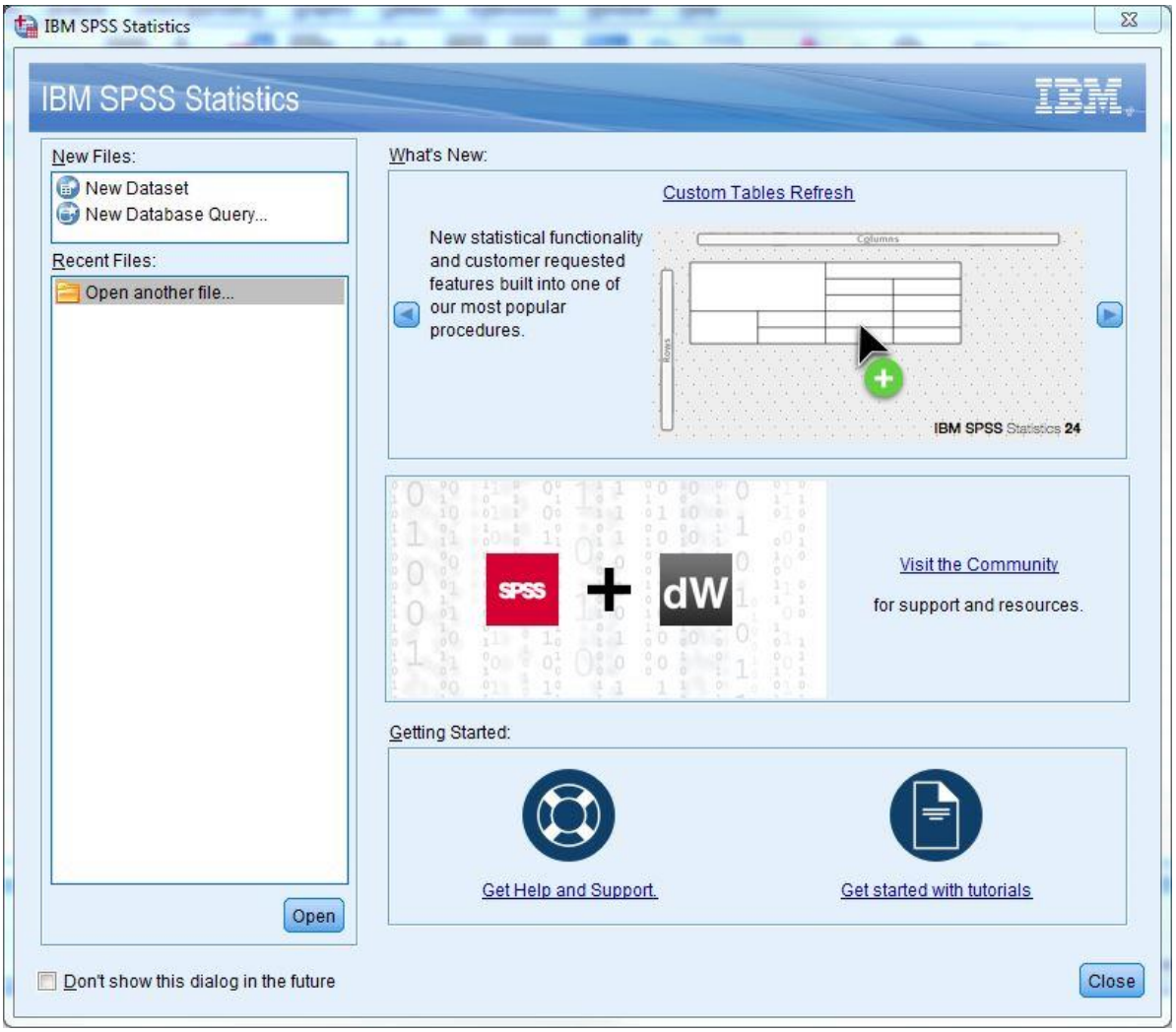
Files: http://www.wiu.edu/CITR/SPSS/SPSS_Workshop.zip

Start -> All Programs -> IBM SPSS Statistics -> IBM SPSS Statistics 24



SPSS opens with options to either:

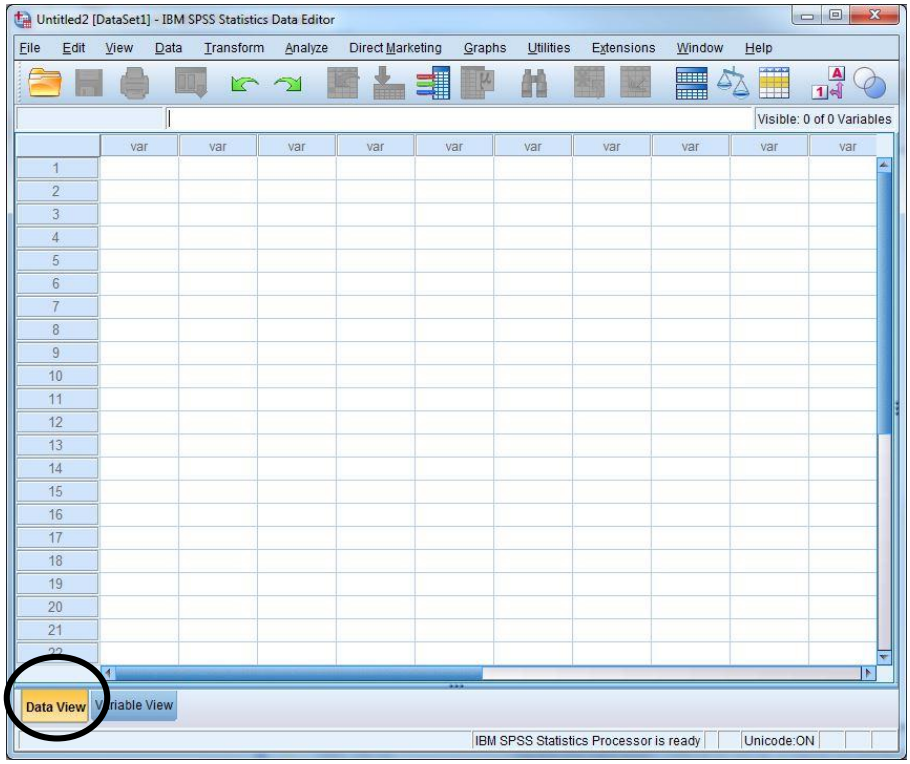
1. Enter new data
2. Open an existing data file

