Chapter 193

Descriptive Statistics – Summary Tables (Old Version)

Introduction

This procedure produces tables of means, medians, percentiles, standard deviations, coefficients of variation, sums, and counts for various combinations of grouping (break) variables. Nine tabular formats are available. The tables are similar in structure to those produced by cross tabulation.

This module is used to summarize data containing a combination of continuous and categorical variables. Large volumes of such data may be summarized in statistical tables of means, counts, or standard deviation. Discussions of these statistics are found in the Descriptive Statistics chapter and will not be reproduced here.

Types of Categorical Variables

Note that we will refer to two types of categorical variables: *Categorical* and *Grouping*. Grouping variables are used to split a database into subgroups. A separate table is generated for each unique set of values of the grouping variables. The values of a categorical variable are used to define the rows and columns of the tabulation table. Up to two categorical variables may be used per table.

Table Layouts

Several table layouts are available in **NCSS**. These layouts are defined in terms of the number of tables, the table rows, the sub-rows within a row, and the table columns.

1 TABLES: One, ROWS: Data, COLUMNS: Statistics

Tables: only one table.

Rows: a row for each Data Variable.

Columns: a column for each Statistic.

An example of this layout is:

Variable	Count	Mean	Median	Std Deviation
X1	122	12.04	12	4.527
X2	124	23.45	25	5.831
X3	133	34.16	38	6.094
X4	126	61.38	63	3.725

2 TABLES: One, ROWS: Categorical, COLUMNS: Data, SUBROWS: Statistics

Tables: only one table.

Rows: a set of rows for each category of the Row Variable.

Sub-Rows: a row for each Statistic.

Columns: a column for each Data Variable.

An example of this layout is:

		Data Variable		
Row Variable		X1	X2	X3
Group1	Mean	12.36	23.77	51.78
	Std Dev	57.62	62.17	79.18
Group2	Mean	87.65	54.32	43.21
	Std Dev	42.89	22.33	64.87

3 TABLES: One, ROWS: Data, COLUMNS: Categorical, SUBROWS: Statistics

Tables: only one table.

Rows: a set of rows for each Data Variable.

Sub-Rows: a row for each Statistic.

Columns: a column for each category of the Column Variable.

An example of this layout is:

		Column Variable		
Data Variables	Group1	Group2	Group3	
X1	Mean	12.36	23.77	51.78
	Std Dev	57.62	62.17	79.18
X2	Mean	87.65	54.32	43.21
	Std Dev	42.89	22.33	64.87

4 TABLES: Statistics, ROWS: Categorical, COLUMNS: Data

Tables: a separate table (and plot) for each Statistic.

Rows: a row for each category of the Row Variable.

Columns: a column for each Data Variable.

An example of this layout is:

Table of Means

	Data Variables		
Row Variable	X1	X2	X3
Group1	12.36	23.77	51.78
Group2	57.62	62.17	79.18
Group3	87.65	54.32	43.21
Group4	42.89	22.33	64.87

5 TABLES: Statistics, ROWS: Data, COLUMNS: Categorical

Tables: a separate table (and plot) for each Statistic.

Rows: a row for each Data Variable.

Columns: a column for each category of the Column Variable.

An example of this layout is:

Table of Means

Data Variables	Group1	Group2	Group3
	Column Variable		
X1	12.36	23.77	51.78
X2	57.62	62.17	79.18
X3	87.65	54.32	43.21
X4	42.89	22.33	64.87

6 TABLES: Statistics, ROWS: Categorical, COLUMNS: Categorical, SUBROWS: Data

Tables: a separate table for each Statistic.

Rows: a set of rows for each category of the Row Variable.

Sub-Rows: a row for each Data Variable.

Columns: a column for each category of the Column Variable.

An example of this layout is:

Table of Means

Row Variable	Data Variable	Group1	Group2	Group3
		Column Variable		
Level1	X1	12.36	23.77	51.78
	X2	57.62	62.17	79.18
Level2	X1	87.65	54.32	43.21
	X2	42.89	22.33	64.87

7 TABLES: Data, ROWS: Categorical, COLUMNS: Categorical, SUBROWS: Statistics

Tables: a separate table for each Data Variable.

