 9), and IAP or not applicable (value 0).  (Not applicable includes people who were not asked the question.) As you will see, these values are very important.

Now you know how to open a preexisting data file in IBM SPSS and how to find out what the variables are in the file.  We will tell you more about this later, but here we just want to give you a brief introduction to IBM SPSS.

## A Brief Tour of IBM SPSS

Now that you have the file opened, let's look at some things you can do with IBM SPSS. You're already familiar with the variable *abany*.  Let's find out what percent of people surveyed thought it ought to be legal for a woman to have an abortion for any reason.  (If

you have the Variables window open showing the variable labels and values, point your mouse at the cancel button and click it.)  On the menu bar you will see Analyze.  Point your mouse at Analyze and click it.  A box opens that looks like Figure 1-5.
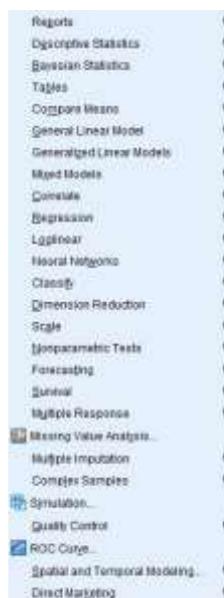
This lists the statistical procedures in IBM SPSS.  We want to use Descriptive Statistics so point your mouse at Descriptive Statistics.  This opens another box listing the statistical procedures you can use to summarize your data.  Point your mouse at Frequencies and click it.  This opens the Frequencies box.  Since *abany* is the first variable in the data, it's already highlighted.[1]  Point your mouse at the right arrow next to the list of variables and click it.  The variable *abany* will move to the box called Variable(s).  This is how you select variables.  Point your mouse at OK and click it.  In a few seconds, a new screen should appear that looks like Figure 1-6.  We are now in a different part of IBM SPSS called the Output Window.  This is where the results, or output, are displayed.

The Output Window is divided into two vertical frames or panes.  The left-hand pane contains the outline of the output or information that IBM SPSS gives you.  This information is in outline form and can be used to select what you want to view.  Simply click on the information you want to look at and that information will appear in the right-hand pane.  You can also collapse and expand the outline by clicking on the plus and minus signs in the left-hand pane.  The plus sign indicates that the information is collapsed (or hidden) and the minus sign indicates that it is expanded (or shown).  You can use the scroll bars on each pane to scroll through the Output Window.

**Figure 1-5**

On the right side, the frequency distribution for *abany* is divided into four parts:  (1) the title, (2) notes on the table (there aren't any for this table), (3) the active dataset (i.e., the dataset you are currently using), (4) statistics (a summary of the number of missing and valid observations), and (5) the actual table showing the frequency distribution.  Click on Statistics in the left-hand pane and you will see that there were 1,825 valid and 1,042 missing cases.  Click on abortion—if woman wants for any

**Statistics**

abany ABORTION IF WOMAN WANTS FOR ANY REASON

| N | Valid | 1825 |
|---|---|---|
| | Missing | 1042 |

**abany ABORTION IF WOMAN WANTS FOR ANY REASON**

| | | Frequency | Percent | Valid Percent | Cumulative Percent |
|---|---|---|---|---|---|
| Valid | 1 YES | 819 | 28.6 | 44.9 | 44.9 |
| | 2 NO | 1006 | 35.1 | 55.1 | 100.0 |
| | Total | 1825 | 63.7 | 100.0 | |
| Missing | 0 IAP | 964 | 33.6 | | |
| | 8 DK | 49 | 1.7 | | |
| | 9 NA | 29 | 1.0 | | |
| | Total | 1042 | 36.3 | | |
| Total | | 2867 | 100.0 | | |

**Figure 1-6**

---

[1] Instead of seeing a list of variables, you may see a list of variable labels.  You can change this so IBM SPSS displays the list of variables.   To do this, click on Edit in the menu bar, then click on Options and on the General tab.  Look for Variable Lists in the General tab and click on Display names.  You can put the variable names in alphabetical order by clicking on Alphabetical just below Display names.  Finally, click on OK.

**reason** and you will see the frequency distribution. In Figure 1-6, you can see that 819 people said yes, 1,006 said no, 49 said they didn't know, 29 didn't answer the question, and 964 were coded not applicable. (These 964 respondents were not asked this question. In survey research it's very common to ask some, but not all, of the respondents a particular question. In this case, only 1,903 of the 2,867 respondents were asked this question.) Of those who had an opinion, we want to know what percent of the respondents said yes or no, so we should look at the Valid Percents in the table. About 45% of the respondents who had an opinion thought it should be legal, while 55% thought it should be not be legal.

