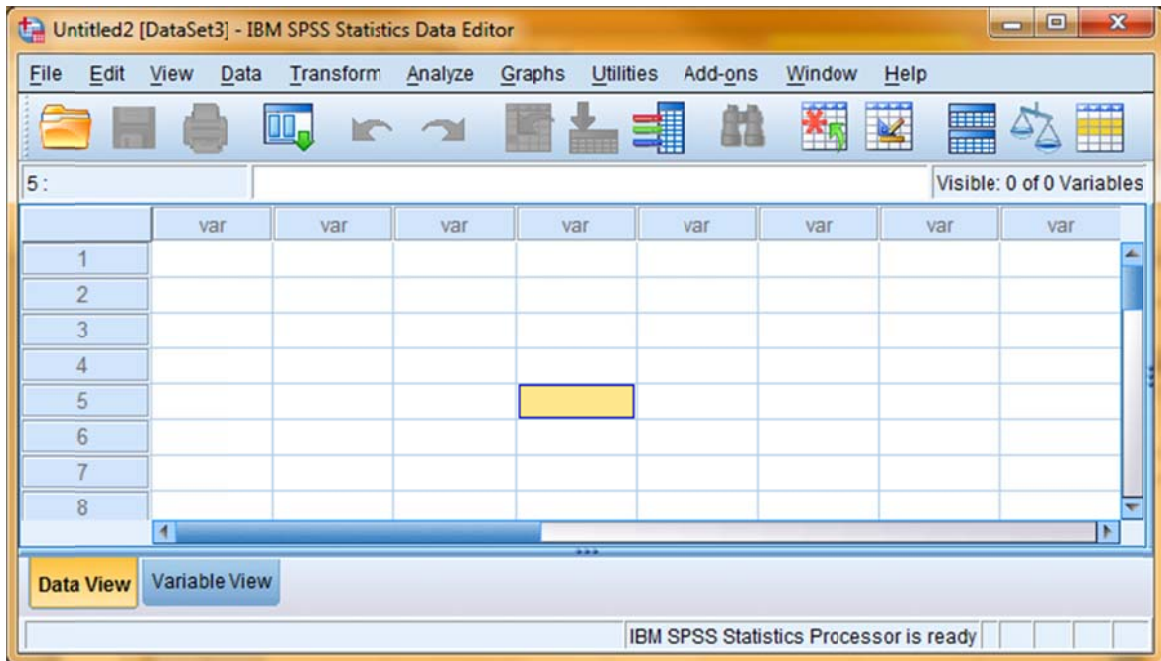


SPSS: Data Entry & Variable Specification



This shows the **Data View** of the editor.

- Each row holds the data from a participant or “case”
- Each column holds the data for a variable

Data can be entered or copied into the cells of the Data View

Remember to save often when entering data !!!

Data for this Example

The initial form of most datasets is a listing of the variable values for each participant. Usually each row is a different participant with a column for each variable.

Below are the data for 12 participants. Each “participant” is a different pet store. The researchers visited each store and recorded information about eight variables.

