

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.

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 8 of the 150 observations are displayed.

Descriptive Statistics – Summary Tables (Old Version)

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
Nev	147500	2	1935	7056
Nev	167200	2	1278	6116
Nev	395700	2	1455	14422

Missing Values

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

Procedure Options

This section describes the options available in this procedure. To find out more about using a procedure, turn to the Procedures chapter.

Variables Tab

This panel specifies the variables that will be used in the analysis.

Table Layout

Layout

This option specifies the table layout. 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

Descriptive Statistics – Summary Tables (Old Version)

- **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:

Row Variable		Data 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:

Data Variables		Column Variable		
		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

Row Variable	Data Variables		
	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	Column Variable		
	Group1	Group2	Group3
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

Descriptive Statistics – Summary Tables (Old Version)

- **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	Column Variable		
		Group1	Group2	Group3
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		Column Variable		
		Group1	Group2	Group3
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

Descriptive Statistics – Summary Tables (Old Version)

Numeric Data Variables

Data Variables

Select one or more data variable. The statistics (means, standard deviations, etc.) generated will be for the values in these variables.

Table Statistics: Count ... 90th Pctile

Check each statistic that is to be reported.

Categorical Variables

Row (Column) Variable(s)

Specify one or more categorical variables for use in table rows (columns). Each unique value in each variable will result in a separate row in the table. The data values themselves may be text (e.g. “Low, Med, High”) or numeric (e.g. “1, 2, 3”), but the data as a whole should be categorical. If more than one variable is entered, a separate table will be created for each variable.

The data values in each variable will be sorted alpha-numerically before the table rows (columns) are created. If you want the values to be displayed in a different order, specify a custom value order for the data column(s) entered here using the Column Info Table on the Data Window.

Create Other Row (Column) Variables from Numeric Data

Check this box to create tables with rows from numeric data. When checked, additional options will be displayed to specify how the numeric data will be classified into categorical variables.

If you choose to create row (column) variables from numeric data, you do not have to enter a categorical row (column) variable in the input box above (but you can). If both numeric and categorical row (column) variables are entered, a separate table and analysis will be calculated for each variable.

Numeric Variable(s) to Categorize for Use in Table Rows (Columns)

Specify one or more variables that have only numeric values to be used in rows (columns) of the table. Numeric values from these variables will be combined into a set of categories using the categorization options that follow. If more than one variable is entered, a separate table will be created for each variable.

For example, suppose you want to tabulate a variable containing individual income values into four categories: “Below 10000”, “10000 to 40000”, “40000 to 80000”, and “Over 80000”. You could select the income variable here, set **Group Numeric Data into Categories Using** to “List of Interval Upper Limits” and set the **List** to “10000 40000 80000”.

Group Numeric Data into Categories Using

Choose the method by which numeric data will be combined into categories for use in table rows or columns.

The choices are:

- **Number of Intervals, Minimum, and/or Width**

This option allows you to specify the categories by entering any combination of the three parameters:

Number of Intervals

Minimum

Width

All three are optional.

Number of Intervals

This is the number of intervals into which the values of the numeric variables are categorized. If not enough intervals are specified to reach past the maximum data value, more will be added.

Descriptive Statistics – Summary Tables (Old Version)

Range

Integer ≥ 2

Minimum

This value is used in conjunction with the Number of Intervals and Width values to construct a set of intervals into which the numeric variables are categorized. This is the minimum value of the first interval.

Range

This value must be less than the minimum data value.

Width

This value is used in conjunction with the Number of Intervals and Minimum values to construct a set of intervals into which the numeric variables are categorized. All intervals will have a width equal to this value. A data value X is in this interval if

$$\text{Lower Limit} < X \leq \text{Upper Limit}.$$

- **List of Interval Upper Limits**

This option allows you to specify the categories for the numeric variable by entering a list of interval boundaries directly, separated by blanks or commas. An interval of the form $LI < X \leq L2$ is generated for each interval. The actual number of intervals is one more than the number of items specified here.