(both options are also available in normal screens)

Views

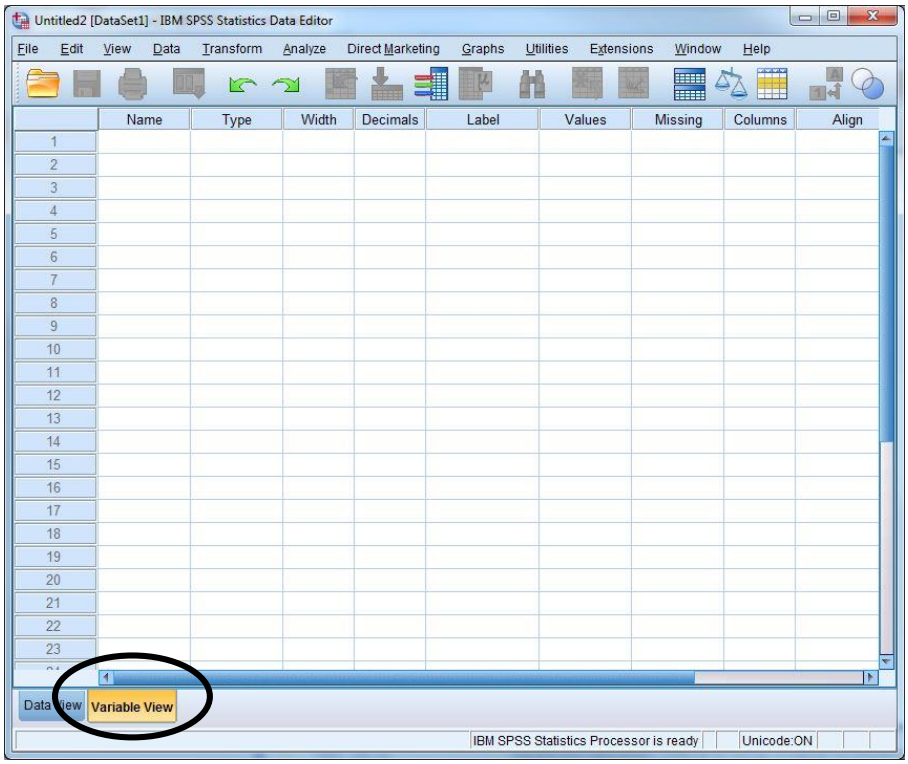
Data View:

- Used to enter data from observations
- Each row is a separate case
- Can have multiple measures (**variables**) for each case (e.g. height and weight)



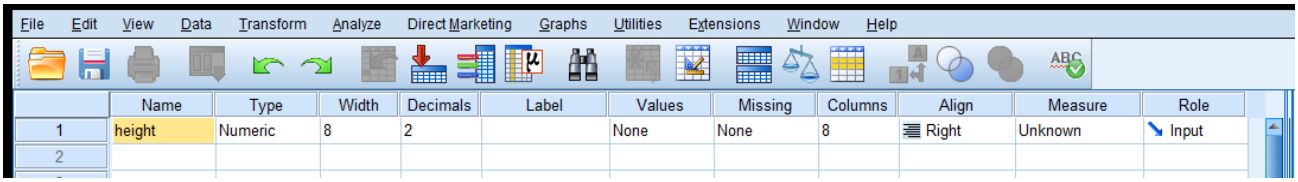
Variable View:

- Used to enter information about each variable
- Rows are separate variables, columns are information about the variables



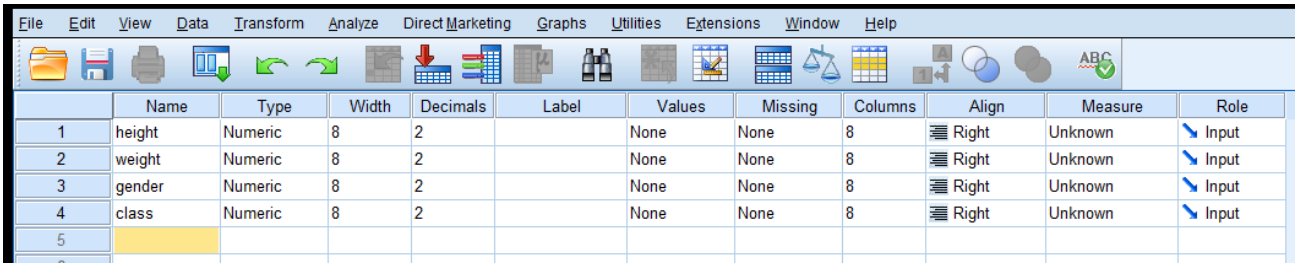
Entering Variable Information

- In the variable view, click in the *Name* column on an empty row
- Type the name of the variable
- Additional information will automatically be entered



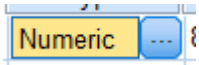
	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	height	Numeric	8	2		None	None	8	Right	Unknown	Input
2											

