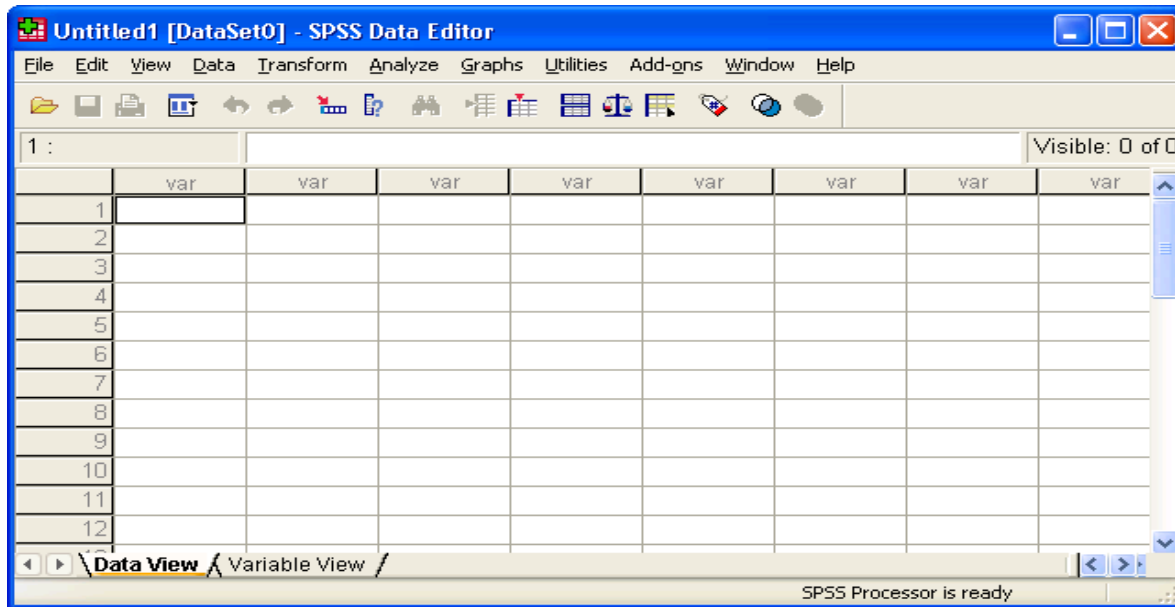


## Creating a New SPSS Dataset -- Variable Specification and Data Entry

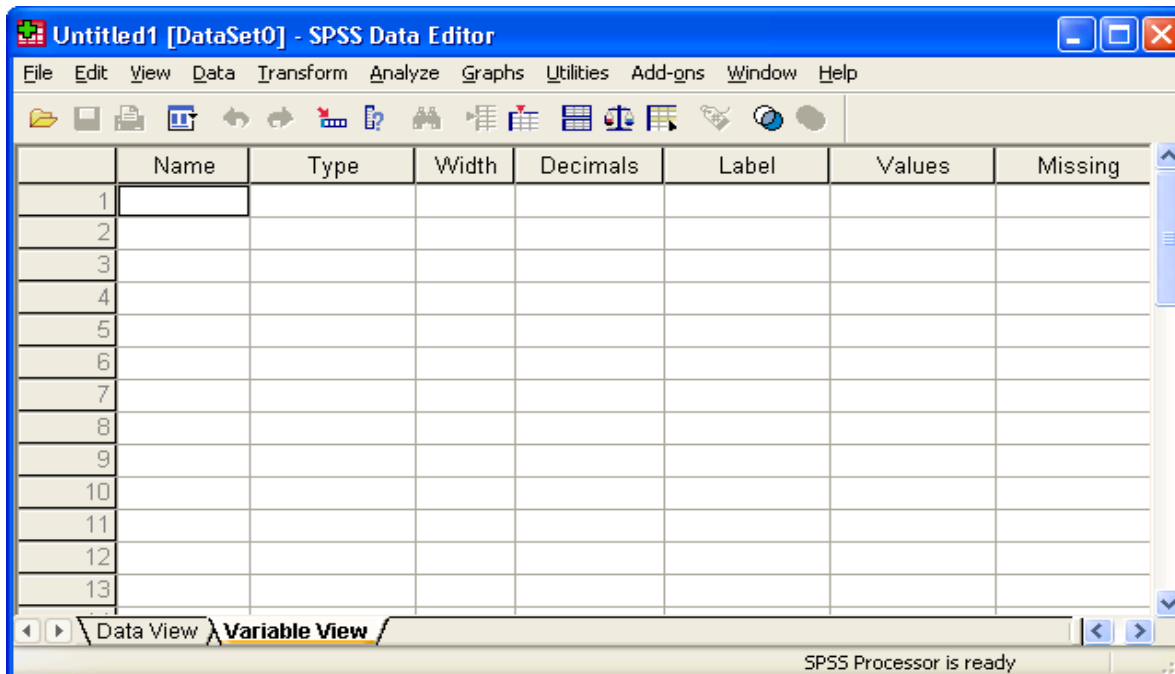
When you open SPSS you will see the Data Editor



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

After specifying the information for each variable using the Variable View of the editor, we will return to this Data View to enter the data from each case.