For example, suppose you want to tabulate a variable containing individual income values into four categories: “Below 10000”, “10000 to 40000”, “40000 to 80000”, and “Over 80000”. You would set **List of Interval Upper Limits** to “10000 40000 80000”. Note that 10000 would be included in the “Below 10000” interval, but not the “10000 to 40000” interval. Also, 80000 would be included in the “40000 to 80000” interval, not the “Over 80000” interval.

Grouping (Break) Variables and Frequency Variable

Number of Grouping Variables

Select the number of grouping (break) variables. All reports and plots will be generated for each unique combination of the values of the grouping variables.

You can select up to 8 grouping variables.

Grouping Variables

Select an optional categorical grouping (or break) variable. All tables, statistical reports, and plots will be generated for each unique value of this variable.

If you specify more than one grouping variable, the tables, statistical reports, and plots will be generated for each unique combination of the values of the variables chosen.

Frequency Variable

This optional variable specifies the number of observations that each row represents. When omitted, each row represents a single observation. If your data is the result of previous summarization, you may want certain rows to represent several observations. Note that negative values are treated as a zero frequency and are omitted. Fractional values may be used. You may also think of this as a weighting variable.

Missing Values Tab

This panel lets you specify up to five missing values (besides the default of blank). For example, '0', '9', or 'NA' may be missing values in your database.

Missing Value Inclusion

Specifies whether to include observations with missing values in the tables.

Delete All indicates that you want the missing values totally ignored.

Include in Counts indicates that you want the number of missing values displayed.

Include in All indicates that you want the missing values treated just like any other category.

Missing Values

Specify individual missing values here.

Report Options Tab

The following options control the format of the reports.

Report Options

Variable Names

This option lets you select whether to display only variable names, variable labels, or both.

Value Labels

This option lets you select whether to display only values, value labels, or both. Use this option if you want the table to automatically attach labels to the values (like 1=Yes, 2=No, etc.). See the section on specifying *Value Labels* elsewhere in this manual.

Show Total

Specify whether to show row and/or column total statistics for those reports that use a by (Table Row or Table Column) variable.

Table Formatting

Column Justification

Specify whether data columns in the tables will be left or right justified.

Column Widths

Specify how the widths of columns in the contingency tables will be determined.

The options are

- **Autosize to Minimum Widths**

Each data column is individually resized to the smallest width required to display the data in the column. This usually results in columns with different widths. This option produces the most compact table possible, displaying the most data per page.

- **Autosize to Equal Minimum Width**

The smallest width of each data column is calculated and then all columns are resized to the width of the widest column. This results in the most compact table possible where all data columns have the same width. This is the default setting.

Descriptive Statistics – Summary Tables (Old Version)

- **Custom (User-Specified)**

Specify the widths (in inches) of the columns directly instead of having the software calculate them for you.

Custom Widths

Enter one or more values for the widths (in inches) of columns in the contingency tables.

- **Single Value**

If you enter a single value, that value will be used as the width for all data columns in the table.

- **List of Values**

Enter a list of values separated by spaces corresponding to the widths of each column. The first value is used for the width of the first data column, the second for the width of the second data column, and so forth. Extra values will be ignored. If you enter fewer values than the number of columns, the last value in your list will be used for the remaining columns.

Type the word "Autosize" for any column to cause the program to calculate its width for you. For example, enter "1 Autosize 0.7" to make column 1 be 1 inch wide, column 2 be sized by the program, and column 3 be 0.7 inches wide.

Wrap Column Headings onto Two Lines

Check this option to make column headings wrap onto two lines. Use this option to condense your table when your data are spaced too far apart because of long column headings.

Show Totals

Specify whether to show row and column total statistics for those reports and plots that use a categorical variable.

Use Short Statistical Names on Reports

Normally, the names of the statistical items in the reports are the complete names, such as *Standard Deviation*. Checking this option causes a shorter name, such as *Std Dev*, to be used so that more columns can be displayed on a single row.

