Overview

IBM SPSS (Statistical Package for the Social Sciences) is a data management and analysis product produced by IBM SPSS, Inc. in Chicago, Illinois. Among its features are modules for statistical data analysis, including descriptive statistics such as plots, frequencies, charts, and lists, as well as sophisticated inferential and multivariate statistical procedures like analysis of variance (ANOVA), factor analysis, cluster analysis, and categorical data analysis. SPSS is particularly well-suited to survey research, though by no means is it limited to just this topic of exploration.

SPSS is a modular product. That is, it requires the Base System module to run, but you may wish to use other modules to carry out specific analyses not supported by the Base product. The current version is 20.0 Lehman College has a site license for a much earlier version which is serviceable and has all the features you might need to handle your data.

Many of the features we will use in this beginner session are intended to introduce you to some common sense data manipulation. These include:

Starting SPSS

Entering small sets of data directly into SPSS

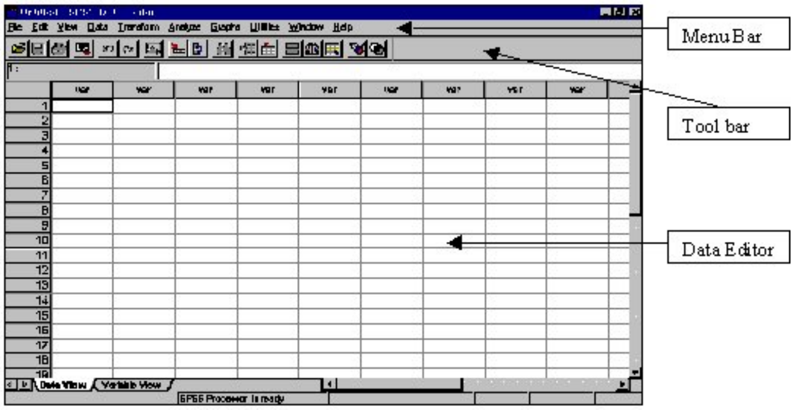
Importing larger sets of data from Excel spreadsheets into SPSS

Using SPSS to create descriptions of your data (frequencies, descriptions, central tendency, variability, correlations)

Starting SPSS

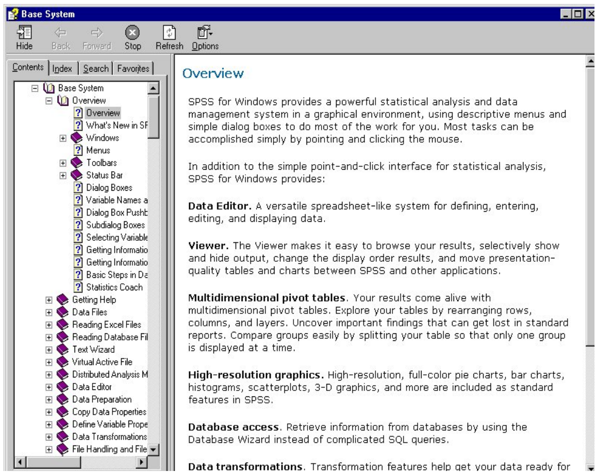
Click the **Start** button, **Programs**, the relevant software group, SPSS, and click on the SPSS icon.

The initial SPSS screen should appear, showing the *Data Editor* window, with the *Data View* window on top, and a tab at the foot of the screen giving access to the *Variable View* window. This is superimposed by a smaller window headed SPSS for Windows which you can temporarily discard by clicking on the **Cancel** button. You can switch between *Data View* and *Variable View* by clicking the appropriate tab.



Getting help

You can obtain help on SPSS at any time during your SPSS session. Features such as being able to search for specific topics are included. To access the online help system, click on the **Help** menu and **Topics** to display the following window:



Overview

SPSS for Windows provides a powerful statistical analysis and data management system in a graphical environment, using descriptive menus and simple dialog boxes to do most of the work for you. Most tasks can be accomplished simply by pointing and clicking the mouse.

In addition to the simple point-and-click interface for statistical analysis, SPSS for Windows provides:

Data Editor. A versatile spreadsheet-like system for defining, entering, editing, and displaying data.

Viewer. The Viewer makes it easy to browse your results, selectively show and hide output, change the display order results, and move presentation-quality tables and charts between SPSS and other applications.

Multidimensional pivot tables. Your results come alive with multidimensional pivot tables. Explore your tables by rearranging rows, columns, and layers. Uncover important findings that can get lost in standard reports. Compare groups easily by splitting your table so that only one group is displayed at a time.

High-resolution graphics. High-resolution, full-color pie charts, bar charts, histograms, scatterplots, 3-D graphics, and more are included as standard features in SPSS.

Database access. Retrieve information from databases by using the Database Wizard instead of complicated SQL queries.

Data transformations. Transformation features help get your data ready for

The *Contents* panel contains a list of broad topics, represented by icons of books. Double-clicking on any book will expand the contents. Selecting any one of these topics by double-clicking will provide you with the information in that topic.

If you know exactly what you want, or you wish to refer to a statistical term or a specific piece of jargon, you may prefer to use the *Index* tab. An alphabetic list of terms and topics will appear, and you can enter a term to search for. If too many similar topics are shown, use the vertical scroll bar to view the rest of the list, and double-click the topic you want.

Use **Search** to locate a **Help** topic. Within the *Search* box, enter a keyword that you would like to find help on. All **Help** topics that contain the keyword will be displayed, not just topics that begin with that word (as in the **Index**).

Opening/Importing files from Excel and opening existing SPSS files

Data organization and the Data Editor

Opening a data file

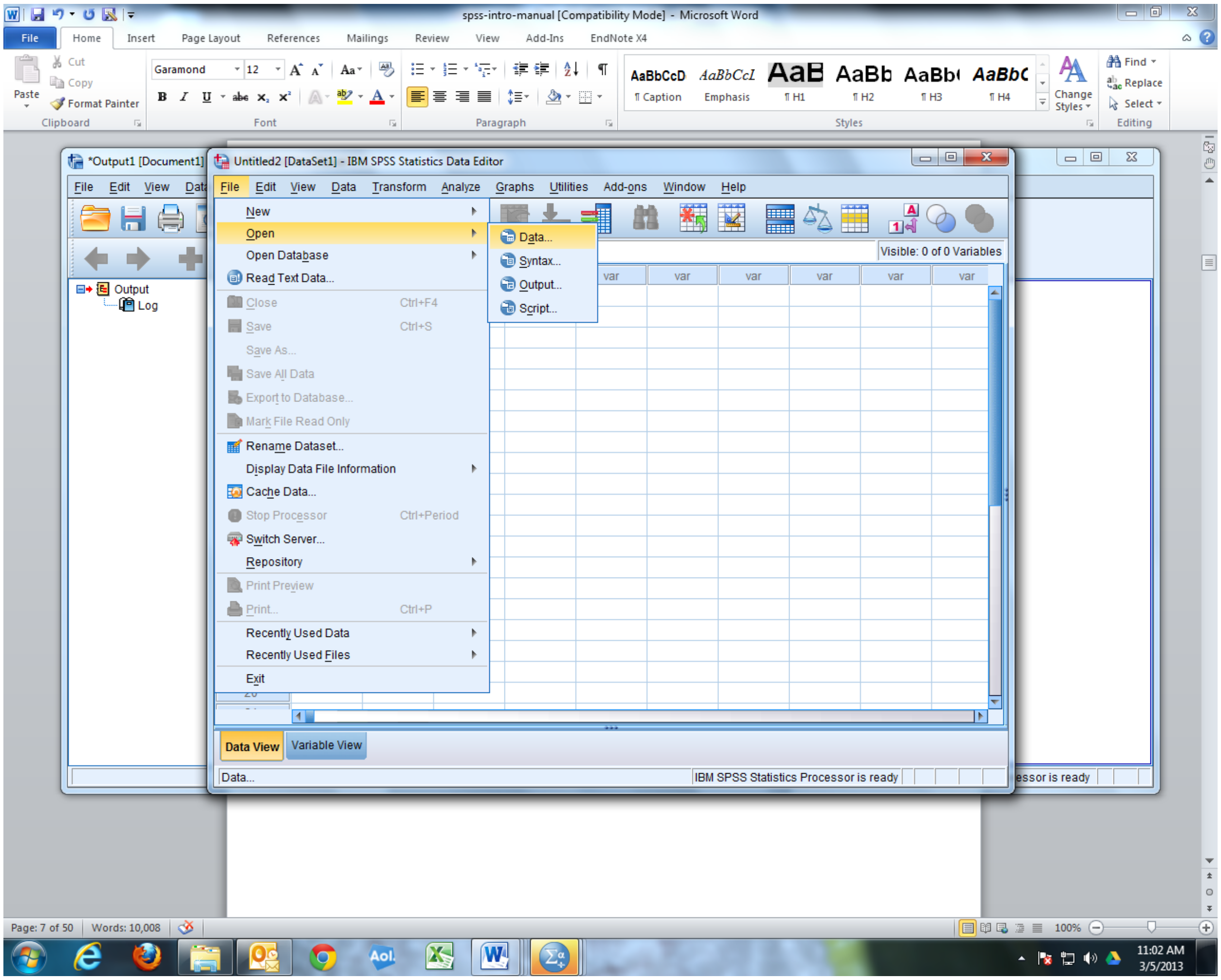
The first task will be to retrieve a simple data file to see how data is organized within SPSS. To do this using the menu:

- From the **File** menu, select **Open**, then select the **Data** from the resulting sub-menu.

The following dialog box will appear:

In the dialog box pictured opposite, you select the file from the list of files. You can select different drives and directories by clicking the drop-down arrow next to the *Look in* box.