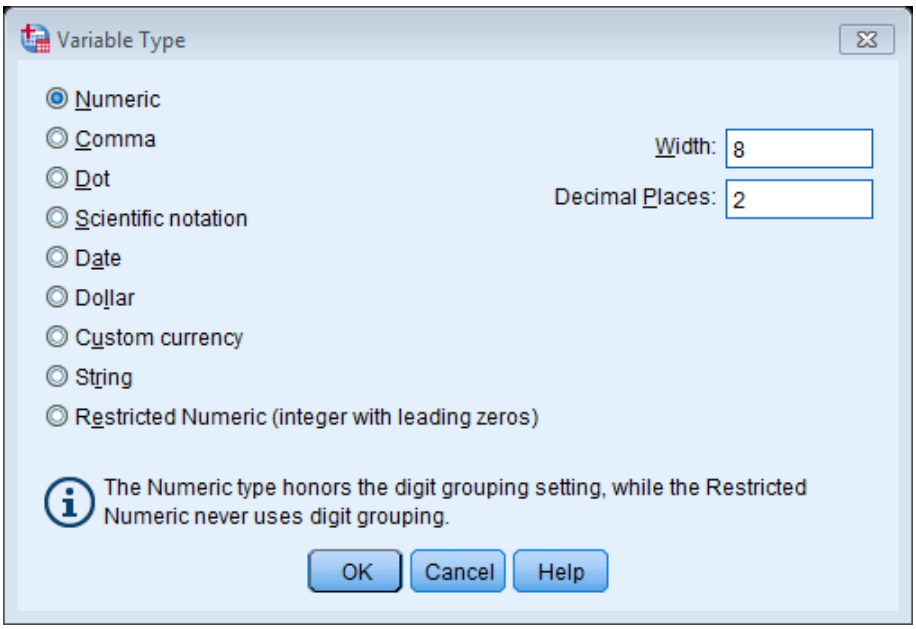
- Enter additional variables on empty rows



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	height	Numeric	8	2		None	None	8	Right	Unknown	Input
2	weight	Numeric	8	2		None	None	8	Right	Unknown	Input
3	gender	Numeric	8	2		None	None	8	Right	Unknown	Input
4	class	Numeric	8	2		None	None	8	Right	Unknown	Input
5											

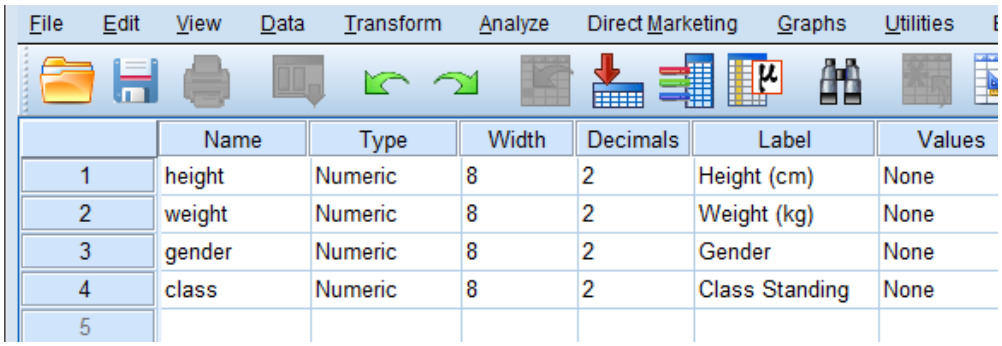
- Change the default parameters for each variable
- *Type*:
  - Click in the type box, then press the ... button





Entering Variable Information

- *Width, Decimals*:
  - Control number of digits (or string length)
- *Label*
  - **Very** useful – provide a descriptive name for your variable



	Name	Type	Width	Decimals	Label	Values
1	height	Numeric	8	2	Height (cm)	None
2	weight	Numeric	8	2	Weight (kg)	None
3	gender	Numeric	8	2	Gender	None
4	class	Numeric	8	2	Class Standing	None
5						

  - All variables have *names* (may default to something like "VAR0001"; a *label* is optional
  - *Labels* are useful for output, e.g. graphs and tables will have the more descriptive *label* rather than the *name*

- *Values*
  - The actual data you enter may be in codes or short strings. Adding *values* allows you to provide more descriptive terms for coded categories.
  - (e.g.) use 1 for male, 2 for female in the **data**, but we want our **output** to be "Male" and "Female"
  - Click the ... button in the *Values* column
  - (continued ….)

Codes for Categories

- Often data includes categories (as opposed to normal numbers)
- You can use codes to indicate different categories ….

Category	One possible code	Another possible code
Male	1	M
Female	2	F

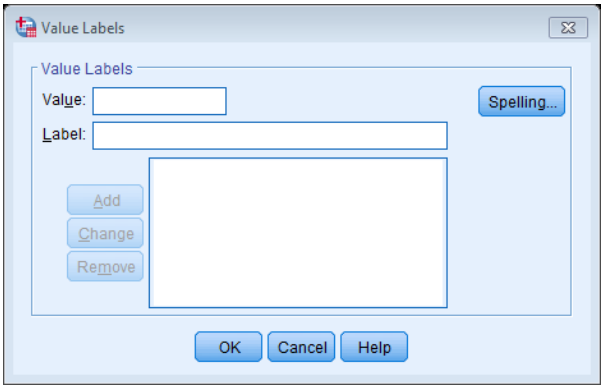
Category	One possible code	Another possible code
Freshman	1	FR
Sophomore	2	SO
Junior	3	JU
Senior	4	SE

- Why use codes?
  - May be using data from another program that uses codes (e.g. Scantron, Qualtrics survey software)
  - Recording device outputs codes
  - When collecting manually, just faster to write codes

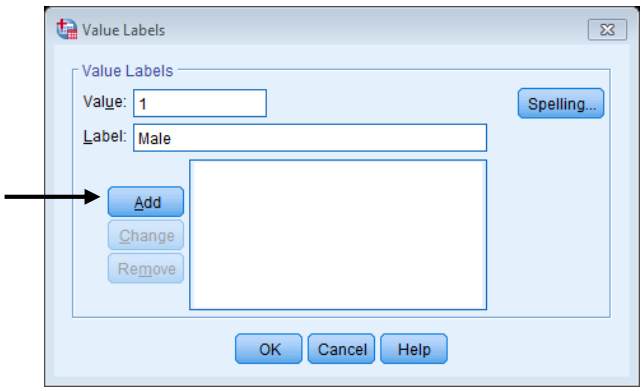
- For SPSS, enough to just have codes that are different from one another
- For *you*, may want to have meaningful labels to go along with the codes ….