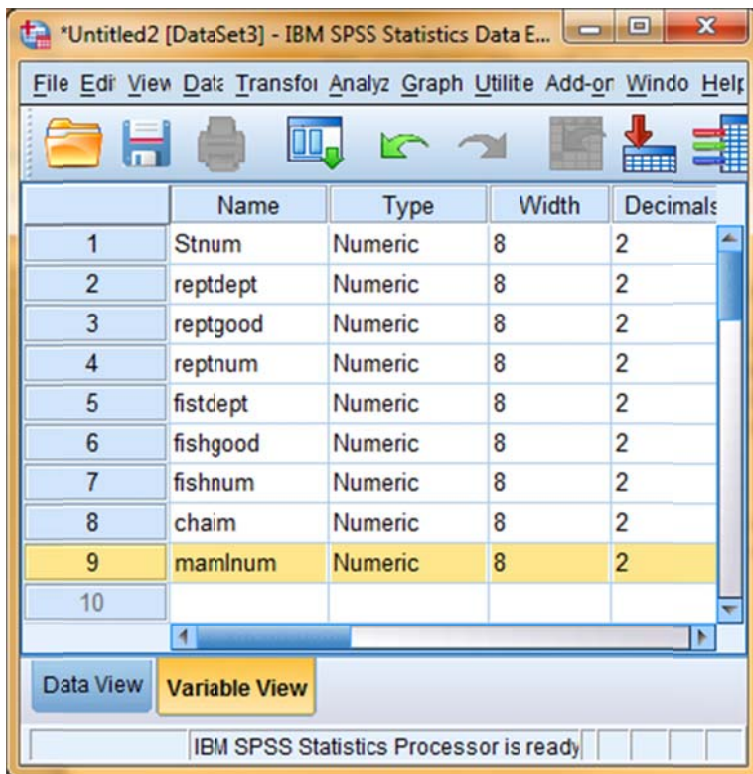
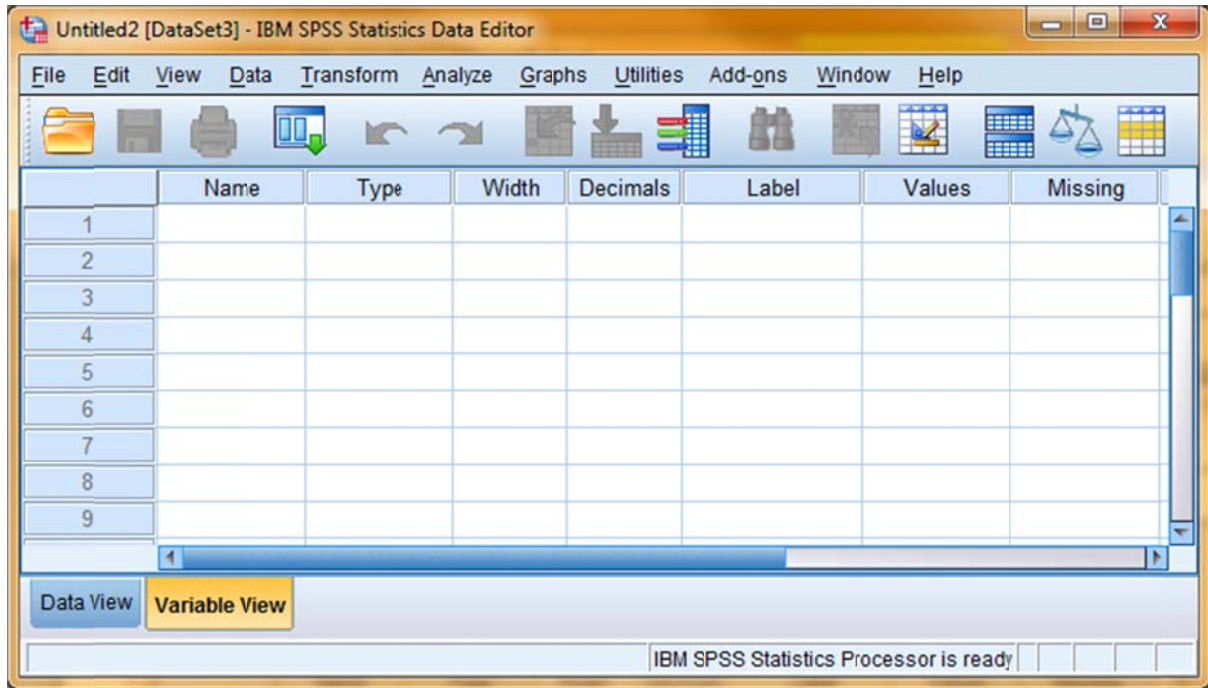
1	1	2	12	1	6	32	3	34
2	2	8	14	2	5	41	3	38
3	2	9	15	2	3	31	3	45
4	2	7	12	2	3	38	3	32
5	1	4	7	1	7	21	1	12
6	1	7	4	2	9	13	1	11
7	2	4	10	2	9	17	2	22
8	1	4	4	1	8	22	1	9
9	1	5	14	1	6	24	2	20
10	2	9	4	1	9	11	1	8
11	2	7	5	2	7	17	2	19
12	1	2	10	1	8	20	1	8

1. Participant number -- usually consecutive numbers starting with 1, 001 or some similar numbering system
2. Type of reptile department -- “1” means the reptile displays are mixed in with other animal and product displays and “2” means that the store has a separate department
3. Rating of the quality of the reptiles -- using a 1-10 scale
-9 was entered if a rating was not available
4. Count of the number of reptiles in the store
5. Type of fish available -- “1” means only freshwater fish are available in the store and “2” means that both freshwater and saltwater fish are available
6. Rating of the quality of the fish -- using a 1-10 scale
-9 was entered if a rating was not available
7. Count of the number of fish in the store
8. The ownership of the store -- “1” means the store is a franchise of a pet store chain, “2” means the store is privately owned, and “3” means the store is part of a pet store cooperative
9. Count of the number of mammals in the store

This shows the **Variable View** of the editor.