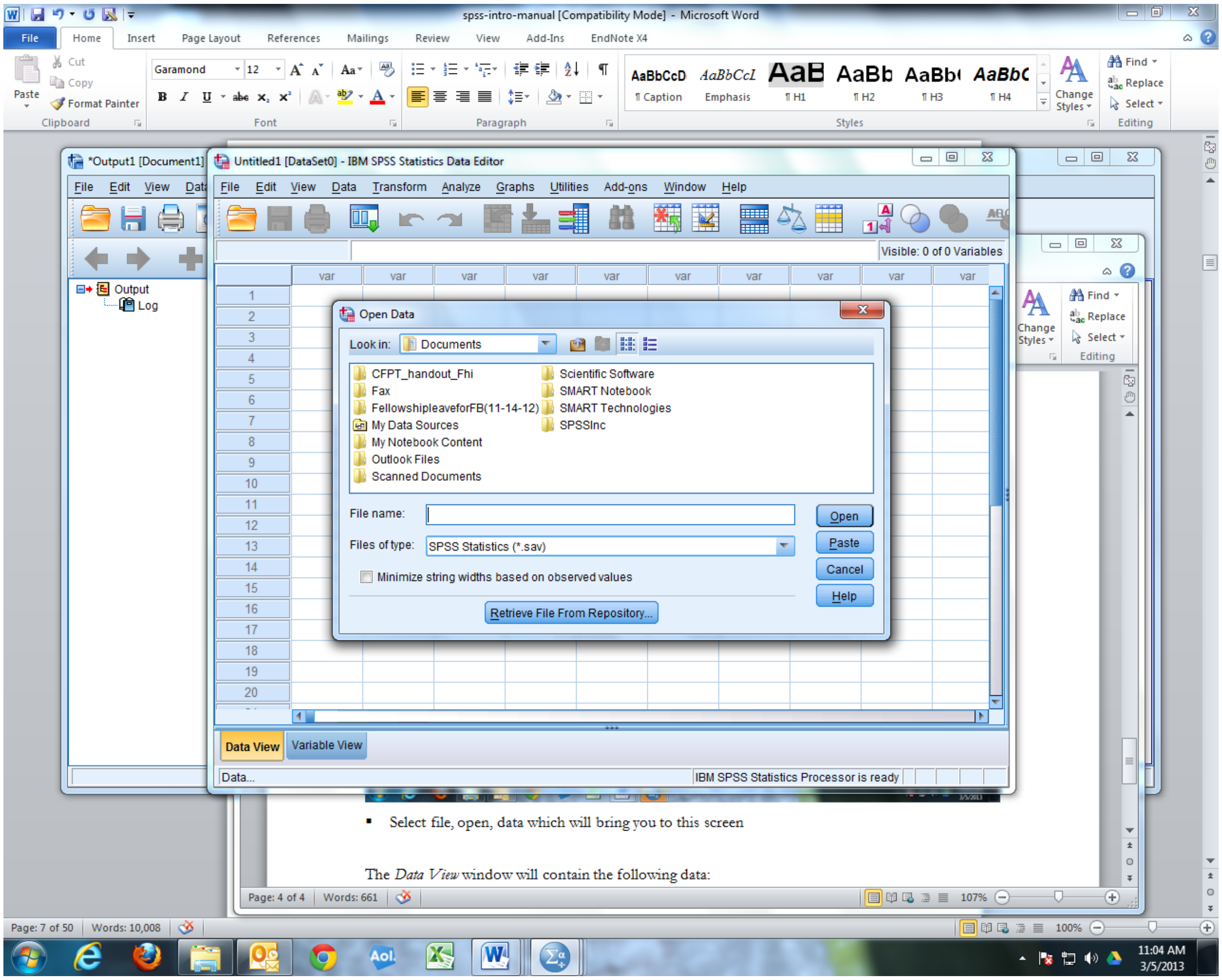
You can retrieve files created by software packages such as Excel by selecting one of the file types from the pull down list in the *Files of Type* box. This will be covered later in the course. The first file we are going to use is called sample.xls



Data...

IBM SPSS Statistics Processor is ready

- Select file, open, data which will bring you to this screen



spss-intro-manual [Compatibility Mode] - Microsoft Word
Untitled1 [DataSet0] - IBM SPSS Statistics Data Editor
*Output1 [Document1]
File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help
Visible: 0 of 0 Variables
Data View
Variable View
IBM SPSS Statistics Processor is ready

Open Data

Look in: Documents

- CFPT_handout_Fhi
- Fax
- FellowshipleaveforFB(11-14-12)
- My Data Sources
- My Notebook Content
- Outlook Files
- Scanned Documents
- Scientific Software
- SMART Notebook
- SMART Technologies
- SPSSInc

File name:

Files of type: SPSS Statistics (*.sav)

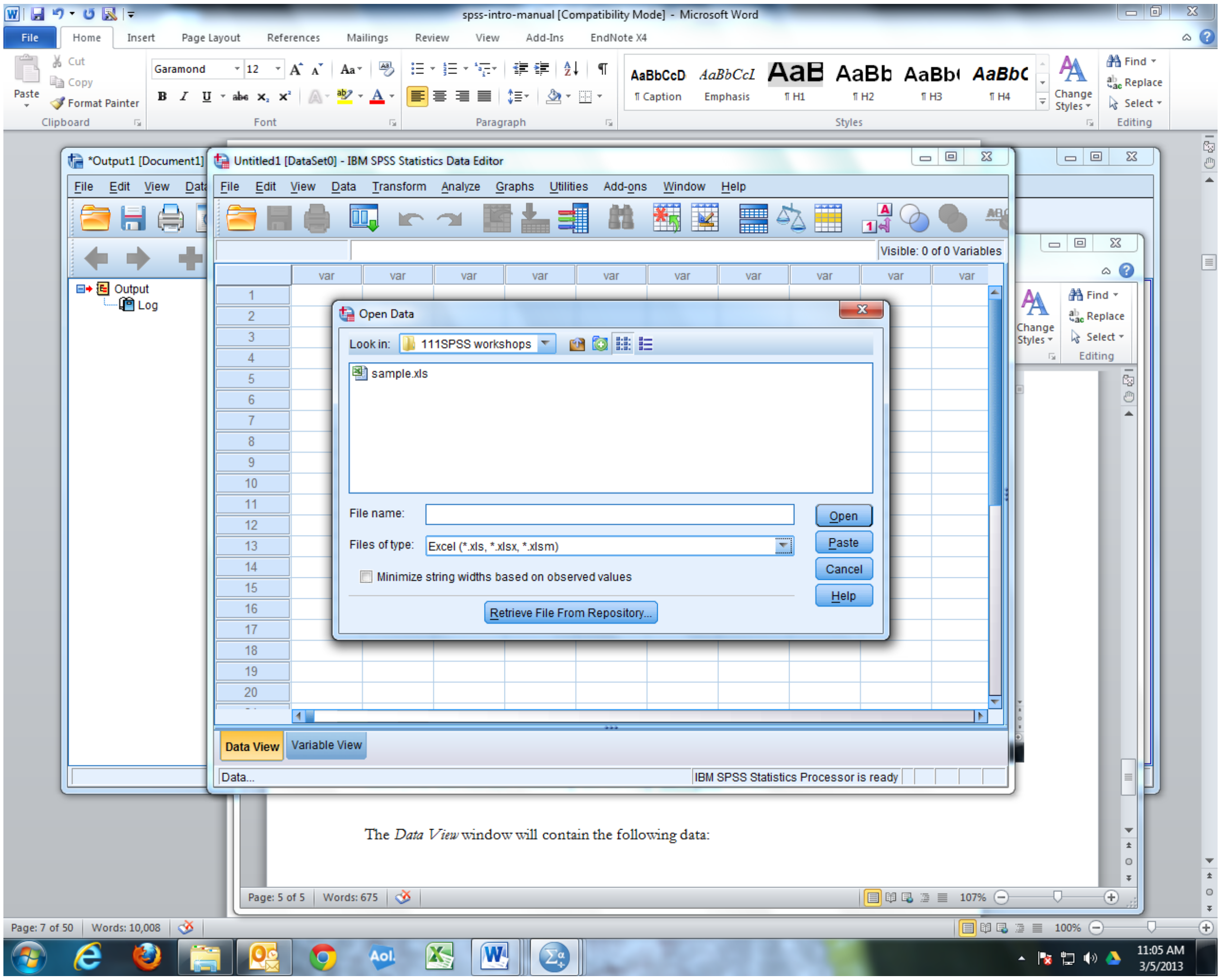
□ Minimize string widths based on observed values

Open
Paste
Cancel
Help
Retrieve File From Repository...

- Select file, open, data which will bring you to this screen

The *Data View* window will contain the following data:

Since your files for this workshop are on a flash/thumb drive, pick the drive and type of file and click OPEN.



spss-intro-manual [Compatibility Mode] - Microsoft Word

Untitled1 [DataSet0] - IBM SPSS Statistics Data Editor

Open Data

Look in: 111SPSS workshops

sample.xls

File name:

Files of type: Excel (*.xls, *.xlsx, *.xlsm)

☐ Minimize string widths based on observed values

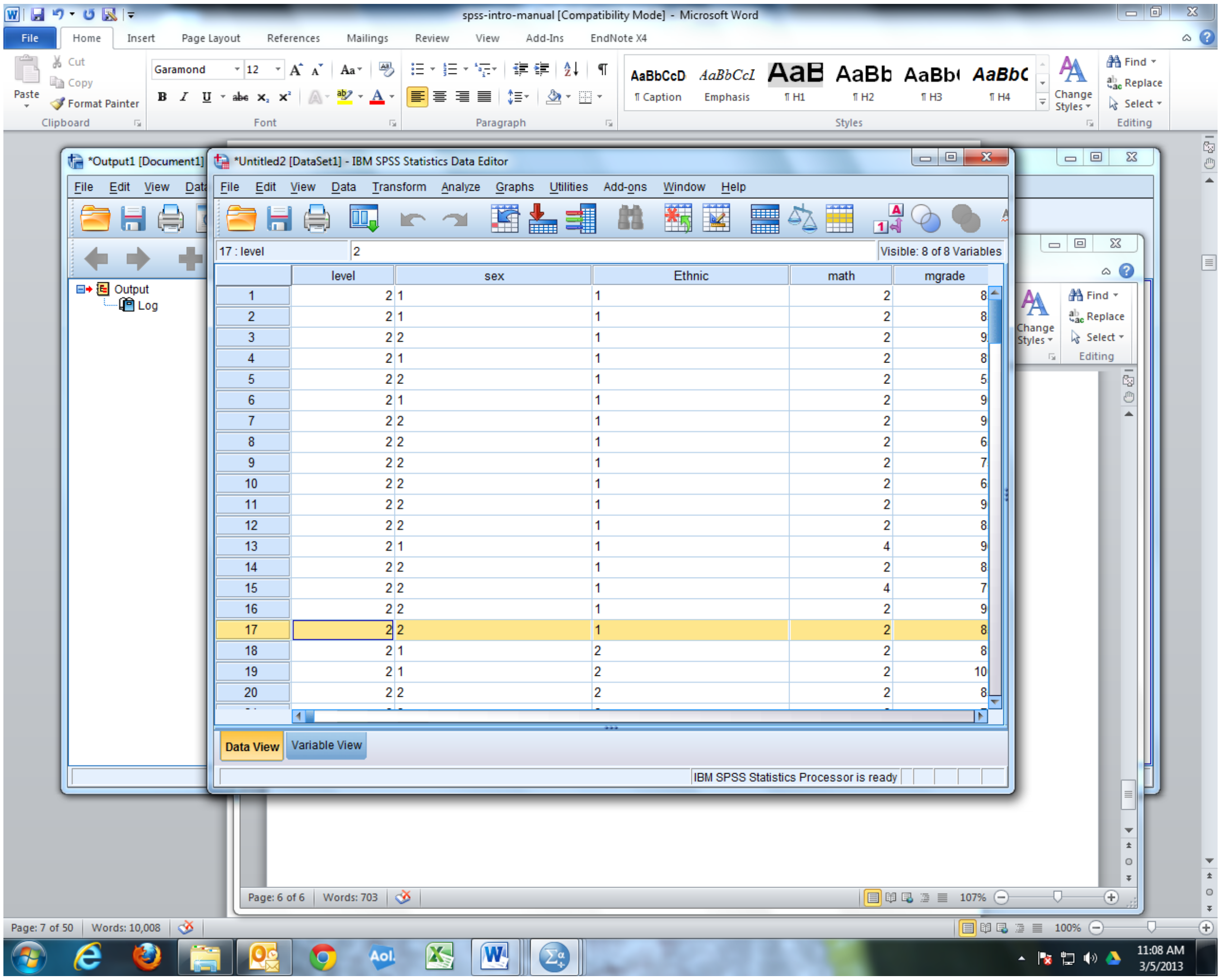
Visible: 0 of 0 Variables

IBM SPSS Statistics Processor is ready

The *Data View* window will contain the following data:

You will see the following asking you if you want to include column heads as variable names, say yes and OK

The *Data View* window will contain the following data:



17 : level 2

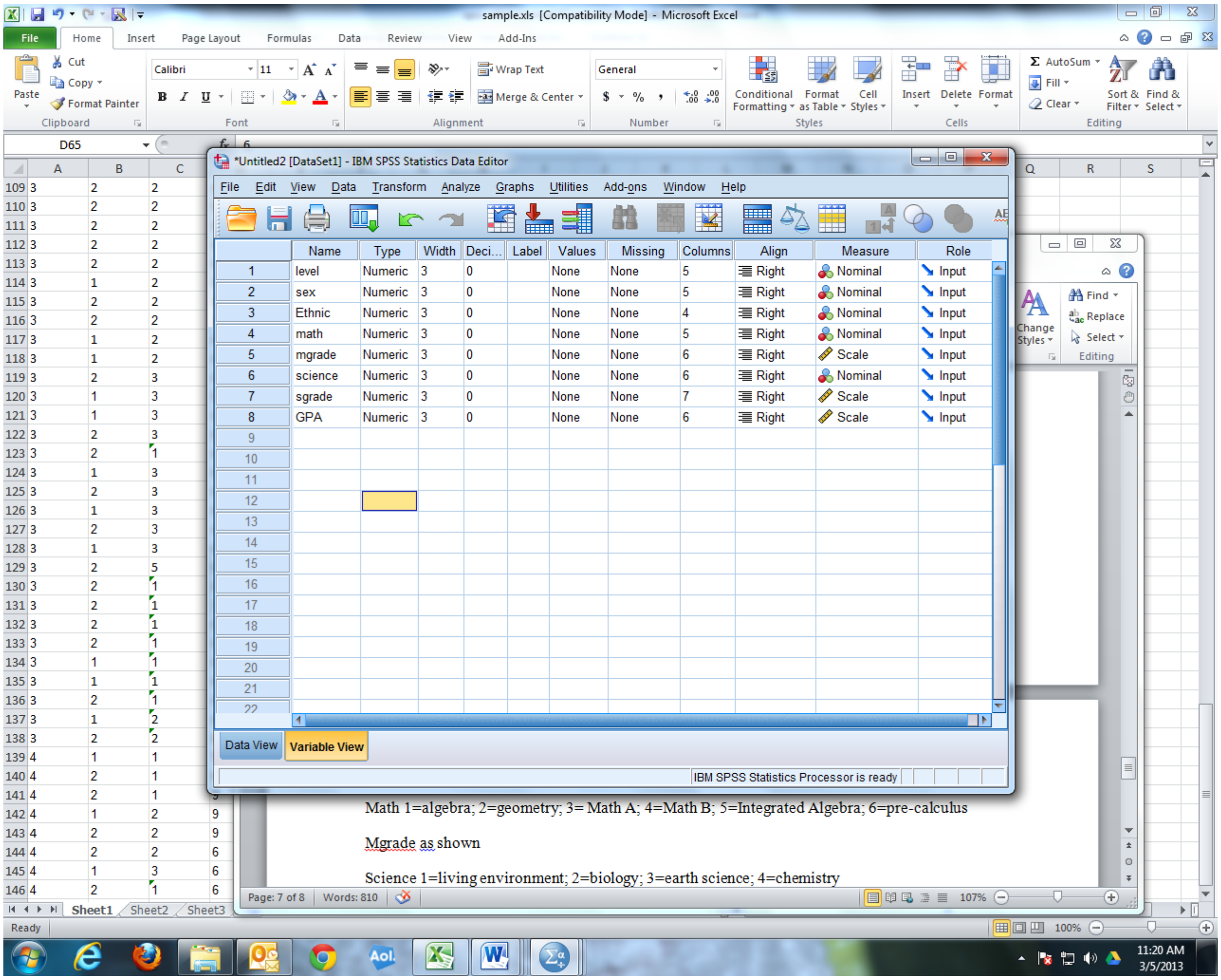
Visible: 8 of 8 Variables

	level	sex	Ethnic	math	mgrade
1	2	1	1	2	8
2	2	1	1	2	8
3	2	2	1	2	9
4	2	1	1	2	8
5	2	2	1	2	5
6	2	1	1	2	9
7	2	2	1	2	9
8	2	2	1	2	6
9	2	2	1	2	7
10	2	2	1	2	6
11	2	2	1	2	9
12	2	2	1	2	8
13	2	1	1	4	9
14	2	2	1	2	8
15	2	2	1	4	7
16	2	2	1	2	9
17	2	2	1	2	8
18	2	1	2	2	8
19	2	1	2	2	10
20	2	2	2	2	8

IBM SPSS Statistics Processor is ready

You can switch between variable view and data view by clicking on the tabs.

I have adjusted the width of the variables in the data view so that all columns are in view by dragging the cells that have the variable names to make them smaller and the variable widths to 3.



	Name	Type	Width	Deci...	Label	Values	Missing	Columns	Align	Measure	Role
1	level	Numeric	3	0		None	None	5	Right	Nominal	Input
2	sex	Numeric	3	0		None	None	5	Right	Nominal	Input
3	Ethnic	Numeric	3	0		None	None	4	Right	Nominal	Input
4	math	Numeric	3	0		None	None	5	Right	Nominal	Input
5	mgrade	Numeric	3	0		None	None	6	Right	Scale	Input
6	science	Numeric	3	0		None	None	6	Right	Nominal	Input
7	sgrade	Numeric	3	0		None	None	7	Right	Scale	Input
8	GPA	Numeric	3	0		None	None	6	Right	Scale	Input

D65 6

	A	B	C
109	3	2	2
110	3	2	2
111	3	2	2
112	3	2	2
113	3	2	2
114	3	1	2
115	3	2	2
116	3	2	2
117	3	1	2
118	3	1	2
119	3	2	3
120	3	1	3
121	3	1	3
122	3	2	3
123	3	2	1
124	3	1	3
125	3	2	3
126	3	1	3
127	3	2	3
128	3	1	3
129	3	2	5
130	3	2	1
131	3	2	1
132	3	2	1
133	3	2	1
134	3	1	1
135	3	1	1
136	3	2	1
137	3	1	2
138	3	2	2
139	4	1	1
140	4	2	1
141	4	2	1
142	4	1	2
143	4	2	2
144	4	2	2
145	4	1	3
146	4	2	1

Math 1=algebra; 2=geometry; 3= Math A; 4=Math B; 5=Integrated Algebra; 6=pre-calculus

Mgrade as shown

Science 1=living environment; 2=biology; 3=earth science; 4=chemistry

We now need to label the variables and indicate the values.

The variable label can be the same as the variable name as the names are already explicit.

The values must be added in the values column using the numbers assigned and the names as in the codes given below. (It is good to have a code book/sheet to remind you of the codes assigned at the beginning of the development of the survey/database.)

Level 1=freshman; 2=sophomore; 3=junior; 4=senior

Sex 1= male; 2=female

Ethnic 1=AA/Black; 2=Latino/Hispanic; 3=White; 4=Asian; 5 = Other

Math 1=algebra; 2=geometry; 3= Math A; 4=Math B; 5=Integrated Algebra; 6=pre-calculus

Mgrade as shown

Science 1=living environment; 2=biology; 3=earth science; 4=chemistry

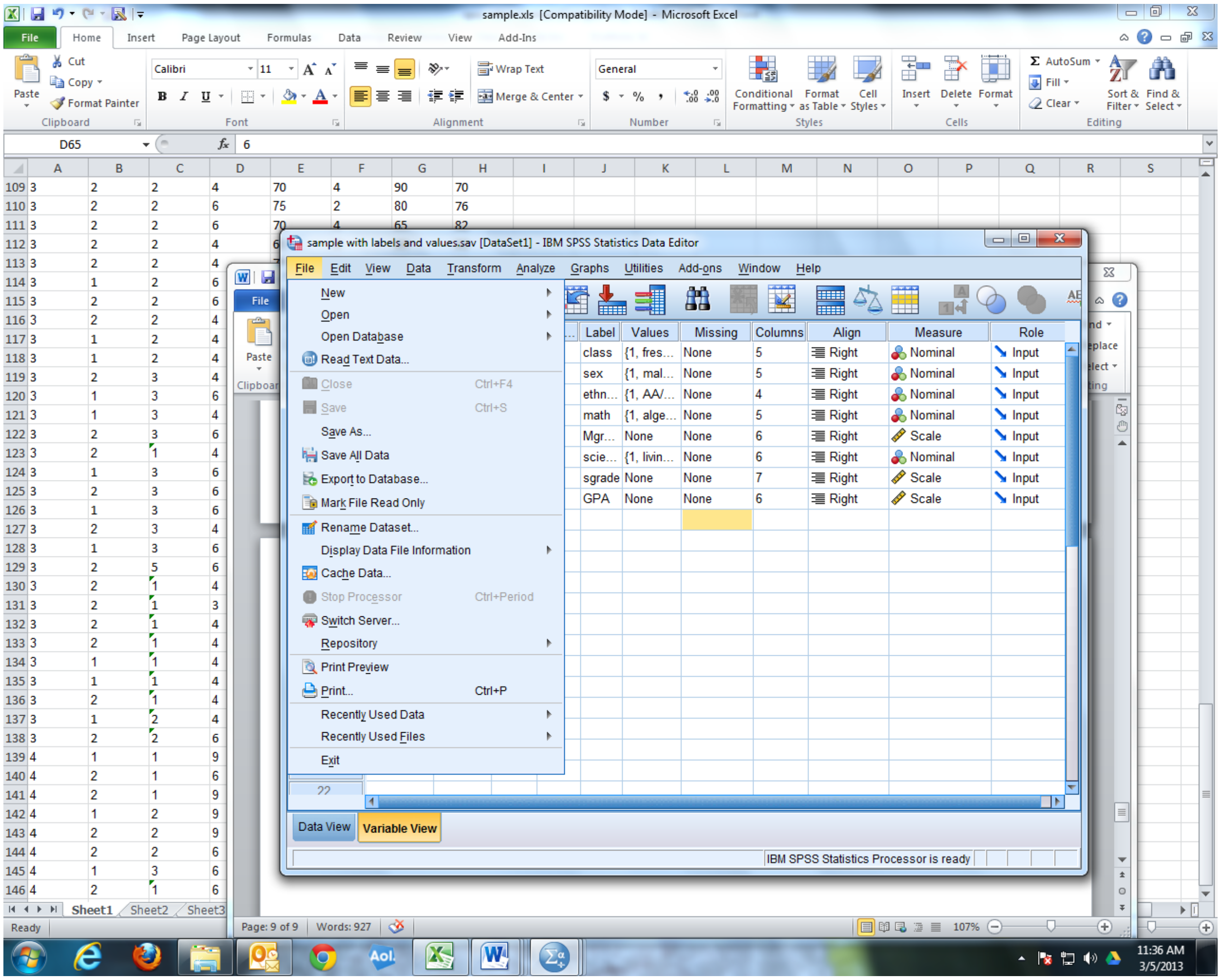
Sgrade as shown

GPA as shown

Once the labels and values are in the SPSS file you should save it. SPSS will save it as a .sav file so that it can be opened for further work in SPSS without having to re-label, etc.

The next step is to explore the data using the various descriptive tools afforded by SPSS.