Decimal Places

Item Decimal Places

These decimal options allow the user to specify the number of decimal places for items in the output. Your choice here will not affect calculations; it will only affect the format of the output.

- **Auto**

If one of the "Auto" options is selected, the ending zero digits are not shown. For example, if "Auto (0 to 7)" is chosen,

0.0500 is displayed as 0.05

1.314583689 is displayed as 1.314584

The output formatting system is not designed to accommodate "Auto (0 to 13)", and if chosen, this will likely lead to lines that run on to a second line. This option is included, however, for the rare case when a very large number of decimals is needed.

Plots Tab

The options on this panel control the appearance of the plots of the statistics that may be displayed. Click the plot format button to change the plot settings.

Show Plots

Check this option to display plots corresponding to the reports that are displayed. These plots are only available when the Layout is set to 4, 5, or 8.

Example 1 – Layout 1: Variable Summary Report

The data used are found in the Resale dataset. You may follow along here by making the appropriate entries or load the completed template **Example 1** by clicking on Open Example Template from the File menu of the Descriptive Statistics - Summary Tables window.

1 Open the Resale dataset.

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Click on the file **Resale.NCSS**.
- Click **Open**.

2 Open the Descriptive Statistics - Summary Tables window.

- On the menus, select **Analysis**, then **Descriptive Statistics**, then **Descriptive Statistics - Summary Tables**. The procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

3 Specify the variables.

- Select the **Variables tab**.
- Set **Layout** to **1. TABLES: One, ROWS: Data, COLUMNS: Statistics**.
- Double-click in the **Data Variables** text box. This will bring up the variable selection window.
- Select **Price** to **LotSize** from the list of variables and then click **Ok**. “Price-LotSize” will appear in the Data Variables box.

4 Specify the statistics.

- In the **Table Statistics** section, check **Count, Mean, Median, Std Dev, COV, and COD**.

5 Run the procedure.

- From the Run menu, select **Run Procedure**. Alternatively, just click the green **Run** button.

Variable Summary Report

<u>Variables</u>	Count	Mean	Median	Standard Deviation	Coef of Variation	Coef of Dispersion
Price	150	174392	158200	97656.81	0.55998	49.050
Year	150	1971.273	1973	13.84667	0.00702	0.572
Bedrooms	150	2.42	2	0.8919476	0.36857	35.000
Bathrooms	150	2.4	2.5	0.8047677	0.33532	24.800
Garage	150	1.266667	1	0.5636252	0.44497	36.000
Fireplace	150	0.96	1	0.6939818	0.72290	48.000
Quality	150	0.7316667	0.75	0.35248	0.48175	39.778
Brick	150	0.5033333	0.5	0.4157013	0.82590	68.667
TotalSqft	150	1893.38	1872.5	754.2496	0.39836	31.980
FinishSqft	150	1597.947	1496	672.1644	0.42064	35.156
LotSize	150	8366.913	8344.5	2376.334	0.28402	23.993

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

The data used are found in the Resale dataset. You may follow along here by making the appropriate entries or load the completed template **Example 2** by clicking on Open Example Template from the File menu of the Descriptive Statistics - Summary Tables window.

1 Open the Resale dataset.

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Click on the file **Resale.NCSS**.
- Click **Open**.

2 Open the Descriptive Statistics - Summary Tables window.

- On the menus, select **Analysis**, then **Descriptive Statistics**, then **Descriptive Statistics - Summary Tables**. The procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

3 Specify the variables.

- Select the **Variables tab**.
- Double-click in the **Data Variables** text box. This will bring up the variable selection window.
- Select **Price**, **TotalSqft**, and **LotSize** from the list of variables and then click **Ok**.
“Price,TotalSqft,LotSize” will appear in the Data Variables box.
- Set **Layout** to **3**. **TABLES: One**, **ROWS: Data**, **COLUMNS: Categorical**, **SUBROWS: Statistics**.
- In the **Table Statistics** section, check **Count**, **Mean**, and **Std Dev**.
- Double-click in **Column Variables** text box. This will bring up the variable selection window.
- Select **State** from the list of variables and then click **Ok**. “State” will appear in the Column Variables box.
- Set **Layout** to **1**. **TABLES: One**, **ROWS: Data**, **COLUMNS: Statistics**.