Rows: a set of rows for each category of the Row Variable.

Sub-Rows: a row for each Statistic.

Columns: a column for each category of the Column Variable.

An example of this layout is:

Summary for X1

Row Variable		Group1	Group2	Group3
		Column Variable		
Level1	Mean	12.36	23.77	51.78
	Std Dev	57.62	62.17	79.18
Level2	Mean	87.65	54.32	43.21
	Std Dev	42.89	22.33	64.87

8 TABLES: Data and Statistics, ROWS: Categorical, COLUMNS: Categorical

Tables: a separate table (and plot) for each Data Variable and Statistic.

Rows: a row for each category of the Row Variable.

Columns: a column for each category of the Column Variable.

An example of this layout is:

Means of X1

Row Variables	Column Variable		
	Group1	**Group2**	**Group3**
Level1	12.36	23.77	51.78
Level2	57.62	62.17	79.18
Level3	87.65	54.32	43.21
Level4	42.89	22.33	64.87

9 Data Summary List

An item-by-item list of the statistics. *Note that only six columns can be displayed.*
Tables: one table

Rows: a row for each Grouping Variable category and Row Variable category.

Columns: a column for each Grouping Variable, Row Variable, and statistic.

An example of this layout is:

Group Variable	Row Variable	Mean	StdDev
Level1	Group1	12.36	23.77
Level1	Group2	57.62	62.17
Level2	Group1	54.32	43.21
Level2	Group2	22.33	64.87

Data Structure

The data below are a subset of the Resale dataset provided with the software. This (computer simulated) data gives the selling price, the number of bedrooms, the total square footage (finished and unfinished), and the size of the lots for 150 residential properties sold during the last four months in two states. Only the first 5 of the 150 observations are displayed.

Resale Dataset (Subset)

State	Price	Bedrooms	TotalSqft	LotSize
Nev	260000	2	2042	10173
Nev	66900	3	1392	13069
Vir	127900	2	1792	7065
Nev	181900	3	2645	8484
Nev	262100	2	2613	8355

Missing Values

Observations with missing values in either the categorical variable or the continuous variable are ignored.

Example 1 – Layout 1: Variable Summary Report

Setup

To run this example, complete the following steps:

1   Open the Resale example dataset
- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **Resale** and click **OK**.

2   Specify the Descriptive Statistics – Summary Tables (Old Version) procedure options
- Find and open the **Descriptive Statistics – Summary Tables (Old Version)** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 1** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

  Variables Tab

Layout	**1. TABLES: One, ROWS: Data, COLUMNS: Statistics**
Data Variables	**Price-LotSize**
Count	**Checked**
Mean	**Checked**
Median	**Checked**
Std Dev	**Checked**
COV	**Checked**
COD	**Checked**

  Report Options (*in the Toolbar*)

Variable Labels	**Column Labels**

3   Run the procedure
- Click the **Run** button to perform the calculations and generate the output.

Variable Summary Report

Variable Summary Report
_____

Section 1

Variables	Count	Mean	Median	Standard Deviation	Coef of Variation
Sales Price	150	174392	158200	97656.81	0.55998
Year Built	150	1971.273	1973	13.84667	0.00702
Bedrooms	150	2.42	2	0.8919476	0.36857
Bathrooms	150	2.4	2.5	0.8047677	0.33532
Garage Size	150	1.266667	1	0.5636252	0.44497
Fireplaces	150	0.96	1	0.6939818	0.72290
Quality Index	150	0.7316667	0.75	0.35248	0.48175
Brick Ratio	150	0.5033333	0.5	0.4157013	0.82590
Total Area (Sqft)	150	1893.38	1872.5	754.2496	0.39836
Finished Area (Sqft)	150	1597.947	1496	672.1644	0.42064
Lot Size (Sqft)	150	8366.913	8344.5	2376.334	0.28402

The definitions of these statistics are identical to those found in the Descriptive Statistics chapter. They will not be repeated here.