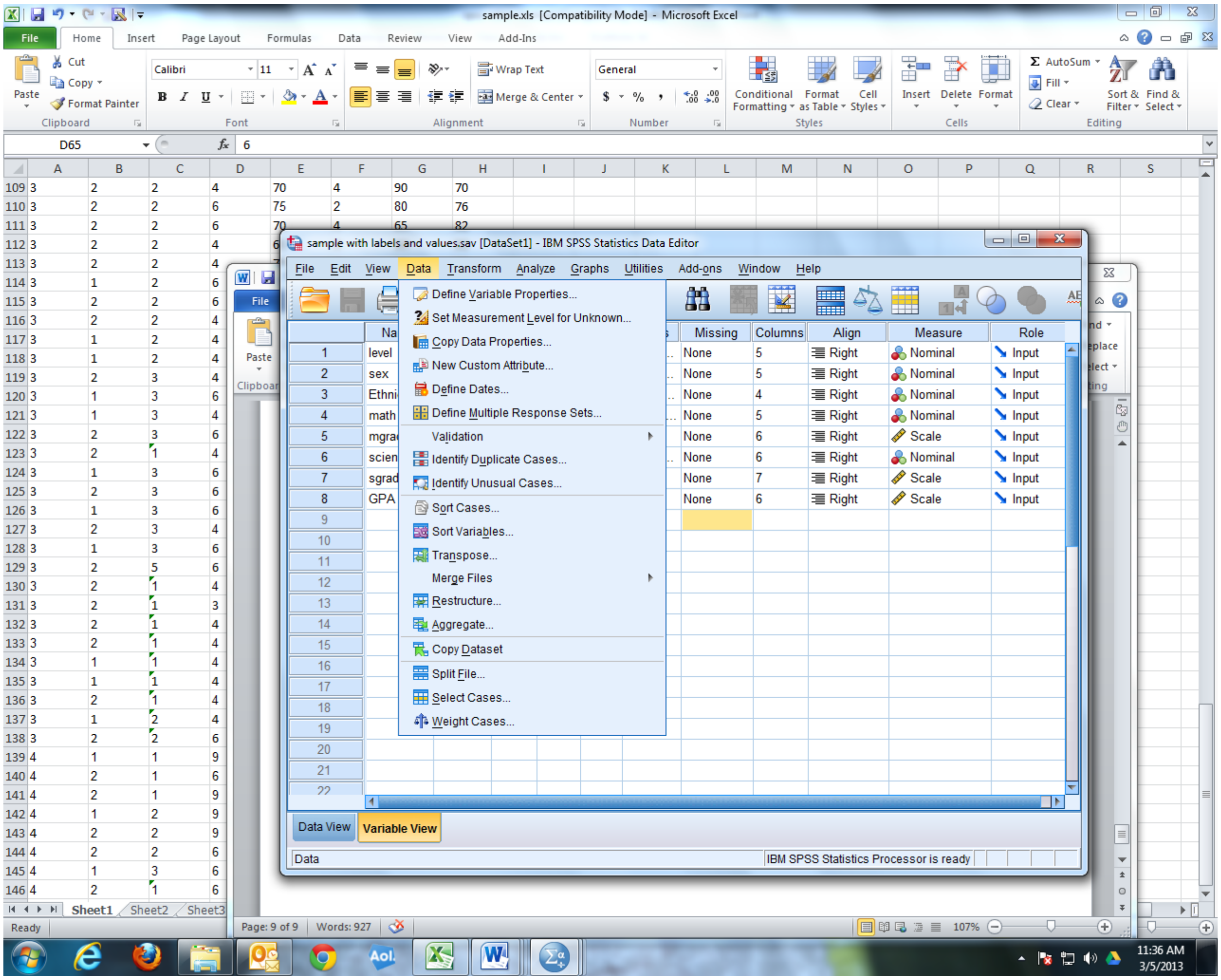
Let's look at some of the tabs: File



	A	B	C	D	E	F	G	H
109	3	2	2	4	70	4	90	70
110	3	2	2	6	75	2	80	76
111	3	2	2	6	70	4	65	82
112	3	2	2	4				
113	3	2	2	4				
114	3	1	2	6				
115	3	2	2	6				
116	3	2	2	4				
117	3	1	2	4				
118	3	1	2	4				
119	3	2	3	4				
120	3	1	3	6				
121	3	1	3	4				
122	3	2	3	6				
123	3	2	1	4				
124	3	1	3	6				
125	3	2	3	6				
126	3	1	3	6				
127	3	2	3	4				
128	3	1	3	6				
129	3	2	5	6				
130	3	2	1	4				
131	3	2	1	3				
132	3	2	1	4				
133	3	2	1	4				
134	3	1	1	4				
135	3	1	1	4				
136	3	2	1	4				
137	3	1	2	4				
138	3	2	2	6				
139	4	1	1	9				
140	4	2	1	6				
141	4	2	1	9				
142	4	1	2	9				
143	4	2	2	9				
144	4	2	2	6				
145	4	1	3	6				
146	4	2	1	6				

Label	Values	Missing	Columns	Align	Measure	Role
class	{1, fres...	None	5	Right	Nominal	Input
sex	{1, mal...	None	5	Right	Nominal	Input
ethn...	{1, AA/...	None	4	Right	Nominal	Input
math	{1, alge...	None	5	Right	Nominal	Input
Mgr...	None	None	6	Right	Scale	Input
scie...	{1, livin...	None	6	Right	Nominal	Input
sgrade	None	None	7	Right	Scale	Input
GPA	None	None	6	Right	Scale	Input

Data:

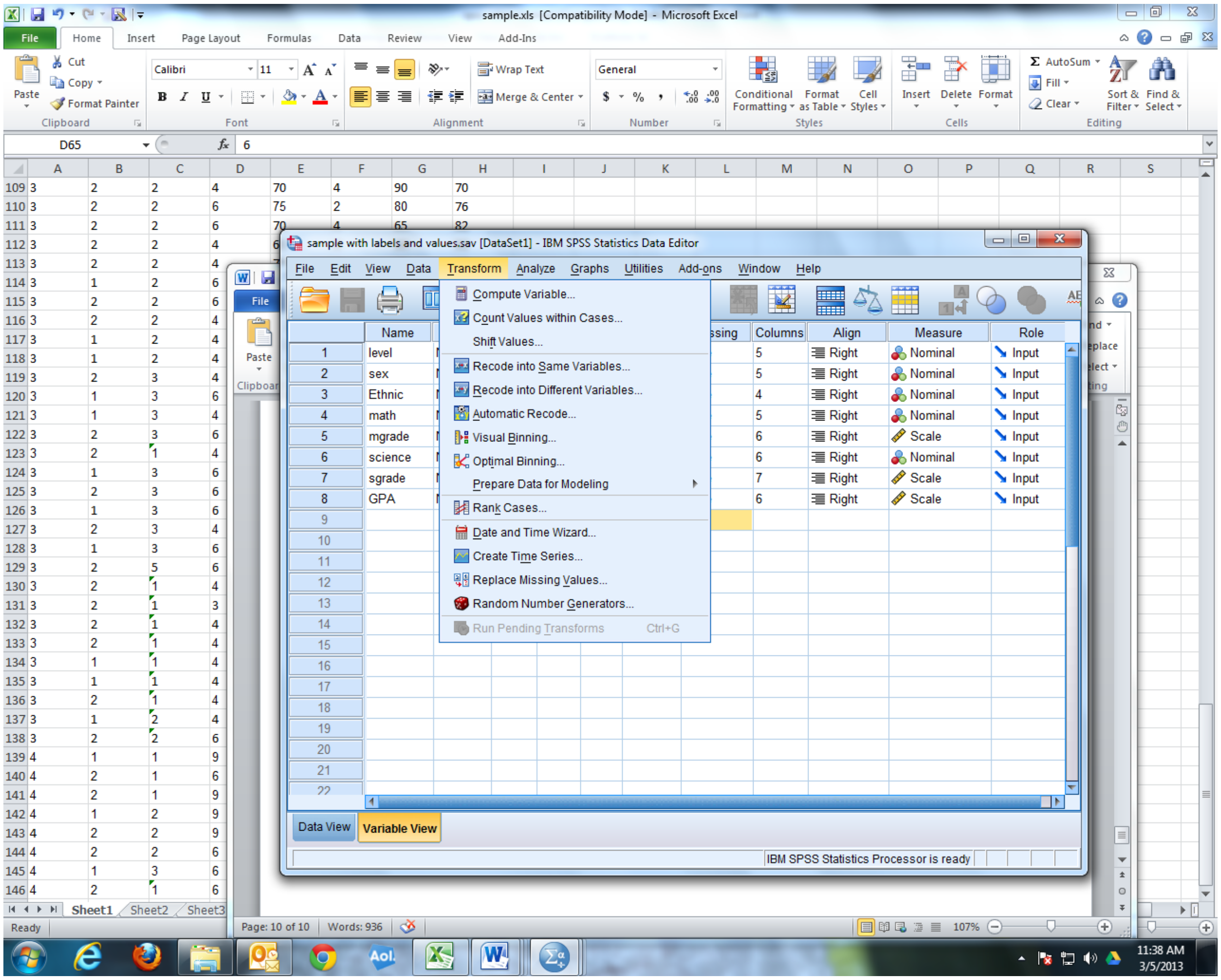


D65 6

	A	B	C	D	E	F	G	H
109	3	2	2	4	70	4	90	70
110	3	2	2	6	75	2	80	76
111	3	2	2	6	70	4	65	82
112	3	2	2	4	6			
113	3	2	2	4				
114	3	1	2	6				
115	3	2	2	6				
116	3	2	2	4				
117	3	1	2	4				
118	3	1	2	4				
119	3	2	3	4				
120	3	1	3	6				
121	3	1	3	4				
122	3	2	3	6				
123	3	2	1	4				
124	3	1	3	6				
125	3	2	3	6				
126	3	1	3	6				
127	3	2	3	4				
128	3	1	3	6				
129	3	2	5	6				
130	3	2	1	4				
131	3	2	1	3				
132	3	2	1	4				
133	3	2	1	4				
134	3	1	1	4				
135	3	1	1	4				
136	3	2	1	4				
137	3	1	2	4				
138	3	2	2	6				
139	4	1	1	9				
140	4	2	1	6				
141	4	2	1	9				
142	4	1	2	9				
143	4	2	2	9				
144	4	2	2	6				
145	4	1	3	6				
146	4	2	1	6				

	Na	Missing	Columns	Align	Measure	Role
1	level	None	5	Right	Nominal	Input
2	sex	None	5	Right	Nominal	Input
3	Ethni	None	4	Right	Nominal	Input
4	math	None	5	Right	Nominal	Input
5	mgra	None	6	Right	Scale	Input
6	scien	None	6	Right	Nominal	Input
7	sgrad	None	7	Right	Scale	Input
8	GPA	None	6	Right	Scale	Input