4 Specify the report format.

- Click on the **Report Options tab**.
- In Variable Names, select **Labels**.
- In Value Labels, select **Value Labels**.

5 Run the procedure.

- From the Run menu, select **Run Procedure**. Alternatively, just click the green Run button.

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

The data used are found in the Resale dataset. You may follow along here by making the appropriate entries or load the completed template **Example 3** by clicking on Open Example Template from the File menu of the Descriptive Statistics - Summary Tables window.

1 Open the Resale dataset.

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Click on the file **Resale.NCSS**.
- Click **Open**.

2 Open the Descriptive Statistics - Summary Tables window.

- On the menus, select **Analysis**, then **Descriptive Statistics**, then **Descriptive Statistics - Summary Tables**. The procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

3 Specify the variables.

- Select the **Variables** tab.
- Double-click in the **Data Variables** text box. This will bring up the variable selection window.
- Select **Bedrooms**, **Bathrooms**, **Garage**, and **Fireplace** from the list of variables and then click **Ok**. “Bedrooms-Fireplace” will appear in the Data Variables box.
- Set **Layout** to **5. TABLES: Statistics, ROWS: Data, COLUMNS: Categorical (Plots Possible)**
- In the **Table Statistics** section, check **Mean**.
- Double-click in **Column Variables** text box. This will bring up the variable selection window.
- Select **State** from the list of variables and then click **Ok**. “State” will appear in the Column Variables box.
- Set **Layout** to **1. TABLES: One, ROWS: Data, COLUMNS: Statistics**.

4 Specify the report format.

- Click on the **Report Options** tab.
- In Variable Names, select **Labels**.
- In Value Labels, select **Value Labels**.
- Set Show Total to **On Reports and Plots**

5 Run the procedure.

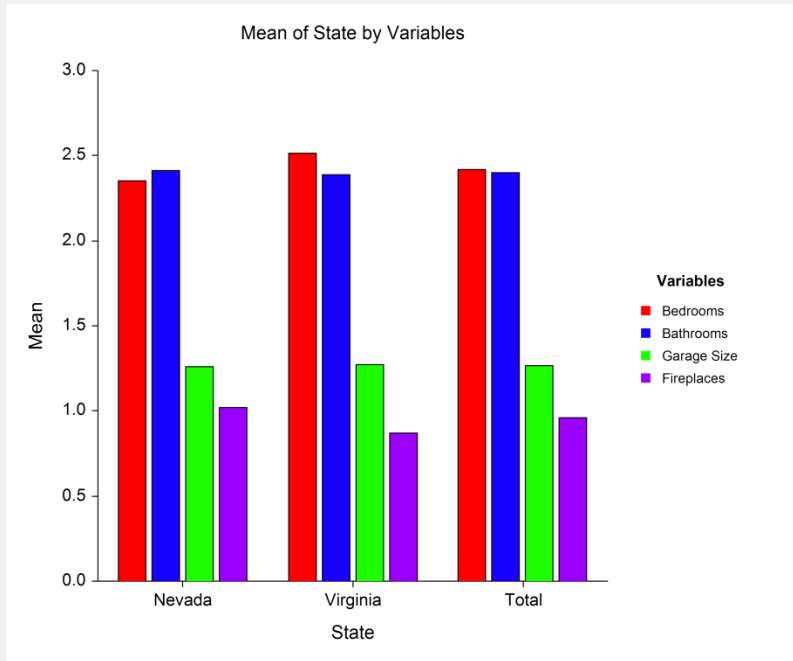
- From the Run menu, select **Run Procedure**. Alternatively, just click the green Run button.