- Each row holds the specification for a variable
- Each column holds a different specification

It is important to specify key aspects of each variable. These specifications will make the statistical analyses and output clearer and easier to read as well as provide a “reminder” of the decisions you made when collecting and entering the data.



Naming Variables

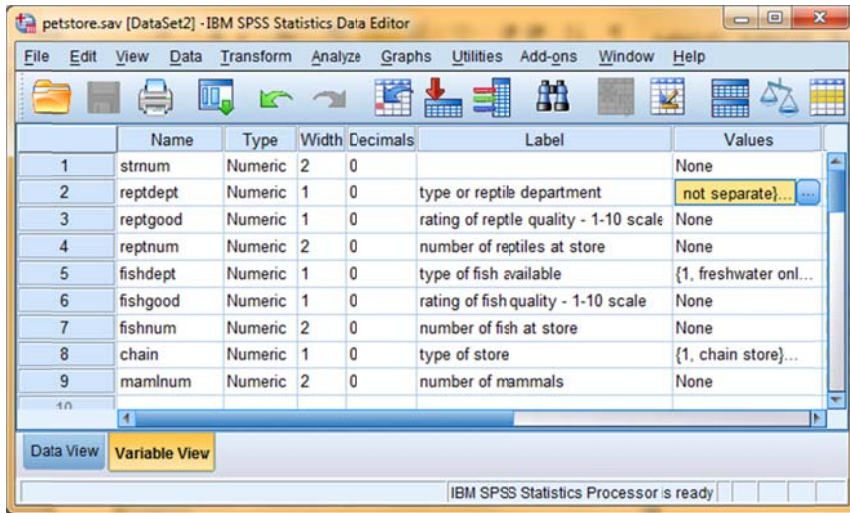
Click a cell under **Variable** and type in the variable name

- Variable names can be really long -- letters, numbers, “_” and “.” work best -- the first character must be a letter.
- When using the GUI to analyze data, only the first 25 or so characters are easy to see. Be sure that the unique part of each variable name is early in the variable name
- When you enter a name for a variable a set of “default” values automatically appears in the other columns -- some of these we will change later
- The variable names for these data are shown to the left, along with the default values

Type, Width & Decimals

The default values for these specifications usually work very well.

- “Numerical” is the most common type for both qualitative & quantitative data
- “Width” tells the number of digits in the largest variable value -- must be at least one more than the number of decimal values “Decimals” tells how many digits of a variable value are decimal values

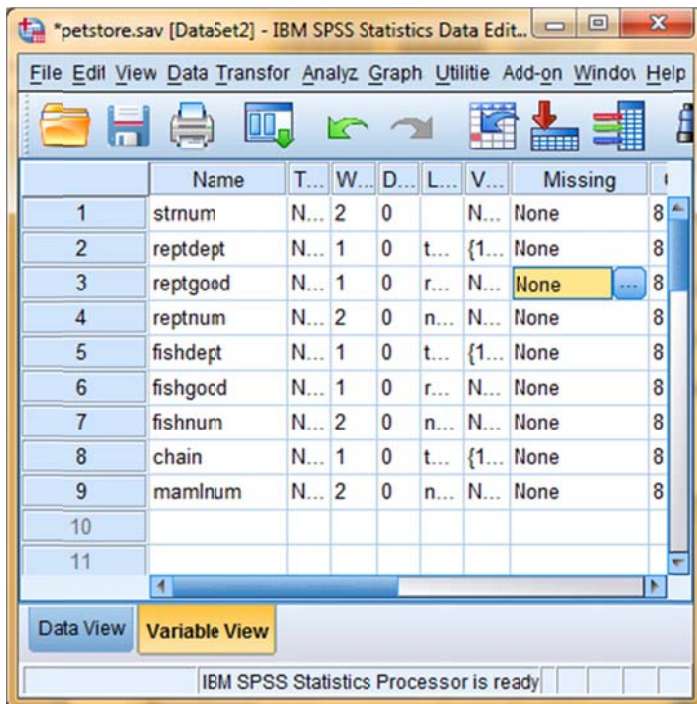
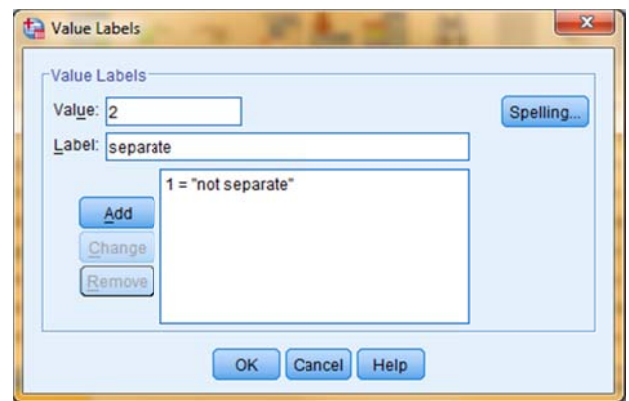