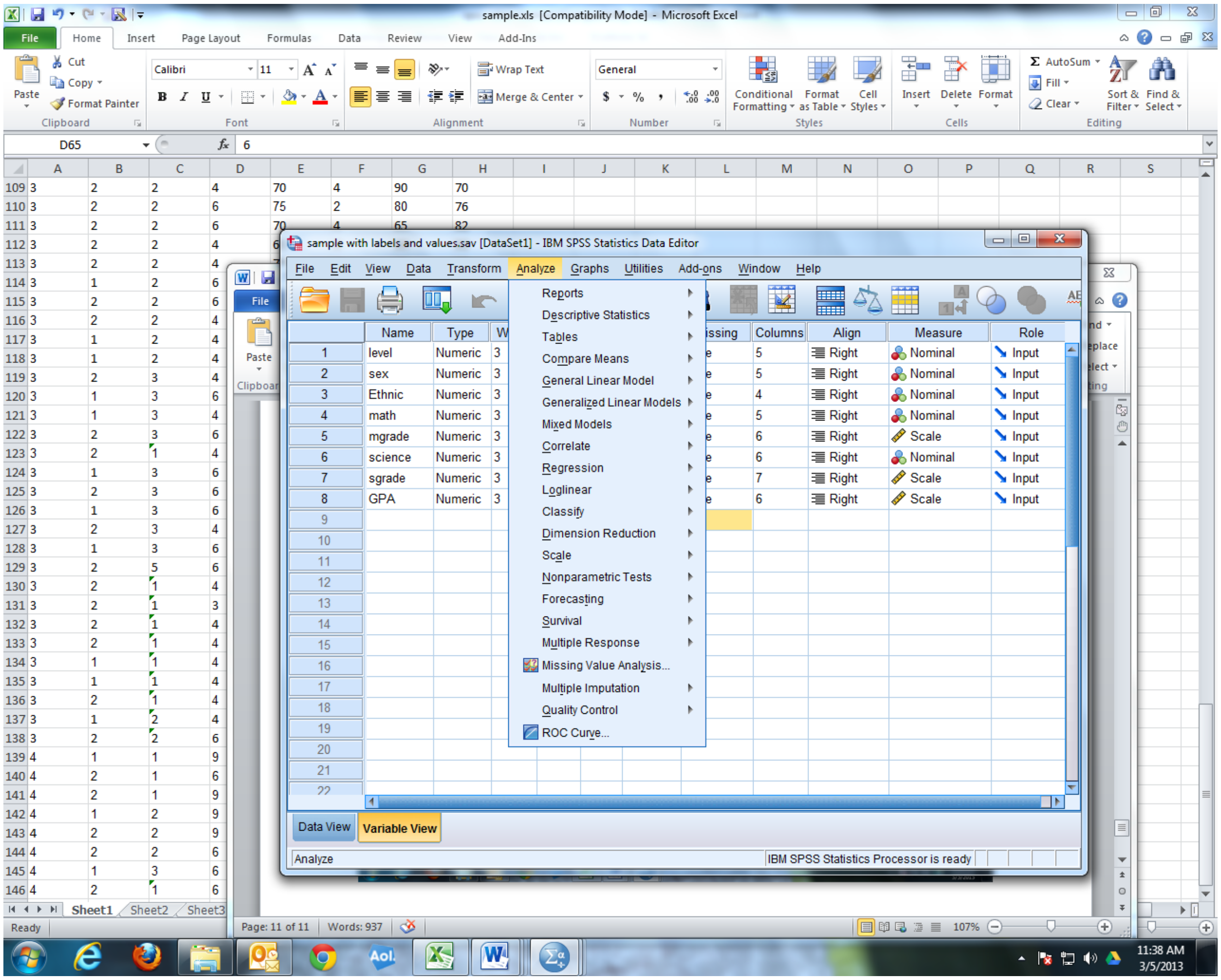
Transform:


sample.xls [Compatibility Mode] - Microsoft Excel
File
Home
Insert
Page Layout
Formulas
Data
Review
View
Add-Ins
D65
6
Sheet1
Sheet2
Sheet3
Ready
sample with labels and values.sav [DataSet1] - IBM SPSS Statistics Data Editor
File
Edit
View
Data
Transform
Analyze
Graphs
Utilities
Add-ons
Window
Help
Compute Variable...
Count Values within Cases...
Shift Values...
Recode into Same Variables...
Recode into Different Variables...
Automatic Recode...
Visual Binning...
Optimal Binning...
Prepare Data for Modeling
Rank Cases...
Date and Time Wizard...
Create Time Series...
Replace Missing Values...
Random Number Generators...
Run Pending Transforms Ctrl+G
Data View
Variable View
IBM SPSS Statistics Processor is ready
Page: 10 of 10
Words: 936
107%
11:38 AM
3/5/2013

	A	B	C	D	E	F	G	H
109	3	2	2	4	70	4	90	70
110	3	2	2	6	75	2	80	76
111	3	2	2	6	70	4	65	82
112	3	2	2	4				
113	3	2	2	4				
114	3	1	2	6				
115	3	2	2	6				
116	3	2	2	4				
117	3	1	2	4				
118	3	1	2	4				
119	3	2	3	4				
120	3	1	3	6				
121	3	1	3	4				
122	3	2	3	6				
123	3	2	1	4				
124	3	1	3	6				
125	3	2	3	6				
126	3	1	3	6				
127	3	2	3	4				
128	3	1	3	6				
129	3	2	5	6				
130	3	2	1	4				
131	3	2	1	3				
132	3	2	1	4				
133	3	2	1	4				
134	3	1	1	4				
135	3	1	1	4				
136	3	2	1	4				
137	3	1	2	4				
138	3	2	2	6				
139	4	1	1	9				
140	4	2	1	6				
141	4	2	1	9				
142	4	1	2	9				
143	4	2	2	9				
144	4	2	2	6				
145	4	1	3	6				
146	4	2	1	6				

	Name	Columns	Align	Measure	Role
1	level	5	Right	Nominal	Input
2	sex	5	Right	Nominal	Input
3	Ethnic	4	Right	Nominal	Input
4	math	5	Right	Nominal	Input
5	mgrade	6	Right	Scale	Input
6	science	6	Right	Nominal	Input
7	sgrade	7	Right	Scale	Input
8	GPA	6	Right	Scale	Input

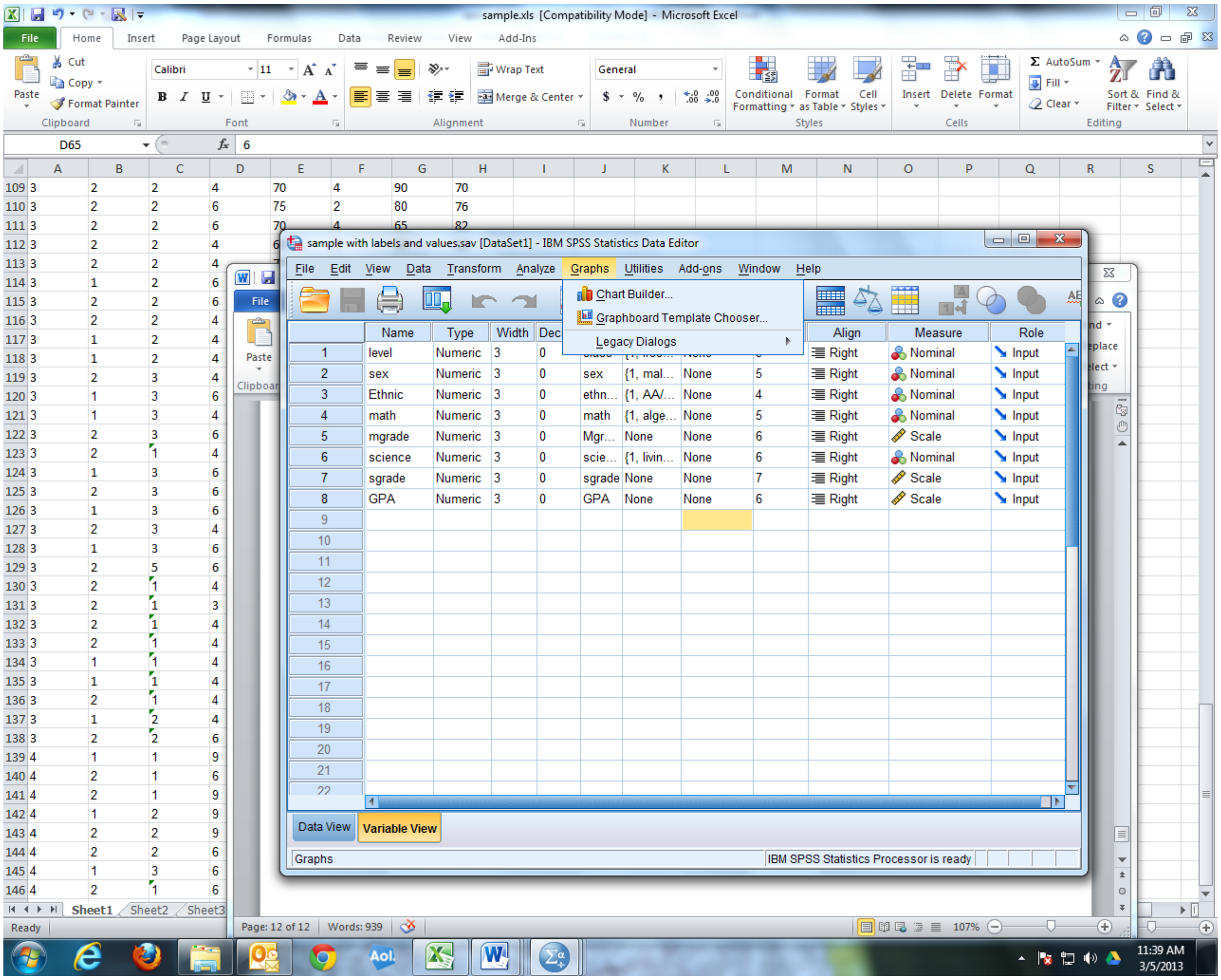
Analyze



	A	B	C	D	E	F	G	H
109	3	2	2	4	70	4	90	70
110	3	2	2	6	75	2	80	76
111	3	2	2	6	70	4	65	82
112	3	2	2	4	6			
113	3	2	2	4				
114	3	1	2	6				
115	3	2	2	6				
116	3	2	2	4				
117	3	1	2	4				
118	3	1	2	4				
119	3	2	3	4				
120	3	1	3	6				
121	3	1	3	4				
122	3	2	3	6				
123	3	2	1	4				
124	3	1	3	6				
125	3	2	3	6				
126	3	1	3	6				
127	3	2	3	4				
128	3	1	3	6				
129	3	2	5	6				
130	3	2	1	4				
131	3	2	1	3				
132	3	2	1	4				
133	3	2	1	4				
134	3	1	1	4				
135	3	1	1	4				
136	3	2	1	4				
137	3	1	2	4				
138	3	2	2	6				
139	4	1	1	9				
140	4	2	1	6				
141	4	2	1	9				
142	4	1	2	9				
143	4	2	2	9				
144	4	2	2	6				
145	4	1	3	6				
146	4	2	1	6				

	Name	Type	W	issing	Columns	Align	Measure	Role
1	level	Numeric	3	e	5	Right	Nominal	Input
2	sex	Numeric	3	e	5	Right	Nominal	Input
3	Ethnic	Numeric	3	e	4	Right	Nominal	Input
4	math	Numeric	3	e	5	Right	Nominal	Input
5	mgrade	Numeric	3	e	6	Right	Scale	Input
6	science	Numeric	3	e	6	Right	Nominal	Input
7	sgrade	Numeric	3	e	7	Right	Scale	Input
8	GPA	Numeric	3	e	6	Right	Scale	Input

Graphs:



D65 f_x 6

	A	B	C	D	E	F	G	H
109	3	2	2	4	70	4	90	70
110	3	2	2	6	75	2	80	76
111	3	2	2	6	70	4	65	82
112	3	2	2	4	6			
113	3	2	2	4				
114	3	1	2	6				
115	3	2	2	6				
116	3	2	2	4				
117	3	1	2	4				
118	3	1	2	4				
119	3	2	3	4				
120	3	1	3	6				
121	3	1	3	4				
122	3	2	3	6				
123	3	2	1	4				
124	3	1	3	6				
125	3	2	3	6				
126	3	1	3	6				
127	3	2	3	4				
128	3	1	3	6				
129	3	2	5	6				
130	3	2	1	4				
131	3	2	1	3				
132	3	2	1	4				
133	3	2	1	4				
134	3	1	1	4				
135	3	1	1	4				
136	3	2	1	4				
137	3	1	2	4				
138	3	2	2	6				
139	4	1	1	9				
140	4	2	1	6				
141	4	2	1	9				
142	4	1	2	9				
143	4	2	2	9				
144	4	2	2	6				
145	4	1	3	6				
146	4	2	1	6				

sample with labels and values.sav [DataSet1] - IBM SPSS Statistics Data Editor

	Name	Type	Width	Dec					Align	Measure	Role
1	level	Numeric	3	0					Right	Nominal	Input
2	sex	Numeric	3	0	sex	{1, mal...	None	5	Right	Nominal	Input
3	Ethnic	Numeric	3	0	ethn...	{1, AA/...	None	4	Right	Nominal	Input
4	math	Numeric	3	0	math	{1, alge...	None	5	Right	Nominal	Input
5	mgrade	Numeric	3	0	Mgr...	None	None	6	Right	Scale	Input
6	science	Numeric	3	0	scie...	{1, livin...	None	6	Right	Nominal	Input
7	sgrade	Numeric	3	0	sgrade	None	None	7	Right	Scale	Input
8	GPA	Numeric	3	0	GPA	None	None	6	Right	Scale	Input

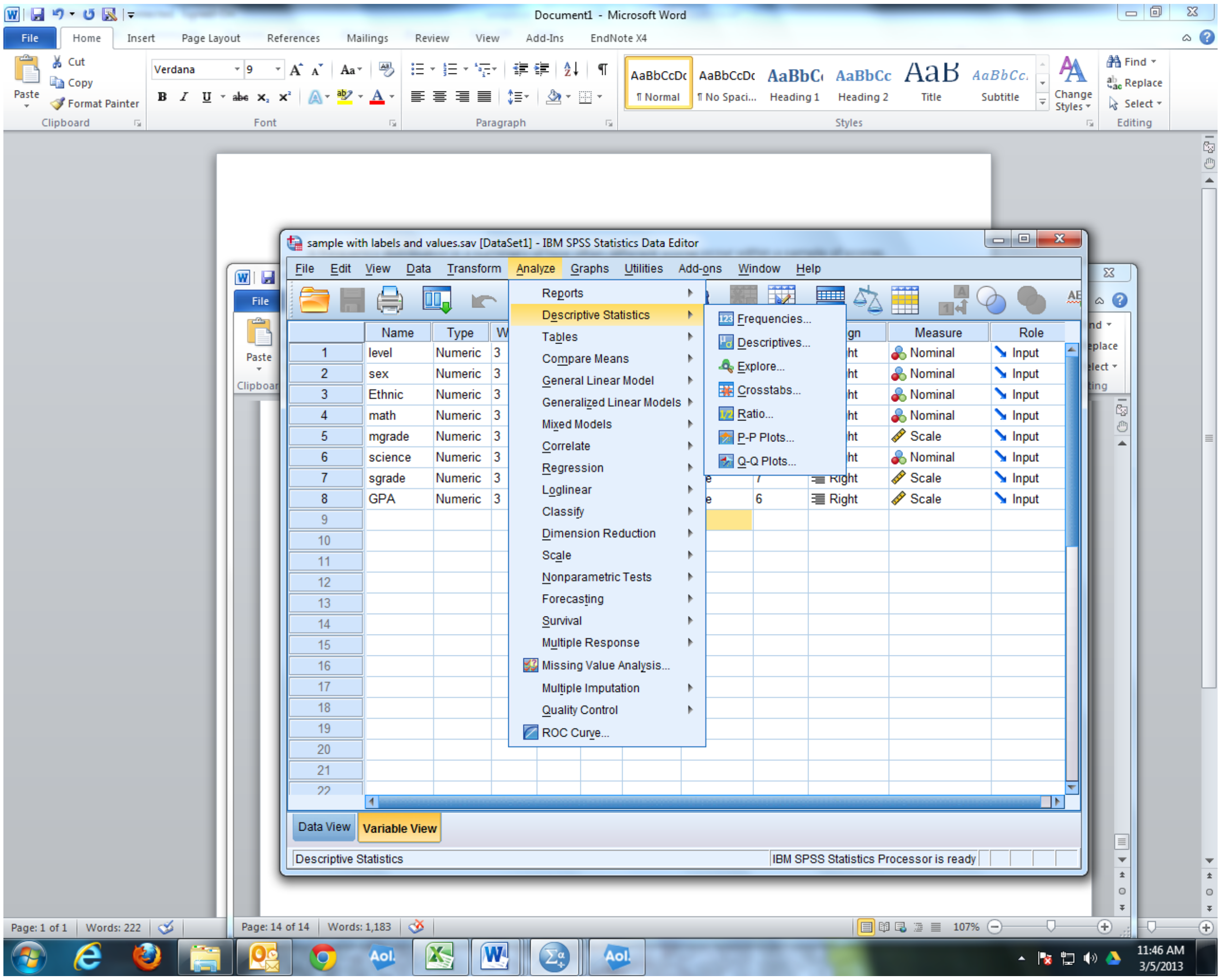
Graphs IBM SPSS Statistics Processor is ready

Ready Page: 12 of 12 Words: 939 107%

11:39 AM 3/5/2013

Using SPSS to create descriptions of your data
(Frequencies, Descriptions, Charts)

Under the Analyze Tab/Descriptive Statistics you will find: Frequencies, Descriptives, Explore, Crosstabs and Ratio.



	Name	Type	W	gn	Measure	Role
1	level	Numeric	3	ht	Nominal	Input
2	sex	Numeric	3	ht	Nominal	Input
3	Ethnic	Numeric	3	ht	Nominal	Input
4	math	Numeric	3	ht	Nominal	Input
5	mgrade	Numeric	3	ht	Scale	Input
6	science	Numeric	3	ht	Nominal	Input
7	sgrade	Numeric	3	Right	Scale	Input
8	GPA	Numeric	3	Right	Scale	Input

Frequencies

A frequency distribution is a summary of how often different scores occur within a sample of scores.

For example, let's suppose that you are collecting data on how many hours of sleep college students get each night. After conducting a survey of 30 of your classmates, you are left with the following set of scores:

7, 5, 8, 9, 4, 10, 7, 9, 9, 6, 5, 11, 6, 5, 9, 10, 8, 6, 9, 7, 9, 8, 4, 7, 8, 7, 6, 10, 4, 8

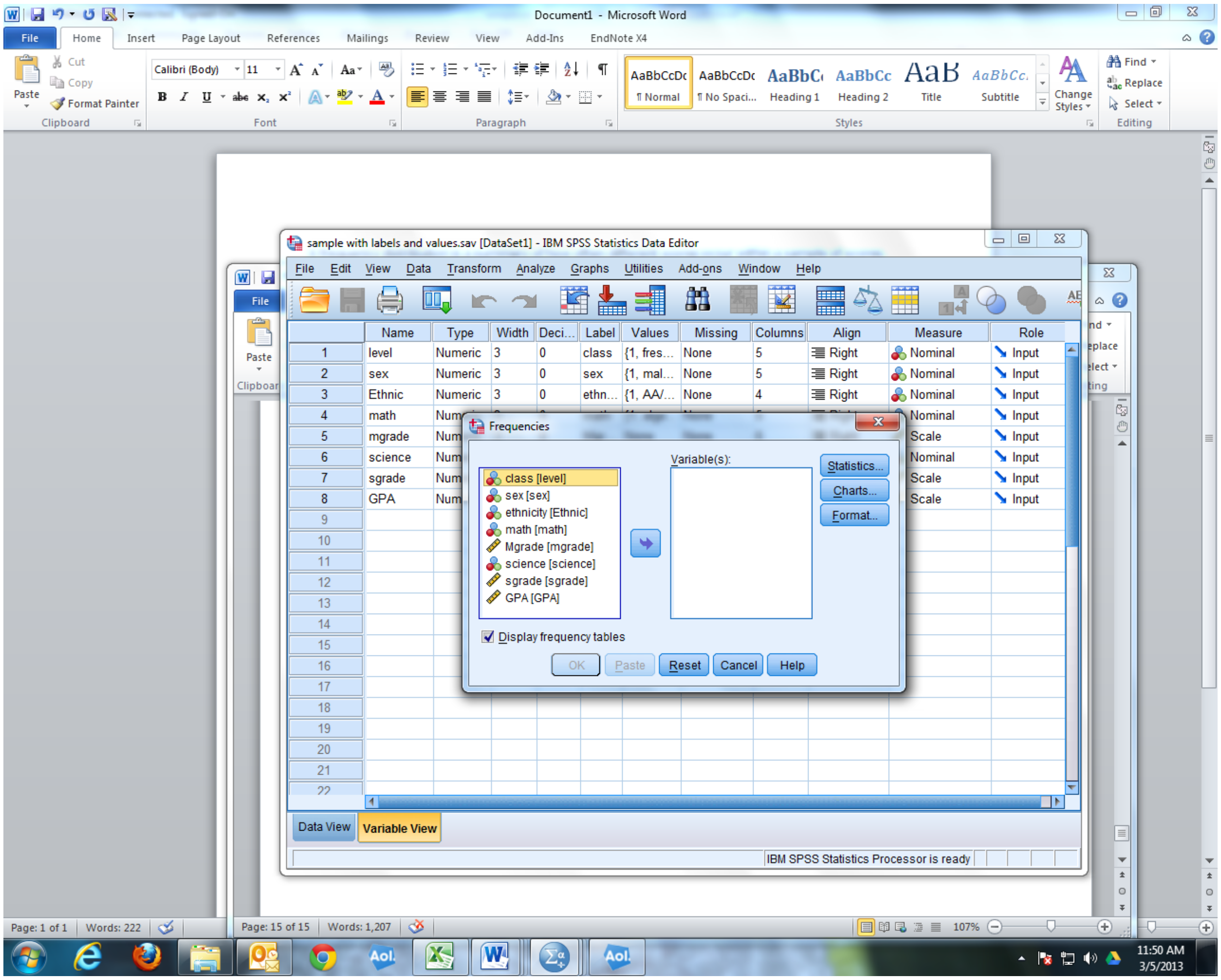
In order to make sense of this information, you need to find a way to organize the data. A frequency distribution is commonly used to categorize information so that it can be interpreted quickly in a visual way. In our example above, the number of hours each week serves as the categories and the occurrences of each number are then tallied.

Using the information from a frequency distribution, researchers can then calculate the mean, median, mode, range and standard deviation. Frequency distributions are often displayed in a table format (as you can see in the example found below), but they can also be presented graphically using a histogram.