Table and Plot of One Statistic

Table of Means

	<u>State</u>		
Variables	Nevada	Virginia	Total
Bedrooms	2.352273	2.516129	2.42
Bathrooms	2.409091	2.387097	2.4
Garage	1.261364	1.274194	1.266667
Fireplace	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

The data used are found in the Resale dataset. You may follow along here by making the appropriate entries or load the completed template **Example 4** by clicking on Open Example Template from the File menu of the Descriptive Statistics - Summary Tables window.

1 Open the Resale dataset.

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Click on the file **Resale.NCSS**.
- Click **Open**.

2 Open the Descriptive Statistics - Summary Tables window.

- On the menus, select **Analysis**, then **Descriptive Statistics**, then **Descriptive Statistics - Summary Tables**. The procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

3 Specify the variables.

- Select the **Variables tab**.
- Double-click in the **Data Variables** text box. This will bring up the variable selection window.
- Select **Price**, **FinishSqft**, and **LotSize** from the list of variables and then click **Ok**. “Price,FinishSqft-LotSize” will appear in the Data Variables box.
- Set **Layout** to **6**. **TABLES: Statistics**, **ROWS: Categorical**, **COLUMNS: Categorical**, **SUBROWS: Data**.
- Check the **Create Other Row Variables from Numeric Data** box.
- Double-click in the **Numeric Variables** text box. This will bring up the variable selection window.
- Select **TotalSqft** from the list of variables and then click **Ok**. “TotalSqft” will appear in the Numeric Variables box.
- Set the **Group Numeric Data into Categories Using** option to **List of Interval Upper Limits**.
- Set the **List** box to **1000 2000 3000**.
- Double-click in **Column Variables** text box. This will bring up the variable selection window.
- Select **State** from the list of variables and then click **Ok**. “State” will appear in the Column Variables box.

4 Specify the report format.

- Click on the **Report Options tab**.
- In **Variable Names**, select **Labels**.
- In **Value Labels**, select **Value Labels**.
- Set **Show Total** to **On Reports and Plots**. Note that no plots are displayed with this table.
- Set the **Rows Decimal Places** to **Auto (Up to 7)**.

5 Run the procedure.

- From the Run menu, select **Run Procedure**. Alternatively, just click the green **Run** button.

Multiple Y's, Two Categoricals, One Statistic

Means of Sales Price, Finished Area (Sqft), Lot Size (Sqft)				
		<u>State</u>		
<u>Total Area (Sqft)</u>		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

The data used are found in the Resale dataset. You may follow along here by making the appropriate entries or load the completed template **Example 5** by clicking on Open Example Template from the File menu of the Descriptive Statistics - Summary Tables window.

1 Open the Resale dataset.

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Click on the file **Resale.NCSS**.
- Click **Open**.

2 Open the Descriptive Statistics - Summary Tables window.

- On the menus, select **Analysis**, then **Descriptive Statistics**, then **Descriptive Statistics - Summary Tables**. The procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

3 Specify the variables.

- Select the **Variables tab**.
- Double-click in the **Data Variables** text box. This will bring up the variable selection window.
- Select **Price** from the list of variables and then click **Ok**. “Price” will appear in the Data Variables box.
- Set **Layout** to **7. TABLES: Data, ROWS: Categorical, COLUMNS: Categorical, SUBROWS: Statistics**.
- Check the **Create Other Row Variables from Numeric Data** box.
- Double-click in the **Numeric Variables** text box. This will bring up the variable selection window.
- Select **TotalSqft** from the list of variables and then click **Ok**. “TotalSqft” will appear in the Numeric Variables box.
- Set the **Group Numeric Data into Categories Using** option to **List of Interval Upper Limits**.
- Set the **List** box to **1000 2000 3000**.
- Double-click in **Column Variables** text box. This will bring up the variable selection window.
- Select **State** from the list of variables and then click **Ok**. “State” will appear in the Column Variables box.