Variable Labels

- Click a cell under **Label** and type up to a 256-character description of the variable -- this label will appear in the output making it easier to understand analyses
- Giving a Label to every variable in the data set prevents forgetting what the shorter variable names mean

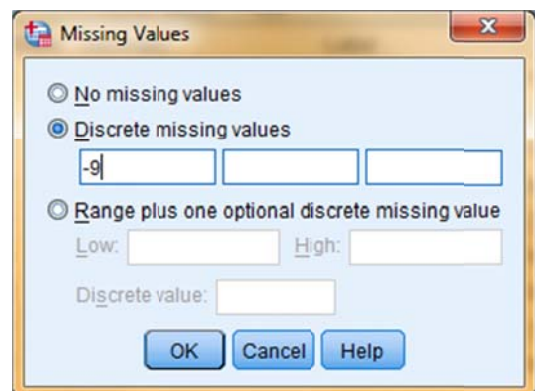
Value Labels

- Specify Value Labels for qualitative/categorical variables
- These values will appear in the output and make the analyses easier to understand
- Click a cell under **Values**. Click the blue box that appears in the cell -- shown above for reptdept in row 2
- The Value Labels window will appear. In it type each "Value," its "Value "Label", and click **Add** -- shown to the right for the value "2" of reptdept
- You can also highlight a specified value label and **Change** or **Remove** it



Missing Values

- If you leave a data cell blank a "." will appear in that cell, signaling a "system missing value"
- Use "user-missing" values to indicate different types of missing data (e.g., -99 = did not respond, -98 = answer could not be scored -- up to 3 different missing values)
- Click the cell under **Missing**. Click the little blue box that appears in the cell -- shown below for reptdept
- Click the "Discrete missing values" button and put missing values in the window
- **Be sure to specify the meaning of each missing value in the Values column !!!**



Entering Data

The “Data View” and “Variable View” for the completed data set is shown below.

Visible: 9 of 9 Variables

	strnum	reptdept	reptgood	reptnum	fishdept	fishgood	fishnum	chain	mamnum
1	1	1	2	12	1	6	32	3	34
2	2	2	8	14	2	5	41	3	38
3	3	2	9	15	2	3	31	3	45
4	4	2	7	12	2	3	38	3	32
5	5	1	4	7	1	7	21	1	12
6	6	1	7	4	2	9	13	1	11
7	7	2	4	10	2	9	17	2	22
8	8	1	4	4	1	8	22	1	9
9	9	1	5	14	1	6	24	2	20
10	10	2	9	4	1	9	11	1	8
11	11	2	7	5	2	7	17	2	19
12	12	1	2	10	1	8	20	1	8
13									

Data View Variable View

IBM SPSS Statistics Processor is ready

	Name	Type	Width	Decimals	Label	Values	Missing
1	strnum	Numeric	2	0		None	None
2	reptdept	Numeric	1	0	type or reptile department	{1, not separate}...	None
3	reptgood	Numeric	1	0	rating of reptile quality - 1-10 scale	None	-9
4	reptrum	Numeric	2	0	number of reptiles at store	None	None
5	fishdept	Numeric	1	0	type of fish available	{1, freshwater only}...	None
6	fishgood	Numeric	1	0	rating of fish quality - 1-10 scale	None	-9
7	fishnum	Numeric	2	0	number of fish at store	None	None
8	chain	Numeric	1	0	type of store	{1, chain store}...	None
9	mamnum	Numeric	2	0	number of mammals	None	None
10							

Data View Variable View

IBM SPSS Statistics Processor is ready

SPSS data set files are saved using the file extension “.sav” Remember to save early and often !!!!