Example of a Frequency Distribution

Hours of Sleep Each Night	Frequency
4	3
5	3
6	4
7	5
8	5
9	6
10	3
11	1
Total	30

In order to generate a frequency distribution for the data we are working with you select Analyze/Descriptives/Frequencies and will see the following selection box.


sample with labels and values.sav [DataSet1] - IBM SPSS Statistics Data Editor
File
Edit
View
Data
Transform
Analyze
Graphs
Utilities
Add-ons
Window
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready

	Name	Type	Width	Deci...	Label	Values	Missing	Columns	Align	Measure	Role
1	level	Numeric	3	0	class	{1, fres...	None	5	Right	Nominal	Input
2	sex	Numeric	3	0	sex	{1, mal...	None	5	Right	Nominal	Input
3	Ethnic	Numeric	3	0	ethn...	{1, AA/...	None	4	Right	Nominal	Input
4	math	Num								Nominal	Input
5	mgrade	Num								Scale	Input
6	science	Num								Nominal	Input
7	sgrade	Num								Scale	Input
8	GPA	Num								Scale	Input

Frequencies

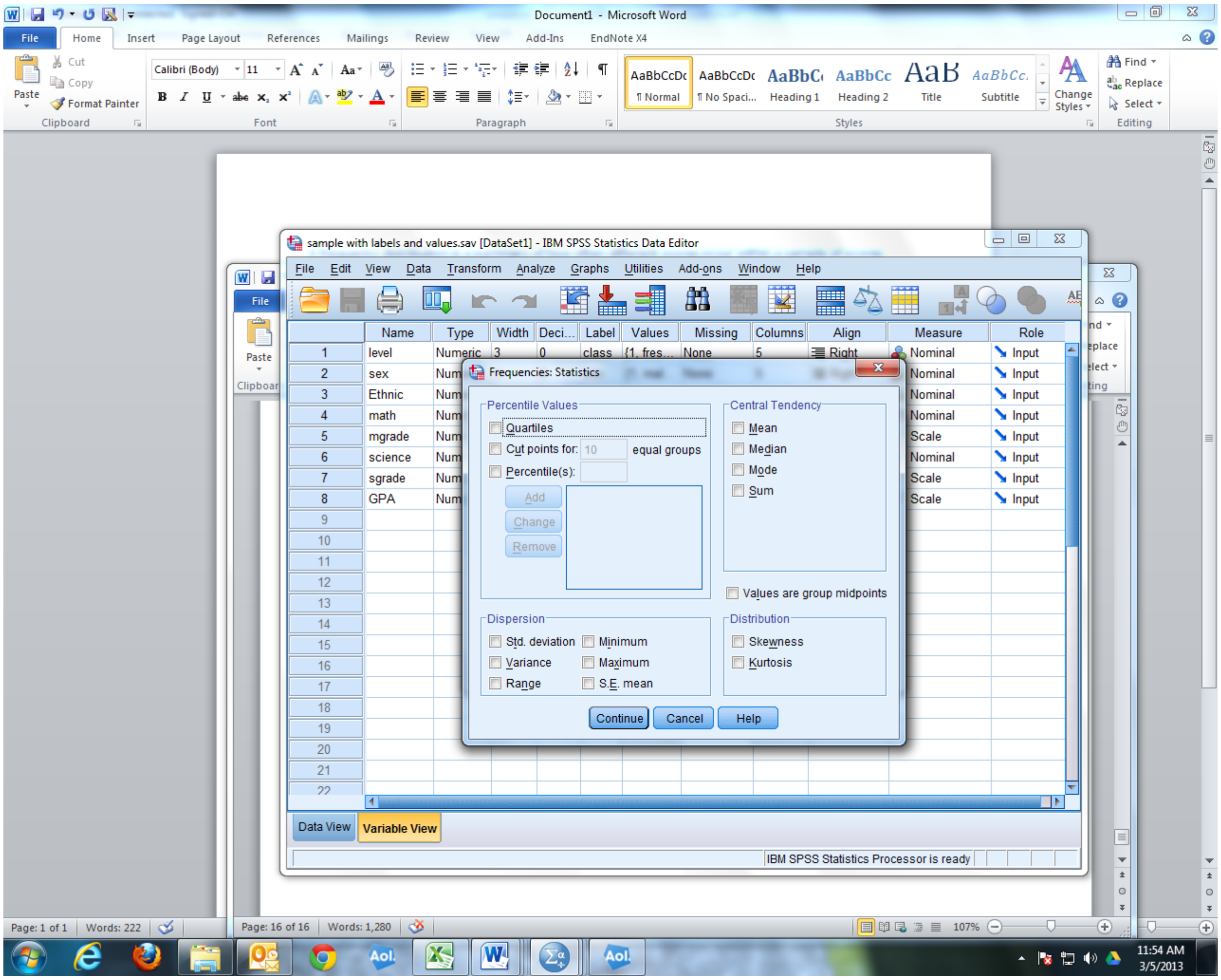
- class [level]
- sex [sex]
- ethnicity [Ethnic]
- math [math]
- Mgrade [mgrade]
- science [science]
- sgrade [sgrade]
- GPA [GPA]

Variable(s):

☑ Display frequency tables

This enables you to select any or all variables by highlighting those you wish to plot and using the arrow in the middle to move them to the right box for use. The program does not care if the variables are nominal or scaled. It will do the same process for all.

After you have moved the variables of interest over you can also select statistics



sample with labels and values.sav [DataSet1] - IBM SPSS Statistics Data Editor

	Name	Type	Width	Deci...	Label	Values	Missing	Columns	Align	Measure	Role
1	level	Numeric	3	0	class	{1, fres...	None	5	Right	Nominal	Input
2	sex	Num								Nominal	Input
3	Ethnic	Num								Nominal	Input
4	math	Num								Nominal	Input
5	mgrade	Num								Scale	Input
6	science	Num								Nominal	Input
7	sgrade	Num								Scale	Input
8	GPA	Num								Scale	Input

Frequencies: Statistics

Percentile Values
- □ Quartiles
- □ Cut points for: 10 equal groups
- □ Percentile(s):

Central Tendency
- □ Mean
- □ Median
- □ Mode
- □ Sum

□ Values are group midpoints

Dispersion
- □ Std. deviation
- □ Minimum
- □ Variance
- □ Maximum
- □ Range
- □ S.E. mean

Distribution
- □ Skewness
- □ Kurtosis

For some of the variables (scaled) these are relevant for the nominal variables the statistics are not relevant. So let’s not choose this now but choose Continue and go back to run the frequencies.

The output is on the next several pages

```
NEW FILE.
DATASET NAME DataSet1 WINDOW=FRONT.
GET
  FILE='F:\111SPSS workshops\sample with labels and values.sav'.
DATASET NAME DataSet2 WINDOW=FRONT.
FREQUENCIES VARIABLES=level sex Ethnic math mgrade science sgrade GPA
  /ORDER=ANALYSIS.
```

Frequencies

Notes

Output Created		05-Mar-2013 12:04:57
Comments		
Input	Data	F:\111SPSS workshops\sample with labels and values.sav
	Active Dataset	DataSet2
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	145
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics are based on all cases with valid data.
Syntax		FREQUENCIES VARIABLES=level sex Ethnic math mgrade science sgrade GPA /ORDER=ANALYSIS.
Resources	Processor Time	00 00:00:00.015
	Elapsed Time	00 00:00:00.077

[DataSet2] F:\111SPSS workshops\sample with labels and values.sav

Statistics

		class	sex	ethnicity	math	Mgrade	science
N	Valid	145	145	145	144	145	145
	Missing	0	0	0	1	0	0

Statistics

		sgrade	GPA
N	Valid	145	145
	Missing	0	0

Frequency Table

class

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	sophomore	48	33.1	33.1	33.1
	junior	89	61.4	61.4	94.5
	senior	8	5.5	5.5	100.0
	Total	145	100.0	100.0	

sex

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	male	53	36.6	36.6	36.6
	female	92	63.4	63.4	100.0
	Total	145	100.0	100.0	

ethnicity

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	AA/Black	60	41.4	41.4	41.4
	Latino/Hispanic	62	42.8	42.8	84.1
	White	19	13.1	13.1	97.2
	Asian	1	.7	.7	97.9
	Other	3	2.1	2.1	100.0
	Total	145	100.0	100.0	

math

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	algebra	2	1.4	1.4	1.4
	geometry	44	30.3	30.6	31.9
	Math A	2	1.4	1.4	33.3
	Math B	48	33.1	33.3	66.7
	Pre-calculus	42	29.0	29.2	95.8
	7	2	1.4	1.4	97.2
	9	4	2.8	2.8	100.0
	Total	144	99.3	100.0	
Missing	System	1	.7		
Total		145	100.0		

Mgrade

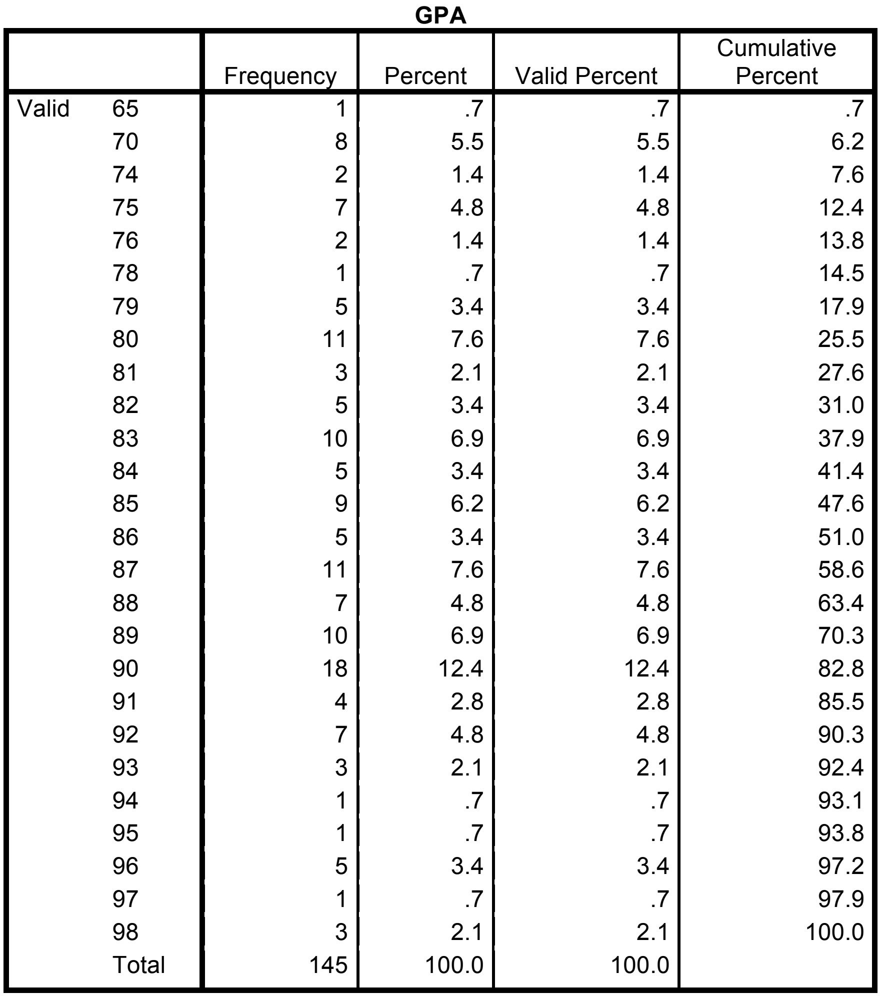
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	45	1	.7	.7	.7
	55	8	5.5	5.5	6.2
	65	20	13.8	13.8	20.0
	68	1	.7	.7	20.7
	70	12	8.3	8.3	29.0
	74	1	.7	.7	29.7
	75	14	9.7	9.7	39.3
	79	2	1.4	1.4	40.7
	80	16	11.0	11.0	51.7
	85	14	9.7	9.7	61.4
	87	1	.7	.7	62.1
	88	1	.7	.7	62.8
	89	4	2.8	2.8	65.5
	90	17	11.7	11.7	77.2
	91	1	.7	.7	77.9
	92	3	2.1	2.1	80.0
	93	2	1.4	1.4	81.4
	94	3	2.1	2.1	83.4
	95	2	1.4	1.4	84.8
	96	4	2.8	2.8	87.6
	97	1	.7	.7	88.3
	98	4	2.8	2.8	91.0
	99	3	2.1	2.1	93.1
	100	10	6.9	6.9	100.0
	Total	145	100.0	100.0	

science

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	living environment	29	20.0	20.0	20.0
	biology	47	32.4	32.4	52.4
	earth science	16	11.0	11.0	63.4
	Chemistry	52	35.9	35.9	99.3
	5	1	.7	.7	100.0
	Total	145	100.0	100.0	

sgrade

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	50	2	1.4	1.4	1.4
	55	2	1.4	1.4	2.8
	65	6	4.1	4.1	6.9
	68	1	.7	.7	7.6
	70	3	2.1	2.1	9.7
	74	1	.7	.7	10.3
	75	8	5.5	5.5	15.9
	77	1	.7	.7	16.6
	79	2	1.4	1.4	17.9
	80	21	14.5	14.5	32.4
	82	1	.7	.7	33.1
	84	1	.7	.7	33.8
	85	33	22.8	22.8	56.6
	88	3	2.1	2.1	58.6
	89	2	1.4	1.4	60.0
	90	26	17.9	17.9	77.9
	91	1	.7	.7	78.6
	92	4	2.8	2.8	81.4
	93	2	1.4	1.4	82.8
	94	1	.7	.7	83.4
	95	5	3.4	3.4	86.9
	96	3	2.1	2.1	89.0
	97	1	.7	.7	89.7
	98	6	4.1	4.1	93.8
	99	2	1.4	1.4	95.2
	100	7	4.8	4.8	100.0
	Total	145	100.0	100.0	