4 Specify the report format.

- Click on the **Report Options tab**.
- In **Variable Names**, select **Labels**.
- In **Value Labels**, select **Value Labels**.
- Set **Show Total** to **On Reports and Plots**. Note that no plots are displayed with this table.
- Set the **Row Decimal Places** to **Auto (Up to 7)**.

5 Run the procedure.

- From the Run menu, select **Run Procedure**. Alternatively, just click the green **Run** button.

Descriptive Statistics – Summary Tables (Old Version)

Complete Summary for Each Data Variable

Statistical Summary of Sales Price		State		
<u>Total Area (Sqft)</u>		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

The data used are found in the Resale dataset. You may follow along here by making the appropriate entries or load the completed template **Example 6** by clicking on Open Example Template from the File menu of the Descriptive Statistics - Summary Tables window.

1 Open the Resale dataset.

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Click on the file **Resale.NCSS**.
- Click **Open**.

2 Open the Descriptive Statistics - Summary Tables window.

- On the menus, select **Analysis**, then **Descriptive Statistics**, then **Descriptive Statistics - Summary Tables**. The procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

3 Specify the variables.

- Select the **Variables** tab.
- Double-click in the **Data Variables** text box. This will bring up the variable selection window.
- Select **Price** from the list of variables and then click **Ok**. “Price” will appear in the Data Variables box.
- Set **Layout** to **8. TABLES: Data and Statistics, ROWS: Categorical, COLUMNS: Categorical (Plots Possible)**.
- Double-click in **Row Variables** text box. This will bring up the variable selection window.
- Select **State** from the list of variables and then click **Ok**. “State” will appear in the Row Variables box.
- Check the **Create Other Column Variables from Numeric Data** box.
- Double-click in the **Numeric Variables** text box. This will bring up the variable selection window.
- Select **TotalSqft** from the list of variables and then click **Ok**. “TotalSqft” will appear in the Numeric Variables box.
- Set the **Group Numeric Data into Categories Using** option to **List of Interval Upper Limits**.
- Set the **List** box to **1000 2000 3000**.

4 Change the report format.

- Click on the **Report Options** tab.
- In **Variable Names**, select **Labels**.
- In **Value Labels**, select **Value Labels**.
- Set **Show Total** to **On Reports and Plots**. Note that no plots are displayed with this table.
- Set the **Row Decimal Places** to **Auto (Up to 7)**.
- Set the **Mean, Sum** Decimal Places to **0**.

5 Change the plot format.

- Click on the **Plots** tab.
- Check **Interactive** box in the upper, right-hand corner of the Plot Format button.

6 Run the procedure.

- From the Run menu, select **Run Procedure**. Alternatively, just click the green **Run** button.
- When the Bar Chart Format window is displayed, click the **Group Axis** tab.
- In the **Ticks** section of the window, click the **Layout** icon of the **Lower Axis Labels**. This button is the second icon to the right.
- Set the **Rotation Angle** to **39**. Click **OK** twice.