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.

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 that was collected.

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. Those variables and details about each are shown on the right

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

Participant number -- usually consecutive numbers starting with 1, 001 or some similar numbering system

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

Rating of the quality of the reptiles -- using a 1-10 scale

Count of the number of reptiles in the store

Type if fish available -- “1” means only freshwater fish are available in the store and “2” means that both freshwater and saltwater fish are available

Rating of the quality of the fish -- using a 1-10 scale

Count of the number of fish in the store

The ownership of the store -- “1” means the store is a franchise of a petstore chain, “2” means the store is privately owned, and “3” means the store is part of a petstore cooperative

Count of the number of mammals in the store

	Name	Type	Width	Decimals	Label	Values
1	strnum	Numeric	8	2		None
2	reptdept	Numeric	8	2		None
3	reptgood	Numeric	8	2		None
4	reptnum	Numeric	8	2		None
5	fishdept	Numeric	8	2		None
6	fishgood	Numeric	8	2		None
7	fishnum	Numeric	8	2		None
8	chain	Numeric	8	2		None
9	mamlnum	Numeric	8	2		None

## Naming Variables

- Click a cell under **Variable** and type in the variable name
- Variable names may have up to 8 characters -- letters, numbers, “-” and “\_” work best -- the first character must be a letter
- 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 right, 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

petstore.sav - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

	Name	Type	Width	Decimals	Label	Values
1	strnum	Numeric	2	0	store number	None
2	reptdept	Numeric	1	0	type or reptile	{1, not sep ...
3	reptgood	Numeric	1	0	rating of reptile	None
4	reptnum	Numeric	2	0	number of repti	None
5	fishdept	Numeric	1	0	type of fish ava	{1, freshwater
6	fishgood	Numeric	1	0	rating of fish q	None
7	fishnum	Numeric	2	0	number of fish	None
8	chain	Numeric	2	0	type of store	{1, chain store}
9	mamlnum	Numeric	2	0	number of ma	None

Data View Variable View

Processor area SPSS Processor is read

## Variable Labels

- Click a cell under **Label** and type up to a 40-character description of the variable -- this label will appear in the output making it easier to understand
- giving a Label to every variable in the data set prevents forgetting what the 8-character variable names mean

## Value Labels

- Specify Values for every qualitative variable -- these will appear in the output
- Click a cell under **Values**. Click the gray box that appears in the cell -- shown on the left 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 later

Value Labels

Value Labels

Value: 2

Value Label: separate department

Add 1 = "not separate"

Change Remove

OK Cancel Help

petstore.sav - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

	Name	Type	Width	Decimals	Label	Values	Missing
1	strnum	Numeric	2	0	store number	None	None
2	reptdept	Numeric	1	0	type or reptile	{1, not separat	None
3	reptgood	Numeric	1	0	rating of reptile	None	None
4	reptnum	Numeric	2	0	number of repti	None	None
5	fishdept	Numeric	1	0	type of fish ava	{1, freshwater	None
6	fishgood	Numeric	1	0	rating of fish q	None	None
7	fishnum	Numeric	2	0	number of fish	None	None
8	chain	Numeric	2	0	type of store	{1, chain store}	None
9	mamlnum	Numeric	2	0	number of ma	None	None

Data View Variable View

Processor area SPSS Processor is ready

## Missing Values

- leaving a cell in the dataset blank tells SPSS that value for that participant has a missing value for that variable -- a "." will appear in that cell, signalling a "system missing value"
- you can also identify alternative "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 gray box that appears in the cell -- shown on the left for reptdept in row 2
- Click the "Discrete missing values" button and enter the missing values you've chosen in the windows -- shown below for "-9" for reptdept
- **Be sure to specify the meaning of each missing value in the "Values" column**

Missing Values

No missing values

Discrete missing values

-9

Range plus one optional discrete missing value

Low: High:

Discrete value:

OK Cancel Help

## Entering Data

- Click the “**Data View**” tab in the bottom-left corner of the screen
- Enter each variable value -- in the correct column -- for each case or participant
- If you leave a space blank a “.” will appear, indicating that the value is missing -- called a “system missing” value

The completed data set for the petstores is shown below. These data will be used for many of the data analysis examples in this handbook

	strnum	reptdept	reptgood	reptnum	fishdept	fishgood	fishnum	chain	mamlnum
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									

## Saving the Dataset

- **File** then **Save As** and proceed with the usual Windows “save” routine
- SPSS data sets have the extension “.sav”