Entering Variable Information

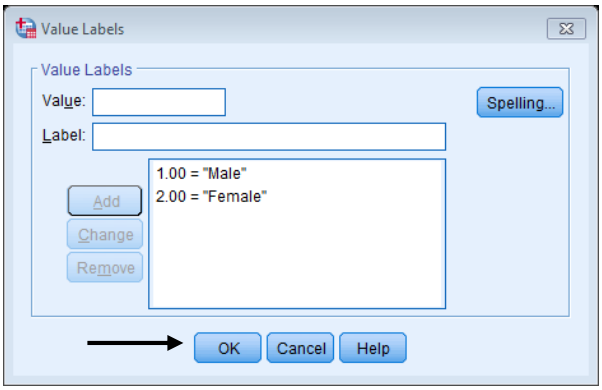
- *Values* (continued)
  - Add labels for the **gender variable**



  - Enter the number or string code in the V*alue* box
  - Enter the descriptive label in the *Label* box
  - Click *Add*



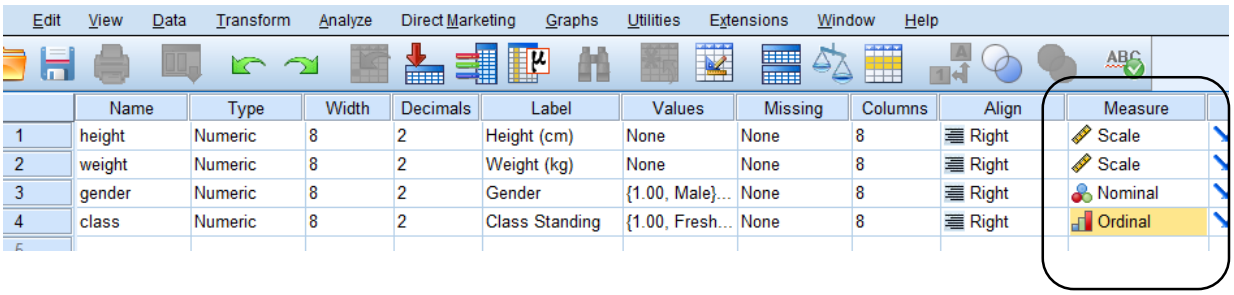
  - Repeat for all possible values, then click OK



  - You can change these later (add more, remove, change label)

Entering Variable Information

- *Missing*
  - You can define specific codes or ranges (e.g. 100 to 105) to indicate missing values. These won't be included in most calculations (e.g. enter 100 for *class* if we forgot to record it)
- *Columns, Align*
  - Affect display
- *Measure*
  - **Very important! SPSS needs to know the *scale of measurement* for each variable**
  - **(very quickly….)**
  - **Scale:** continuous values; numbers you can add and subtract; often have "meaningful zeros"
    - *height* and *weight*: can put on a number line, zero means something, can do math (e.g. 180cm x 2 = 360cm)
  - **Ordinal**: ranks; values can be put in order, but only know "greater than" and "less than"
    - *class:* "Freshman" < "Sophomore" < "Junior"
    - (note that if we used *credit hours* it would be **scale)**
  - **Nominal**
    - categories; letter or number codes that just define *groups*
    - *gender*: 1 = "Male", 2 = "Female" (note: no order)

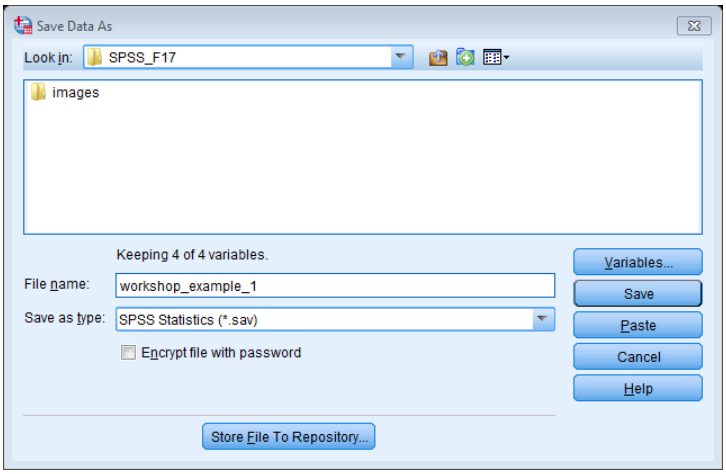


	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	height	Numeric	8	2	Height (cm)	None	None	8	Right	Scale
2	weight	Numeric	8	2	Weight (kg)	None	None	8	Right	Scale
3	gender	Numeric	8	2	Gender	{1.00, Male}...	None	8	Right	Nominal
4	class	Numeric	8	2	Class Standing	{1.00, Fresh...	None	8	Right	Ordinal

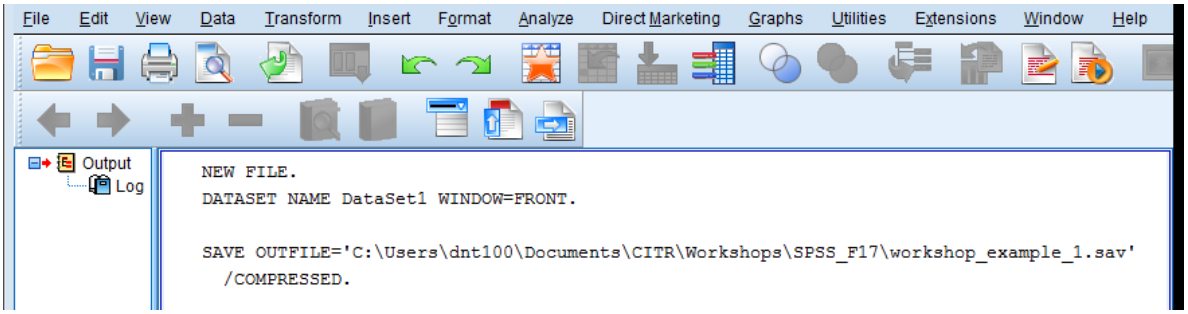
- *Role*: can specify what the variable "does" (e.g. independent)