Example 2 – Layout 3: Categorical Variable as Columns

Setup

To run this example, complete the following steps:

1   Open the Resale example dataset
- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **Resale** and click **OK**.

2   Specify the Descriptive Statistics – Summary Tables (Old Version) procedure options
- Find and open the **Descriptive Statistics – Summary Tables (Old Version)** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 2** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Variables Tab

Layout	**3. TABLES: One, ROWS: Data, COLUMNS: Categorical, SUBROWS: Statistics**
Data Variables	**Price,TotalSqft,LotSize**
Count	**Checked**
Mean	**Checked**
Std Dev	**Checked**
Column Variables	**State**

Report Options (*in the Toolbar*)

Variable Labels	**Column Labels**
Data Labels	**Value Labels**

3   Run the procedure
- Click the **Run** button to perform the calculations and generate the output.

Statistical Summary

Statistical Summary

Variables		State		
		Nevada	**Virginia**	**Total**
Sales Price	Count	88	62	150
	Mean	170762.5	179543.5	174392
	Standard Deviation	98665.72	96771.49	97656.81
Total Area (Sqft)	Count	88	62	150
	Mean	1881.33	1910.484	1893.38
	Standard Deviation	788.569	708.6572	754.2496
Lot Size (Sqft)	Count	88	62	150
	Mean	8571.454	8076.597	8366.913
	Standard Deviation	2419.88	2301.226	2376.334

The definitions of these statistics are identical to those found in the Descriptive Statistics chapter. They will not be repeated here.

Example 3 – Layout 5: Table of One Statistic

Setup

To run this example, complete the following steps:

1   Open the Resale example dataset
- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **Resale** and click **OK**.

2   Specify the Descriptive Statistics – Summary Tables (Old Version) procedure options
- Find and open the **Descriptive Statistics – Summary Tables (Old Version)** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 3** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

  Variables Tab

  Layout ....**5. TABLES: Statistics, ROWS: Data, COLUMNS: Categorical (Plots Possible)**
  Data Variables ....**Bedrooms-Fireplace**
  Mean...**Checked**
  Column Variables ....**State**

  Report Options Tab

  Show Total...**On Reports and Plots**

  Report Options (*in the Toolbar*)

  Variable Labels...**Column Labels**
  Data Labels...**Value Labels**
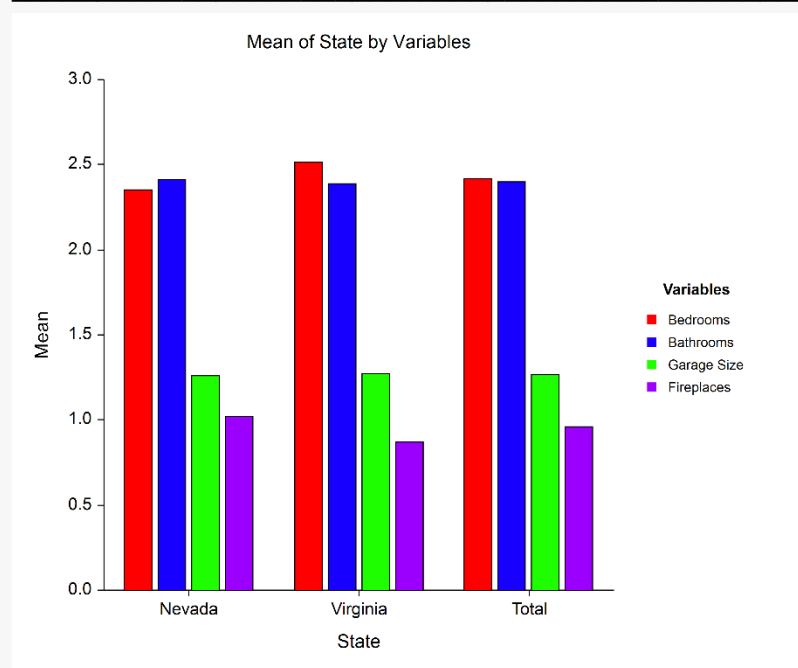
3   Run the procedure
- Click the **Run** button to perform the calculations and generate the output.