GPA

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	65	1	.7	.7	.7
	70	8	5.5	5.5	6.2
	74	2	1.4	1.4	7.6
	75	7	4.8	4.8	12.4
	76	2	1.4	1.4	13.8
	78	1	.7	.7	14.5
	79	5	3.4	3.4	17.9
	80	11	7.6	7.6	25.5
	81	3	2.1	2.1	27.6
	82	5	3.4	3.4	31.0
	83	10	6.9	6.9	37.9
	84	5	3.4	3.4	41.4
	85	9	6.2	6.2	47.6
	86	5	3.4	3.4	51.0
	87	11	7.6	7.6	58.6
	88	7	4.8	4.8	63.4
	89	10	6.9	6.9	70.3
	90	18	12.4	12.4	82.8
	91	4	2.8	2.8	85.5
	92	7	4.8	4.8	90.3
	93	3	2.1	2.1	92.4
	94	1	.7	.7	93.1
	95	1	.7	.7	93.8
	96	5	3.4	3.4	97.2
	97	1	.7	.7	97.9
	98	3	2.1	2.1	100.0
	Total	145	100.0	100.0	

Now it is clear that in cases where we have an extensive range of values it is difficult to interpret the output. Therefore you can recode the variables so that you get ranges. In the case of grades you can **recode** into categories of grades:

<65

65-74

75-84

85-94

95-100

This is done under the Transform tab. Since we may later want to use the scaled grades as they are we should recode into new variables. We will name these variables as follows:

Mgrade = mathgr
Sgrade = sciegr
GPA = average

Then we can rerun the data for the mathgr, sciegr and average and will get the following

```
FREQUENCIES VARIABLES=mathgr sciegr average
  /ORDER=ANALYSIS.
```

Frequencies

Notes

Output Created		05-Mar-2013 13:06:19
Comments		
Input	Data	F:\111SPSS workshops\sample with labels and values.sav
	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	145
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics are based on all cases with valid data.
Syntax		FREQUENCIES VARIABLES=mathgr sciegr average /ORDER=ANALYSIS.
Resources	Processor Time	00 00:00:00.000
	Elapsed Time	00 00:00:00.000

[DataSet1] F:\111SPSS workshops\sample with labels and values.sav

Statistics

		mathgr	sciegr	average
N	Valid	145	145	145
	Missing	0	0	0

Frequency Table

mathgr

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	9	6.2	6.2	6.2
	2	34	23.4	23.4	29.7
	3	32	22.1	22.1	51.7
	4	46	31.7	31.7	83.4
	5	24	16.6	16.6	100.0
	Total	145	100.0	100.0	

sciegr

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	4	2.8	2.8	2.8
	2	11	7.6	7.6	10.3
	3	34	23.4	23.4	33.8
	4	72	49.7	49.7	83.4
	5	24	16.6	16.6	100.0
	Total	145	100.0	100.0	

average

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	2	11	7.6	7.6	7.6
	3	49	33.8	33.8	41.4
	4	75	51.7	51.7	93.1
	5	10	6.9	6.9	100.0
	Total	145	100.0	100.0	

Descriptives

Once your data is in SPSS you may want to be able to characterize the set as to central tendency (including the mean, median and mode) and variability (standard deviation, range).

Central Tendency

MEAN. Score that represents the balance point of the distribution. The sum of all scores divided by the number of scores.

MEDIAN. The middle score in the distribution; 50% of scores fall above this score and 50% fall below.

MODE. The mode is the score which appears the most frequently. Distributions may be bi- or tri-modal.

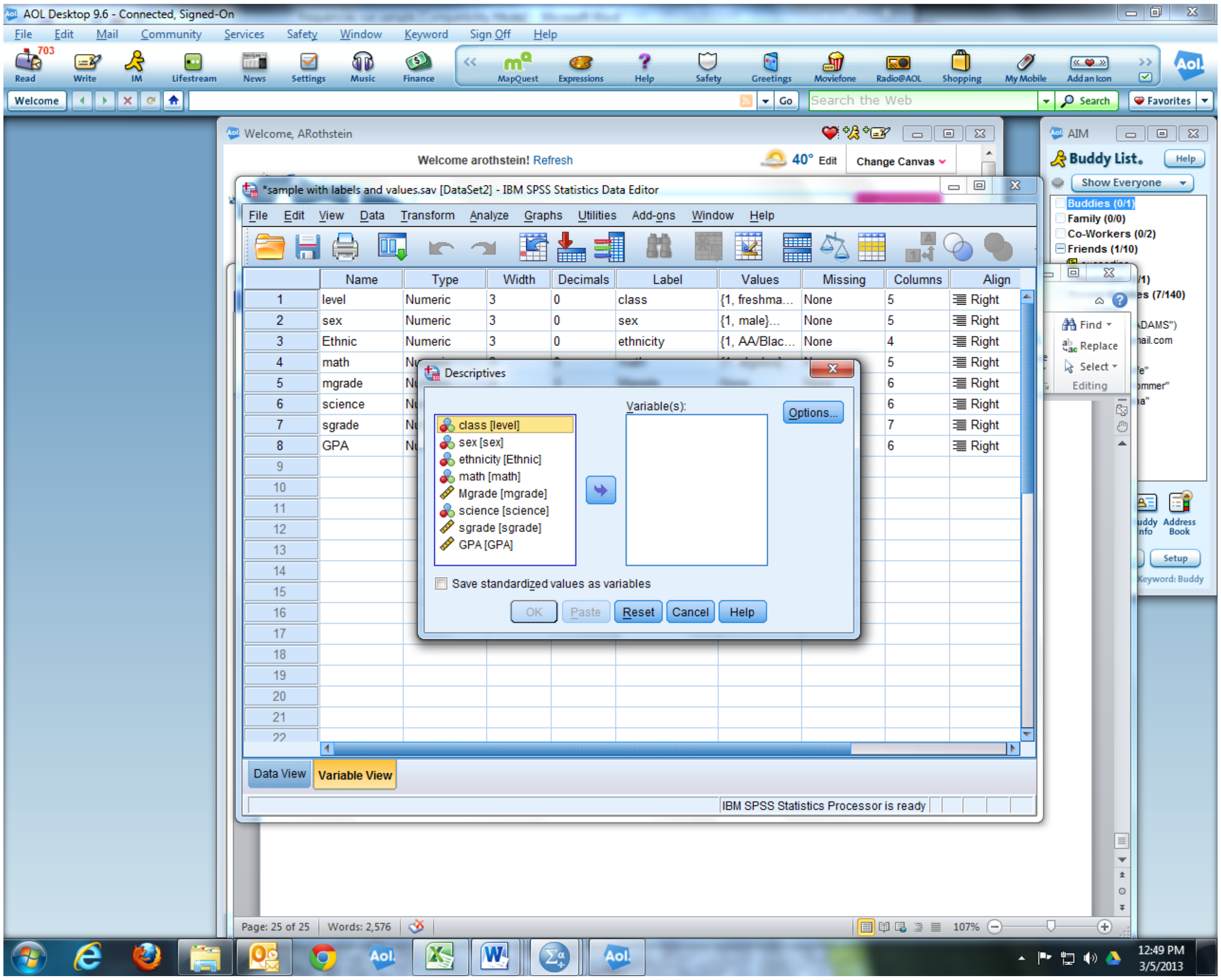
Variability

RANGE. The range is the total distance in scores from the highest to the lowest score. It can be affected by spurious scores and outliers; unusually low or high scores.

INTERQUARTILE RANGE. The IQ provides a more stable measure of the variability by using the 75% and 25% percentile scores to determine the range.

VARIANCE AND STANDARD DEVIATION. The most widely used and respected measures of variability or spread. They are based on squared deviation scores for the differences between each score and the mean. Variance and Standard Deviation are central to many inferential statistical tests. In some cases where means are too close to reveal significant differences variability among groups can be exceedingly high and lead to interesting conclusions.

The descriptives can be found at Analyze/Descriptive Statistics/Descriptives. This opens a window that looks much like that for the Frequencies. Choose and move over the three variables: mgrade, sgrade and GPA that are scaled variables and then click on options to choose the output you want.


AOL Desktop 9.6 - Connected, Signed-On
Welcome, ARothstein
Welcome arothstein! Refresh
*sample with labels and values.sav [DataSet2] - IBM SPSS Statistics Data Editor
File
Edit
View
Data
Transform
Analyze
Graphs
Utilities
Add-ons
Window
Help
Data View
Variable View
IBM SPSS Statistics Processor is ready

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align
1	level	Numeric	3	0	class	{1, freshma...	None	5	Right
2	sex	Numeric	3	0	sex	{1, male}...	None	5	Right
3	Ethnic	Numeric	3	0	ethnicity	{1, AA/Blac...	None	4	Right
4	math	Nu						5	Right
5	mgrade	Nu						6	Right
6	science	Nu						6	Right
7	sgrade	Nu						7	Right
8	GPA	Nu						6	Right

Descriptives

- class [level]
- sex [sex]
- ethnicity [Ethnic]
- math [math]
- Mgrade [mgrade]
- science [science]
- sgrade [sgrade]
- GPA [GPA]

Variable(s):

Options...

□ Save standardized values as variables

OK Paste Reset Cancel Help

```
DESCRIPTIVES VARIABLES=mgrade sgrade GPA
  /STATISTICS=MEAN STDDEV MIN MAX.
```

Descriptives

Notes

Output Created		05-Mar-2013 13:09:32
Comments		
Input	Data	F:\111SPSS workshops\sample with labels and values.sav
	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	145
Missing Value Handling	Definition of Missing	User defined missing values are treated as missing.
	Cases Used	All non-missing data are used.
Syntax		DESCRIPTIVES VARIABLES=mgrade sgrade GPA /STATISTICS=MEAN STDDEV MIN MAX.
Resources	Processor Time	00 00:00:00.000
	Elapsed Time	00 00:00:00.000

```
[DataSet1] F:\111SPSS workshops\sample with labels and values.sav
```

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Mgrade	145	45	100	80.70	13.042
sgrade	145	50	100	84.86	9.980
GPA	145	65	98	85.04	6.864
Valid N (listwise)	145				