The Output File

- Almost everything you do in SPSS will be pasted into the Output:
  - Tables and graphs from analysis
  - *Syntax from commands
- If you haven't already, save the contents of the Data Editor …



  - Note that doing this adds text to the Output window (in this case, the *syntax* of the *Save* command)

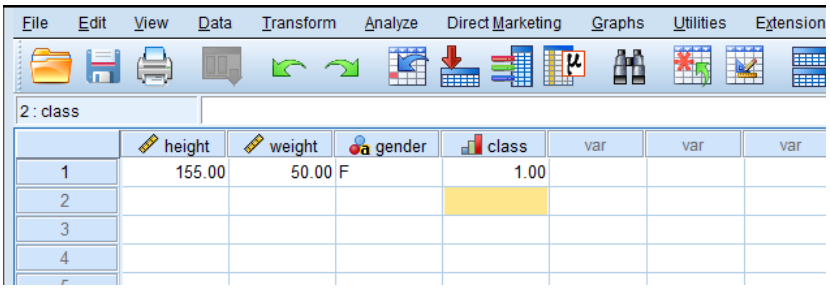


- The Output is treated like a document, e.g. SPSS will ask you if you want to save it if you haven't already and try to close the program.
- You don't necessarily need to save it, though this is where the output of your analyses will go!

*SPSS can be programmed to do certain tasks, e.g. running specific analysis, opening, saving, and closing files, etc. This can be useful if you want to repeat the same task, and some things that SPSS does can only be accomplished by programming. You can copy and paste syntax from the output into a syntax file, run it, and it will (usually!) do the same thing you did when you used the menus, drop-downs, and dialog boxes.

Entering Data

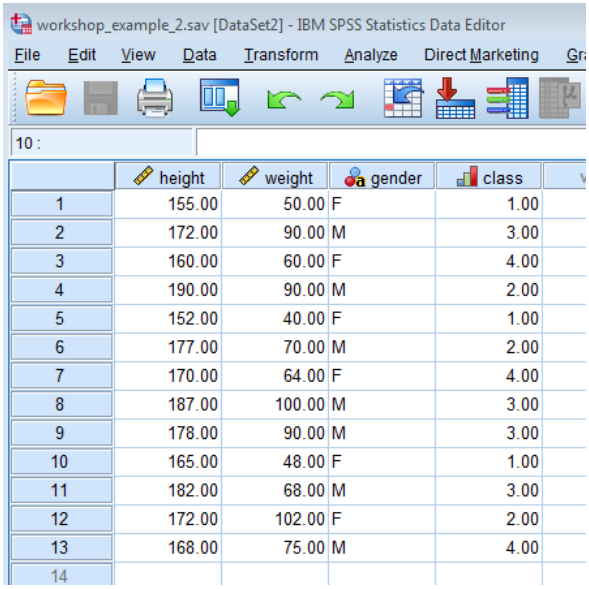
- To make things a bit easier to understand in the data view, change **gender** *Type* to **String**; for *Values,* change “Males” to “M”, “Females” to “F”
  - Note that it may make sense to keep number codes in some situations, e.g. opening data from a survey that uses numbers
  - (May see warning about existing labels – need to go back to *Values* and change 1=Male, 2=Female to M=Male, F=Female)
- Use the tab at the bottom to switch to *Data View*
- Enter your data; **each row is a case, and there may be multiple observations / variables per row**



- The data for the first row has been entered. All observed values for this case have been entered (note that in this example, each case represents a person)
  - Note that we don’t enter units (cm or kg) for our scale variables!
- Fill in the rest of the data ….

Entering Data

Subject	Height (cm)	Weight (kg)	Gender	Class Standing
1	155	50	female	freshman
2	172	90	male	junior
3	160	60	female	senior
4	190	90	male	sophomore
5	152	40	female	freshman
6	177	70	male	sophomore
7	170	64	female	senior
8	187	100	male	junior
9	178	90	male	junior
10	165	48	female	freshman
11	182	68	male	junior
12	172	102	female	sophomore
13	168	75	male	senior



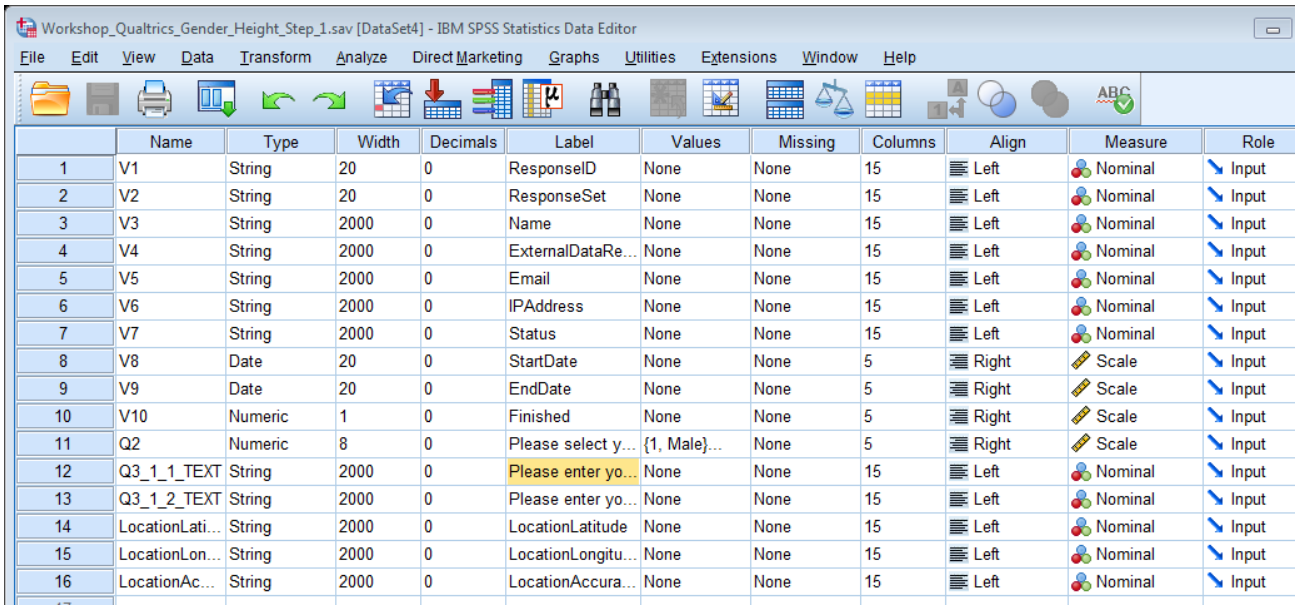
	height	weight	gender	class
1	155.00	50.00	F	1.00
2	172.00	90.00	M	3.00
3	160.00	60.00	F	4.00
4	190.00	90.00	M	2.00
5	152.00	40.00	F	1.00
6	177.00	70.00	M	2.00
7	170.00	64.00	F	4.00
8	187.00	100.00	M	3.00
9	178.00	90.00	M	3.00
10	165.00	48.00	F	1.00
11	182.00	68.00	M	3.00
12	172.00	102.00	F	2.00
13	168.00	75.00	M	4.00
14				