Table and Plot of One Statistic

Table of Means

	State		
Variables	**Nevada**	**Virginia**	**Total**
Bedrooms	2.352273	2.516129	2.42
Bathrooms	2.409091	2.387097	2.4
Garage Size	1.261364	1.274194	1.266667
Fireplaces	1.022727	0.8709677	0.96

Plot of Means

The definitions of these statistics are identical to those found in the Descriptive Statistics chapter. They will not be repeated here.

Example 4 – Layout 6: Multiple Y's, Two Categoricals, One Statistic

Setup

To run this example, complete the following steps:

1   Open the Resale example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **Resale** and click **OK**.

2   Specify the Descriptive Statistics – Summary Tables (Old Version) procedure options

- Find and open the **Descriptive Statistics – Summary Tables (Old Version)** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 4** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Variables Tab

Layout..	**6. TABLES: Statistics, ROWS: Categorical, COLUMNS: Categorical, SUBROWS: Data**
Data Variables	**Price,FinishSqft-LotSize**
Create Other Row Variables from Numeric Data	**Checked**
Numeric Variables to Categorize for Use in Table Rows.......	**TotalSqft**
Group Numeric Data into Categories Using.............	**List of Interval Upper Limits**
List	**1000 2000 3000**
Column Variables	**State**

Report Options Tab

Show Total..	**On Reports and Plots**
Rows Decimals..	**Auto (Up to 7)**

Report Options (*in the Toolbar*)

Variable Labels...	**Column Labels**
Data Labels..	**Value Labels**

3   Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Multiple Y's, Two Categoricals, One Statistic

Means of Sales Price, Finished Area (Sqft), Lot Size (Sqft)

		State		
Total Area (Sqft)		**Nevada**	**Virginia**	**Total**
Up To 1000	Sales Price	160475	142850	152921.4
	Finished Area (Sqft)	738.125	739.6667	738.7857
	Lot Size (Sqft)	8816	9857.833	9262.5
1000 To 2000	Sales Price	153293.3	172992.9	160849.3
	Finished Area (Sqft)	1234.311	1247.179	1239.247
	Lot Size (Sqft)	9094.8	7674.286	8549.945
2000 To 3000	Sales Price	197200	186461.5	192029.6
	Finished Area (Sqft)	1974.214	2086.077	2028.074
	Lot Size (Sqft)	7503.179	8129.808	7804.889
Over 3000	Sales Price	189071.4	291400	211811.1
	Finished Area (Sqft)	3375.143	2871	3263.111
	Lot Size (Sqft)	9200.714	7673.5	8861.333
Total	Sales Price	170762.5	179543.5	174392
	Finished Area (Sqft)	1594.92	1602.242	1597.947
	Lot Size (Sqft)	8571.454	8076.597	8366.913

The definitions of these statistics are identical to those found in the Descriptive Statistics chapter. They will not be repeated here.

Example 5 – Layout 7: Complete Summary for each Data Variable

Setup

To run this example, complete the following steps:

1   Open the Resale example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **Resale** and click **OK**.

2   Specify the Descriptive Statistics – Summary Tables (Old Version) procedure options

- Find and open the **Descriptive Statistics – Summary Tables (Old Version)** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 5** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Variables Tab

Layout...**7. TABLES: Data, ROWS: Categorical,**	
	COLUMNS: Categorical, SUBROWS: Statistics
Data Variables ....**Price**	
Create Other Row Variables from....................**Checked**	
Numeric Data	
Numeric Variables to Categorize for Use.......**TotalSqft**	
in Table Rows	
Group Numeric Data into Categories.............**List of Interval Upper Limits**	
Using	
List **1000 2000 3000**	
Column Variables ....**State**	

Report Options Tab

Show Total...**On Reports and Plots**	
Rows Decimals...**Auto (Up to 7)**	

Report Options (*in the Toolbar*)