It would be interesting to know if men or women were more likely to favor allowing a legal abortion when the woman wants it for any reason. We're going to use a crosstab to determine this. Point your mouse at Analyze and press the left mouse button. Then point your mouse at Descriptive Statistics and finally, point your mouse at Crosstabs and press the mouse button. Your screen should look like Figure 1-7. The list of variables in your data set is on the



**Figure 1-7**

left of the screen. We want to move the variable *abany* into the box next to the list of variables where it says Rows. Click on the variable *abany*, which will highlight it.
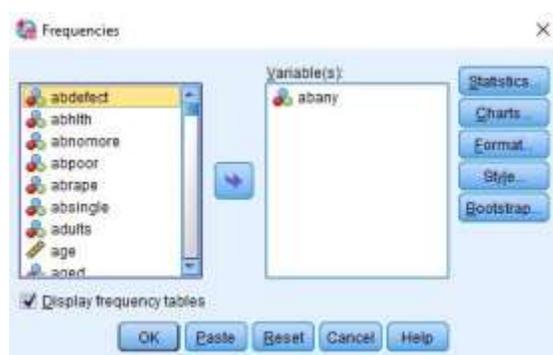
Now click on the arrow pointing to the right, which is next to the Rows box. Notice that this moves *abany* into the Rows box. We also need to move the variable *sex* into the Columns box. You will have to use the scroll bar in the box containing the list of variables to find this variable. (You can also click anywhere in this box and then type the letter **S** to move to the first variable starting with the letter S.) Point your mouse at the down arrow next to the list of variables and click. If you keep pressing the mouse button, the list of variables will move down and eventually you



**Figure 1-8**

will see the variable *sex*. Highlight it and click on the arrow pointing to the right, which is next to the Columns box. This moves *sex* into the Columns box. You screen should look like Figure 1-8.

Frequencies by themselves are seldom useful. You'll need to compute percents. To get

IBM SPSS to compute percentages, point your mouse at the button labeled Cells at the right of the screen and click on it. This will open the Crosstabs: Cell Display box. Find the box called Column Percentages and click on this box.

Now click on Continue and you will be back to the Crosstabs box. To tell IBM SPSS to run the Crosstabs procedure, click on OK. In a second your screen should look like Figure 1-9. Use the scroll bar to look at all the information that IBM SPSS gives you in the Output Window. Figure 1-9 shows the results, or "output." It shows, for example, that males and females differ very little in their opinions about a woman obtaining a legal abortion for any reason. Forty-five percent of the males and 44 percent of the females approve of a woman obtaining an abortion for any reason.



**Figure 1-9**

You can also examine other items in the survey to compare men and women. Who has more education? Is the average age at birth of first child younger for women than for men? Comparing means will answer these questions. Click on Analyze, point your mouse to Compare Means, and then click on Means. Your screen should look like Figure 1-10.



**Figure 1-10**

Now put age at birth of first child (*agekdbrn*) and years of school completed (*educ*) in the Dependent List box and *sex* in the Independent List box. By now you have a good idea how to do this. Highlight *agekdbrn* in the list of variables on the left of the screen by pointing your mouse at it and clicking. Then click on the arrow next to the Dependent List box. Do the same for *educ*. Now highlight *sex* and click on the arrow next to the Independent List box. This should move



**Figure 1-11**

*agekdbrn* and *educ* into the Dependent List box and *sex* into the Independent List box and your screen should look like Figure 1-11. Then click on OK and the output should look like Figure 1-12.

Women had their first child at an average age of 23.45 years, while the average for men is 25.84 years, a difference of a little more than two years. Now look at the mean years of school completed for men and women. There's virtually no difference (about 0.11 of a year) between men and women.

Another way of examining relationships is to look at Pearson Correlation Coefficients. One might hypothesize that respondents' education is correlated with the educational achievements of their parents. The Pearson Correlation Coefficient will tell us the strength of the linear relationship between father's education, mother's education, and the respondent's education. The closer the correlation is to 1, the stronger the linear relationship, and the closer it is to 0, the weaker the linear relationship.

Point your mouse at Analyze and press the mouse button. Now point your mouse at Correlate and then click on Bivariate. Your screen should look like Figure 1-13.

Now move the following three variables into the Variables box: *educ*, *maeduc*, and *paeduc*. These variables refer to the number of years of school completed by the respondent and the respondent's parents. Highlight each of these variables and click the arrow next to the Variables box. The screen should look like Figure 1-14. Click on OK and the correlations will appear in your Output box.
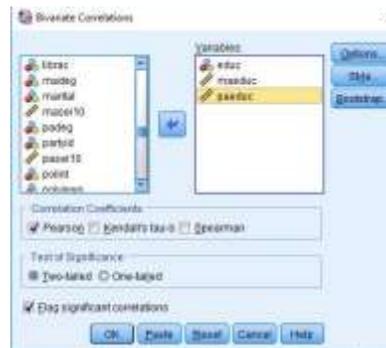
The strongest correlation is between father's and mother's education. As we predicted, there is also a strong correlation between respondent's education and parent's education.

**Figure 1-12**

**Figure 1-13**

**Figure 1-14**

## Overview of Chapters

Chapter 2 will acquaint you with how to enter new data into IBM SPSS using the Data Editor. Chapter 3 explains how to take your data, or preexisting data, and modify it in a way that makes it easier to understand. Chapter 4 starts the sections where you really get to see the results of your work. In Chapter 4 you will learn how to look at each variable,

one at a time.  We call this univariate analysis.  Chapters 5 through 7 will show you how to look at two variables at a time, or what we call bivariate analysis.  Chapter 5 will show you how to create a cross tabulation.  Chapter 6 shows you alternative ways of comparing two variables and Chapter 7 will show you how to do this using linear regression techniques.  Chapter 8 shows you how to explore relationships among sets of variables using multivariate cross tabulation and multiple regression.  This is called multivariate analysis.  Finally, Chapter 9 will show you how to present your data effectively and will include charts and tables.