- Remember we are using *codes*, e.g. “freshman” = 1, “male” = “M”
- Note that you don’t *have* to put subject number as a variable, but you may want to if you have subjects (or observations in general) coded

SPSS file: workshop_example_2.sav

Other Sources of Data

- You can import data from other sources (rather than type it in)
- *File -> Import*
  - e.g. Excel
    - Put cases in rows, variables in columns
    - If you put variable names in the top row you can import those (though you will still need to set other variable parameters)
    - **example: Pretest_Posttest_SPSS_F17_Workshop.xlsx**
  - Text files from other sources, e.g. tab- or comma-delimited text
  - Qualtrics (survey software) can export SPSS files



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	V1	String	20	0	ResponseID	None	None	15	Left	Nominal	Input
2	V2	String	20	0	ResponseSet	None	None	15	Left	Nominal	Input
3	V3	String	2000	0	Name	None	None	15	Left	Nominal	Input
4	V4	String	2000	0	ExternalDataRe...	None	None	15	Left	Nominal	Input
5	V5	String	2000	0	Email	None	None	15	Left	Nominal	Input
6	V6	String	2000	0	IPAddress	None	None	15	Left	Nominal	Input
7	V7	String	2000	0	Status	None	None	15	Left	Nominal	Input
8	V8	Date	20	0	StartDate	None	None	5	Right	Scale	Input
9	V9	Date	20	0	EndDate	None	None	5	Right	Scale	Input
10	V10	Numeric	1	0	Finished	None	None	5	Right	Scale	Input
11	Q2	Numeric	8	0	Please select y...	{1, Male}...	None	5	Right	Scale	Input
12	Q3_1_1_TEXT	String	2000	0	Please enter yo...	None	None	15	Left	Nominal	Input
13	Q3_1_2_TEXT	String	2000	0	Please enter yo...	None	None	15	Left	Nominal	Input
14	LocationLati...	String	2000	0	LocationLatitude	None	None	15	Left	Nominal	Input
15	LocationLon...	String	2000	0	LocationLongitu...	None	None	15	Left	Nominal	Input
16	LocationAc...	String	2000	0	LocationAccura...	None	None	15	Left	Nominal	Input

Raw data from Qualtrics usually needs to be cleaned up, but typically has question text as labels and answer text as values

Computing New Variables

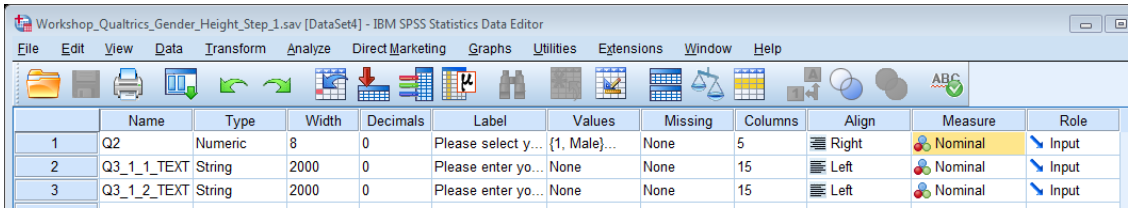
- Occasionally you may need to create new variables based on the values of existing variables.
- In this example, we will use data from a Qualtrics survey:
  - Ask about height, have respondent enter feet and inches
  - Compute inches and centimeters

Please select your gender:

Male
Female

Please enter your height. Use digits (0 through 9), and write fractions as decimals (e.g. "6.5" instead of "6 1/2").

	Feet	Inches
Height		



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	Q2	Numeric	8	0	Please select y...	{1, Male}...	None	5	Right	Nominal	Input
2	Q3_1_1_TEXT	String	2000	0	Please enter yo...	None	None	15	Left	Nominal	Input
3	Q3_1_2_TEXT	String	2000	0	Please enter yo...	None	None	15	Left	Nominal	Input

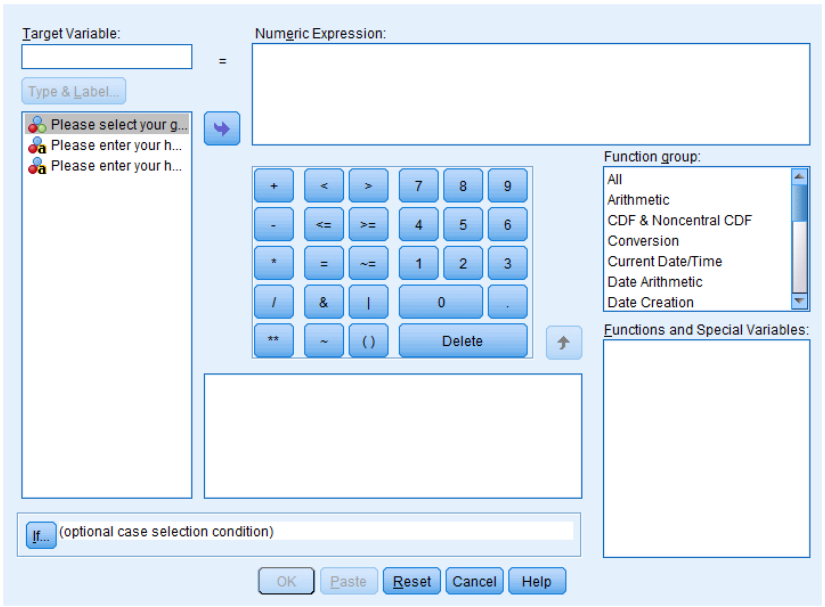
File: Workshop_Qualtrics_Gender_Height_Start.sav