Variable Labels...**Column Labels**	
Data Labels..**Value Labels**	

3   Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

Complete Summary for Each Data Variable

Statistical Summary of Sales Price

Total Area (Sqft)		State		
		Nevada	Virginia	Total
Up To 1000	Count	8	6	14
	Mean	160475	142850	152921.4
	Median	136050	85200	110500
	Standard Deviation	110945.7	107838.2	105747.5
1000 To 2000	Count	45	28	73
	Mean	153293.3	172992.9	160849.3
	Median	123400	163000	150100
	Standard Deviation	91336.91	71798.73	84405.74
2000 To 3000	Count	28	26	54
	Mean	197200	186461.5	192029.6
	Median	182850	145550	176250
	Standard Deviation	106136.7	111024.2	107621.7
Over 3000	Count	7	2	9
	Mean	189071.4	291400	211811.1
	Median	150900	291400	168500
	Standard Deviation	94037.06	173806.8	111554.4
Total	Count	88	62	150
	Mean	170762.5	179543.5	174392
	Median	151050	162800	158200
	Standard Deviation	98665.72	96771.49	97656.81

The definitions of these statistics are identical to those found in the Descriptive Statistics chapter. They will not be repeated here.

Example 6 – Layout 8: One Data Variable and Statistic, Two Categories

Setup

To run this example, complete the following steps:

1   Open the Resale example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **Resale** and click **OK**.

2   Specify the Descriptive Statistics – Summary Tables (Old Version) procedure options

- Find and open the **Descriptive Statistics – Summary Tables (Old Version)** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 6** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Variables Tab

Layout..**8. TABLES: Data and Statistics, ROWS: Categorical, COLUMNS: Categorical (Plots Possible)**

Data Variables ....**Price**

Row Variables..**State**

Column Variables ....**<Empty>**

  Create Other Column Variables.....................**Checked**
  from Numeric Data

    Numeric Variables to Categorize for Use.....**TotalSqft**
    in Table Columns

    Group Numeric Data into Categories...........**List of Interval Upper Limits**
    Using

      List **1000 2000 3000**

Reports Options Tab

Show Total...**On Reports and Plots**

Rows Decimals...**0**

Mean, Sum Decimals.......................................**0**

Plots Tab

Bar Chart Format (*Click the Button*)

Group Axis Tab

Lower Axis Tick Label Layout (*Click the Button*)

    Alignment ....**Right**

    Rotation Angle...**45**

    Margin Above the Text **10**

Report Options (*in the Toolbar*)

Variable Labels...**Column Labels**
Data Labels...**Value Labels**

3   Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

One Data Variable and Statistic, Two Categories

Means of Sales Price

State	Total Area (Sqft)				
	Up To 1000	**1000 To 2000**	**2000 To 3000**	**Over 3000**	**Total**
Nevada	160475	153293	197200	189071	170763
Virginia	142850	172993	186462	291400	179544
Total	152921	160849	192030	211811	174392