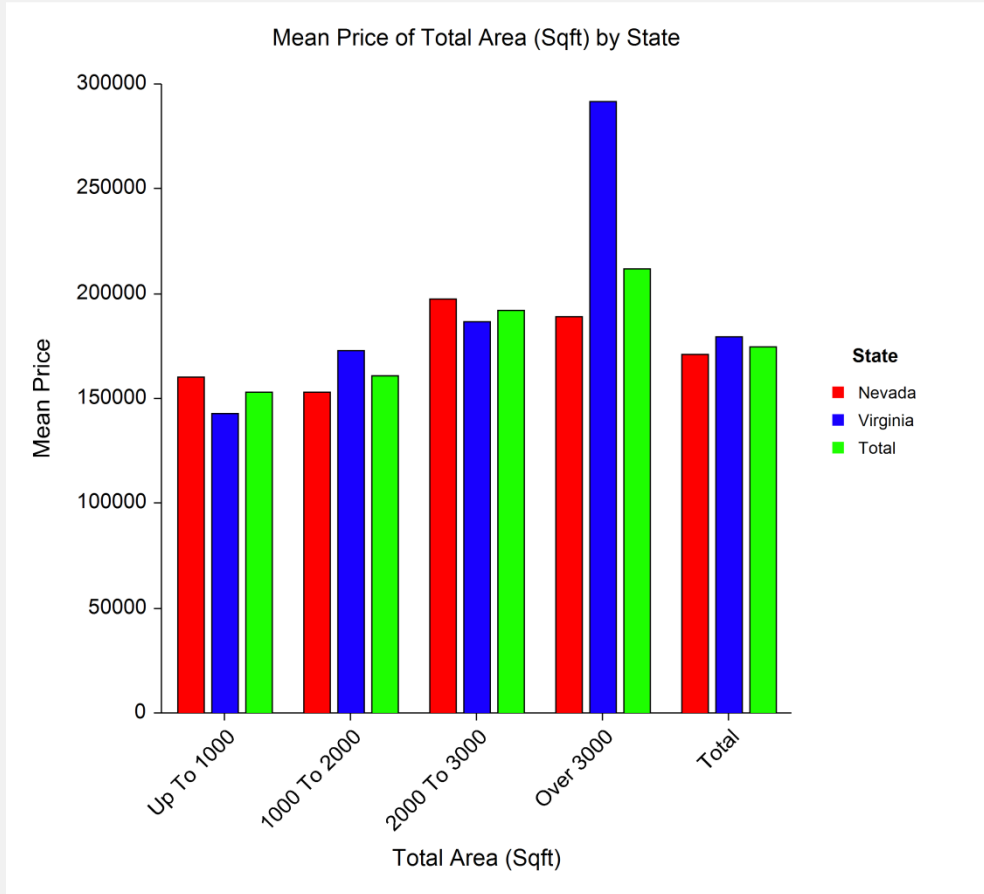
One Data Variable and Statistic, Two Categories

Means of Sales Price

Total Area (Sqft)

State	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



Example 7 – List Format

The data used are found in the Resale dataset. You may follow along here by making the appropriate entries or load the completed template **Example 7** by clicking on Open Example Template from the File menu of the Descriptive Statistics - Summary Tables window.

1 Open the Resale dataset.

- From the File menu of the NCSS Data window, select **Open Example Data**.
- Click on the file **Resale.NCSS**.
- Click **Open**.

2 Open the Descriptive Statistics - Summary Tables window.

- On the menus, select **Analysis**, then **Descriptive Statistics**, then **Descriptive Statistics - Summary Tables**. The Descriptive Statistics - Summary Tables procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

3 Specify the variables.

- Select the **Variables** tab.
- In Table Layout, select **9. DATA SUMMARY LIST**.
- Double-click in the **Data Variables** text box. This will bring up the variable selection window.
- Select **Price** from the list of variables and then click **Ok**. “Price” will appear in the Data Variables box.
- Double-click in the **Row Variables** text box. This will bring up the variable selection window.
- Select **Neighborhood** from the list of variables and then click **Ok**. “Neighborhood” will appear in the Row Variables box.
- Set the **Number of Grouping Variables** to **2**.
- Double-click in the **Grouping Variable 1** text box. This will bring up the variable selection window.
- Select **State** from the list of variables and then click **Ok**.
- Double-click in the **Grouping Variable 2** text box. This will bring up the variable selection window.
- Select **City** from the list of variables and then click **Ok**.

4 Click on the Report Options tab.

- In Variable Names, select **Names**.
- In Value Labels, select **Data Values**.

5 Run the procedure.

- From the Run menu, select **Run Procedure**. Alternatively, just click the green **Run** button.

List Format Report

Summary List					
State	City	Neighborhood	Price Count	Price Mean	Price StdDev
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.