Removed some unnecessary variables, **gender** stored as code in variable **Q2**, height stored as *text* in Q3_1_1_TEXT (feet) and Q3_1_2_TEXT (inches)

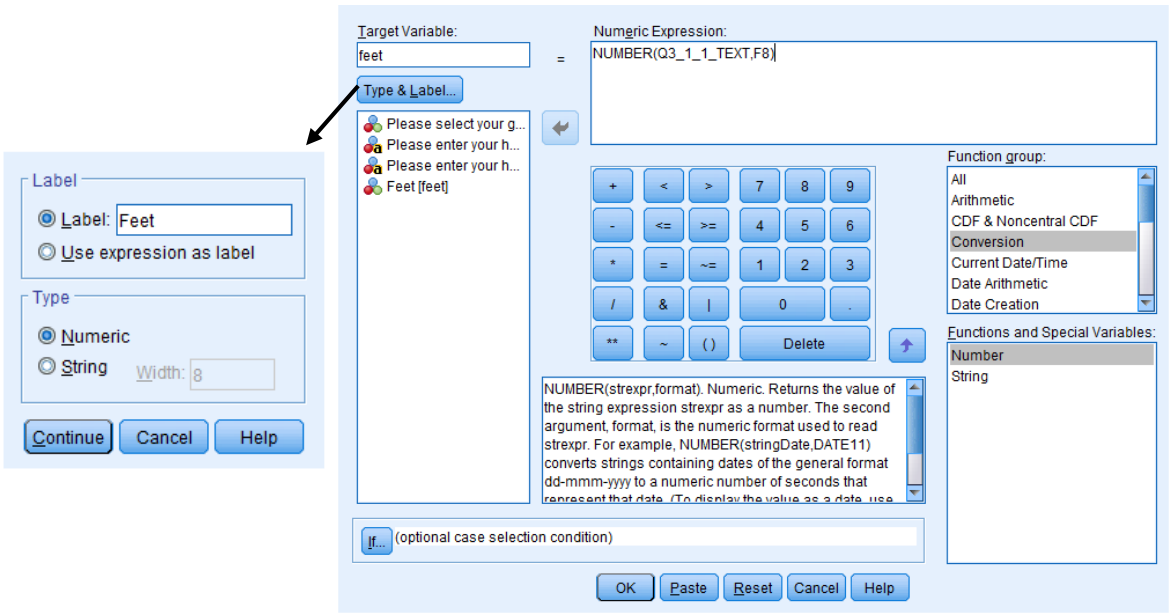
(the "Q" variable names are generated by Qualtrics, we'll give new variables better names)

Computing New Variables

- Use *Transform -> Compute Variable*
  - You will have list of existing variables (conveniently, as their labels), plus some math and conversion functions
  - You can create a new variable, give it a label and a type

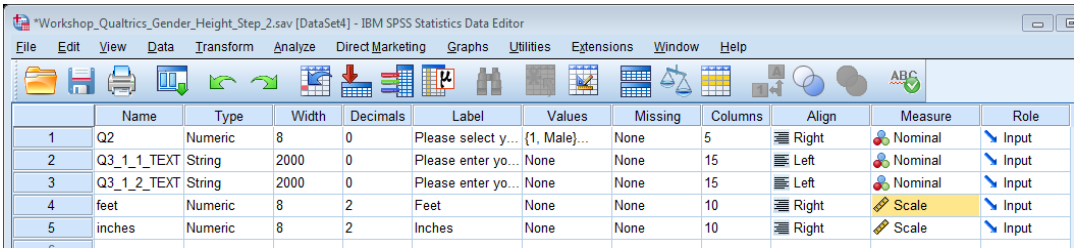


  - Original feet and inches are text – convert them to numbers:

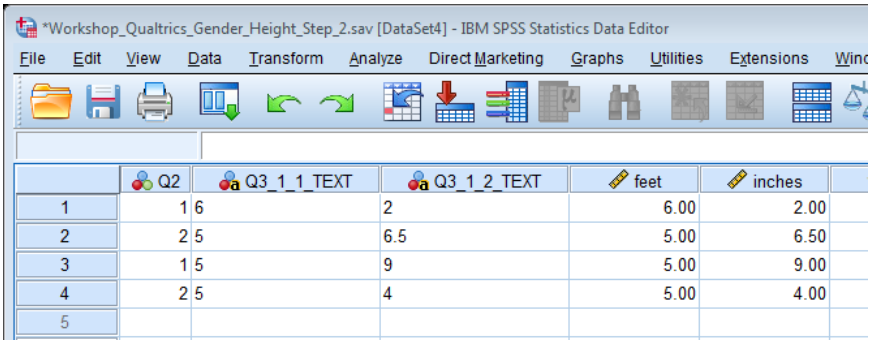


Computing New Variables

- Convert from text to number for both feet and inches

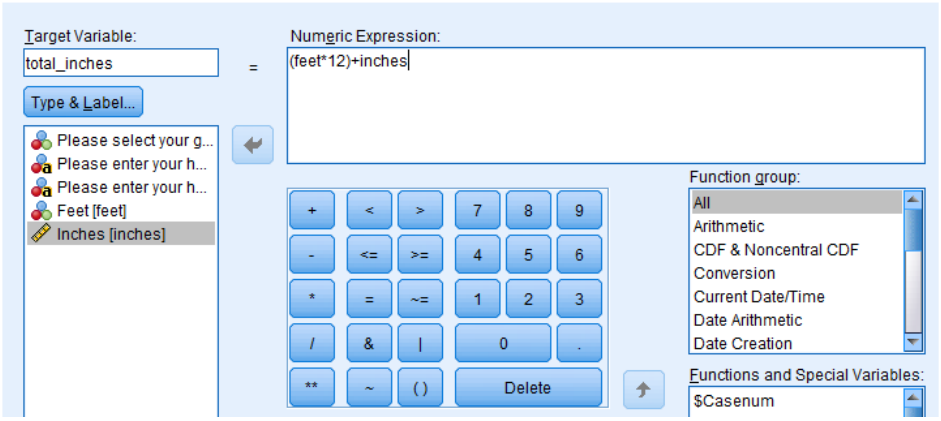


	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	Q2	Numeric	8	0	Please select y...	{1, Male}...	None	5	Right	Nominal	Input
2	Q3_1_1_TEXT	String	2000	0	Please enter yo...	None	None	15	Left	Nominal	Input
3	Q3_1_2_TEXT	String	2000	0	Please enter yo...	None	None	15	Left	Nominal	Input
4	feet	Numeric	8	2	Feet	None	None	10	Right	Scale	Input
5	inches	Numeric	8	2	Inches	None	None	10	Right	Scale	Input



	Q2	Q3_1_1_TEXT	Q3_1_2_TEXT	feet	inches
1	1	6	2	6.00	2.00
2	2	5	6.5	5.00	6.50
3	1	5	9	5.00	9.00
4	2	5	4	5.00	4.00
5					

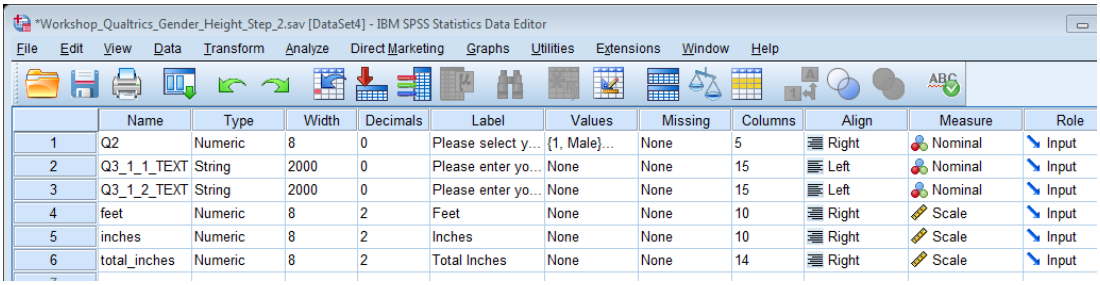
- Note that that the *scale of measurement* is (or should be!) **scale** for the transformed feet and inches – they can be plotted on a number line, have a meaningful zero, etc. The original was text, so had to be **nominal**
- Now compute *total* inches (i.e. remove feet)



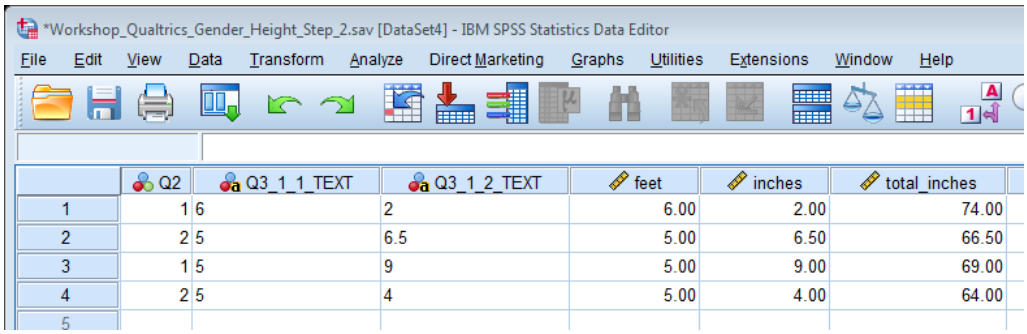
- Careful with missing values!
  - e.g. case where *feet* is a missing value above -> (0*12)+inches!
  - There are ways around this problem (e.g. use "If.." and check for missing variables)

Computing New Variables

- Our new variable computes *total inches = (feet * 12) + inches*



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	Q2	Numeric	8	0	Please select y...	{1, Male}...	None	5	Right	Nominal	Input
2	Q3_1_1_TEXT	String	2000	0	Please enter yo...	None	None	15	Left	Nominal	Input
3	Q3_1_2_TEXT	String	2000	0	Please enter yo...	None	None	15	Left	Nominal	Input
4	feet	Numeric	8	2	Feet	None	None	10	Right	Scale	Input
5	inches	Numeric	8	2	Inches	None	None	10	Right	Scale	Input
6	total_inches	Numeric	8	2	Total Inches	None	None	14	Right	Scale	Input



	Q2	Q3_1_1_TEXT	Q3_1_2_TEXT	feet	inches	total_inches
1	1	6	2	6.00	2.00	74.00
2	2	5	6.5	5.00	6.50	66.50
3	1	5	9	5.00	9.00	69.00
4	2	5	4	5.00	4.00	64.00
5						

(Why do this? Imagine we wanted to compare height between men and women. We can't average 6'2", 5'9", etc., but we *can* average 74", 69", etc.)

A bit about syntax (advanced)…

- The Output window shows the "programming" version of this (syntax). You can use this to repeat similar operations over and over again, or save a record of everything you have done so you could automatically repeat it with other data:

Automatically recorded to output window →

```
COMPUTE total_inches=(feet*12)+inches.
VARIABLE LABELS  total_inches 'Total Inches'.
EXECUTE.
```

Program something similar to compute centimeters →

```
COMPUTE total_cm=total_inches*2.54.
VARIABLE LABELS  total_cm 'Total Centimeters'.
EXECUTE.
```