Plot of Means

Mean Price of Total Area (Sqft) by State
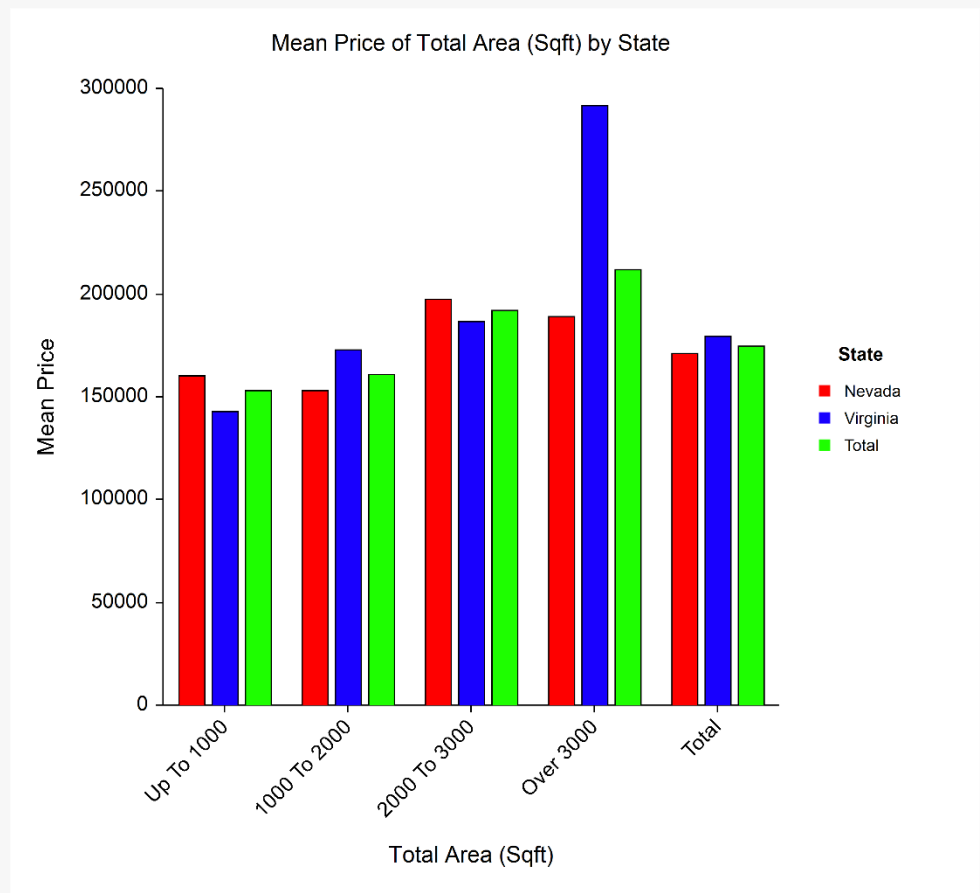
Example 7 – List Format

Setup

To run this example, complete the following steps:

1   Open the Resale example dataset

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Select **Resale** and click **OK**.

2   Specify the Descriptive Statistics – Summary Tables (Old Version) procedure options

- Find and open the **Descriptive Statistics – Summary Tables (Old Version)** procedure using the menus or the Procedure Navigator.
- The settings for this example are listed below and are stored in the **Example 7** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Variables Tab
_____

Layout..**9. DATA SUMMARY LIST**
Data Variables ....**Price**
Row Variables...**Neighborhood**

Breaks Tab
_____

Number of Grouping Variables **2**
Grouping Variable 1 ....**State**
Grouping Variable 2..**City**

Report Options (*in the Toolbar*)
_____

Variable Labels...**Column Names**
Data Labels..**Data Values**

3   Run the procedure

- Click the **Run** button to perform the calculations and generate the output.

List Format Report

Summary List

State	City	Neighborhood	Price Count	Price Mean	Price SD
Nev	1	1	11	203727.3	105805.4
Nev	1	2	16	183625	105754.7
Nev	2	3	16	135018.8	94628.04
Nev	2	4	13	156192.3	93304.72
Nev	2	5	20	192190	100400.5
Nev	3	6	12	151125	88063.07
Vir	4	7	13	197307.7	80288.13
Vir	4	8	14	168700	86626.27
Vir	5	9	6	178716.7	107857.3
Vir	5	10	9	159511.1	132957.2
Vir	5	11	9	150488.9	70977.03
Vir	6	12	11	212963.6	112784.7

The definitions of these statistics are identical to those found in the Descriptive Statistics chapter. They will not be repeated here.

This format is especially useful for creating a database containing only summary information such as the means, standard deviations, etc. To create a summary database, take the following steps:

1. Run this report on the data, summarizing across the categorical variables of interest.

2. Copy the output report to the clipboard.

3. Open a new database (or spreadsheet).

4. Paste the data from the clipboard to this new database by placing the cursor in the upper-left cell and pasting. The paste can use the Ctrl-V key or Paste from the Edit menu.

5. Label the columns in the Variable Info sheet.