

Chapter 500

Frequency Tables

Introduction

This procedure produces tables of frequency counts and percentages for categorical and continuous variables. This procedure serves as a summary reporting tool and is often used to analyze survey data. This procedure also calculates the **multinomial chi-square** test.

Frequency Tables

Frequency tables are generally produced on individual variables. For categorical data, the table records the number of observations (the frequency) for each unique value of the variable. For continuous data, you must specify a set of intervals (or bins). The frequency table records the number of observations falling in each interval.

Frequency tables are useful for analyzing categorical data and for screening data for data entry errors.

Types of Categorical Variables

Note that we will refer to two types of categorical variables: *Categorical* and *Grouping or Break*. *Grouping* variables are used to split a database into subgroups. A separate set of reports is generated for each unique set of values of the *Grouping* variables. The values of a *Categorical* variable are used to define the rows of the frequency table.

Data Structure

The data below are a subset of the real estate sales (**Resale**) dataset provided with the software. This 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

Missing values may be ignored or included in the table's counts and percentages.

Example 1 – Standard Frequency Tables

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 Frequency Tables procedure options

- Find and open the **Frequency Tables** 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

Type of Data Input **Column(s) in the Database**
 Categorical Variables..... **State-City**

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.

Output

Frequency Distribution of State

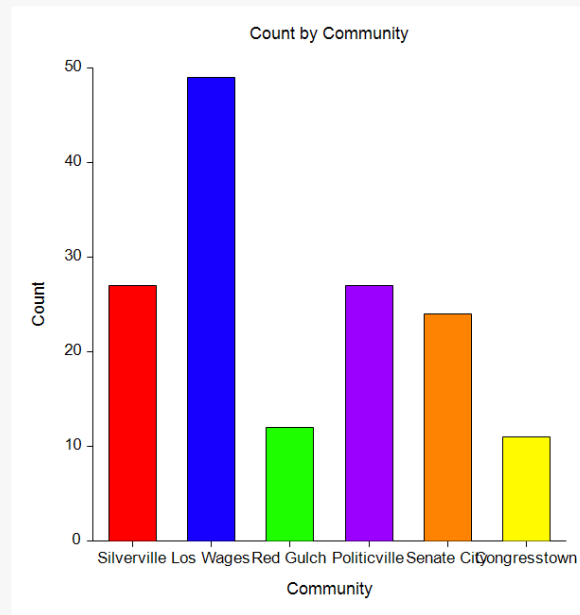
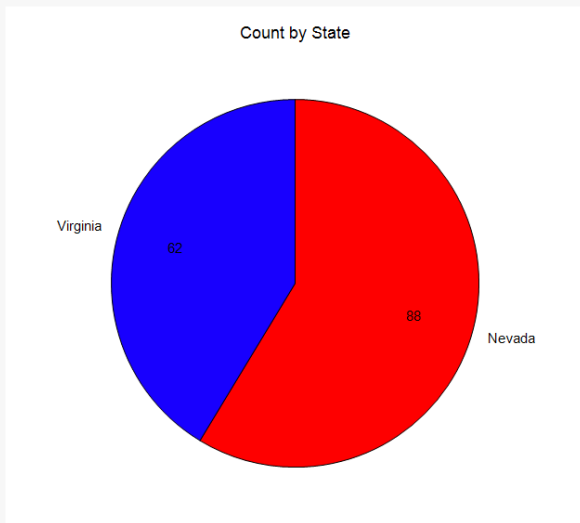
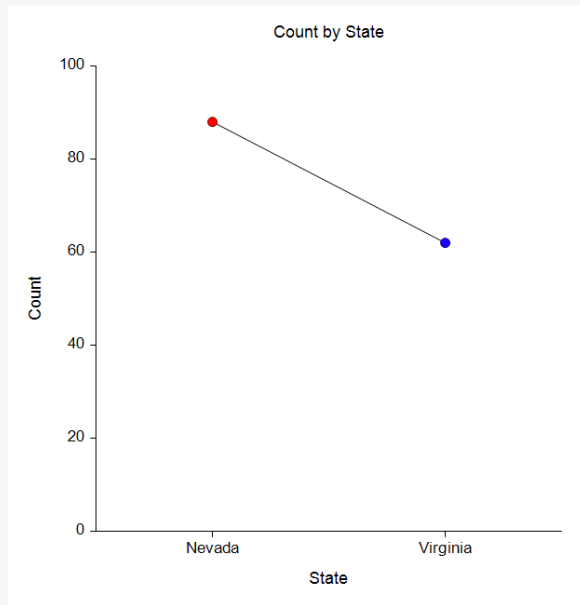
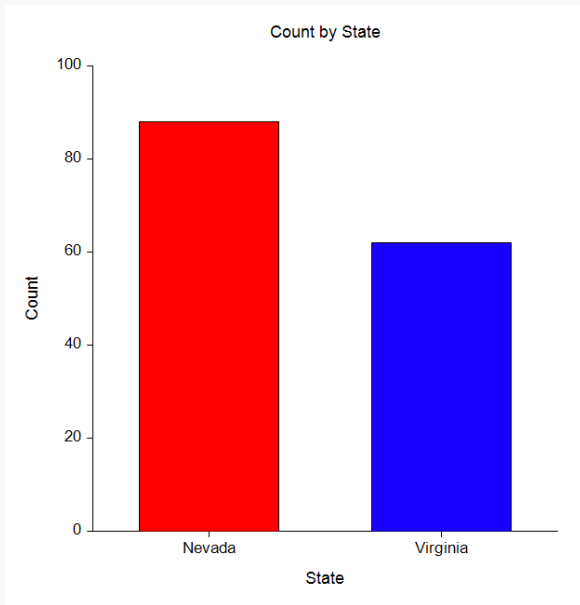
| State | Count | Cumulative Count | Percent | Cumulative Percent | Graph of Percent |
|----------|-------|------------------|---------|--------------------|------------------|
| Nevada | 88 | 88 | 58.67% | 58.67% | |
| Virginia | 62 | 150 | 41.33% | 100.00% | |

Frequency Tables

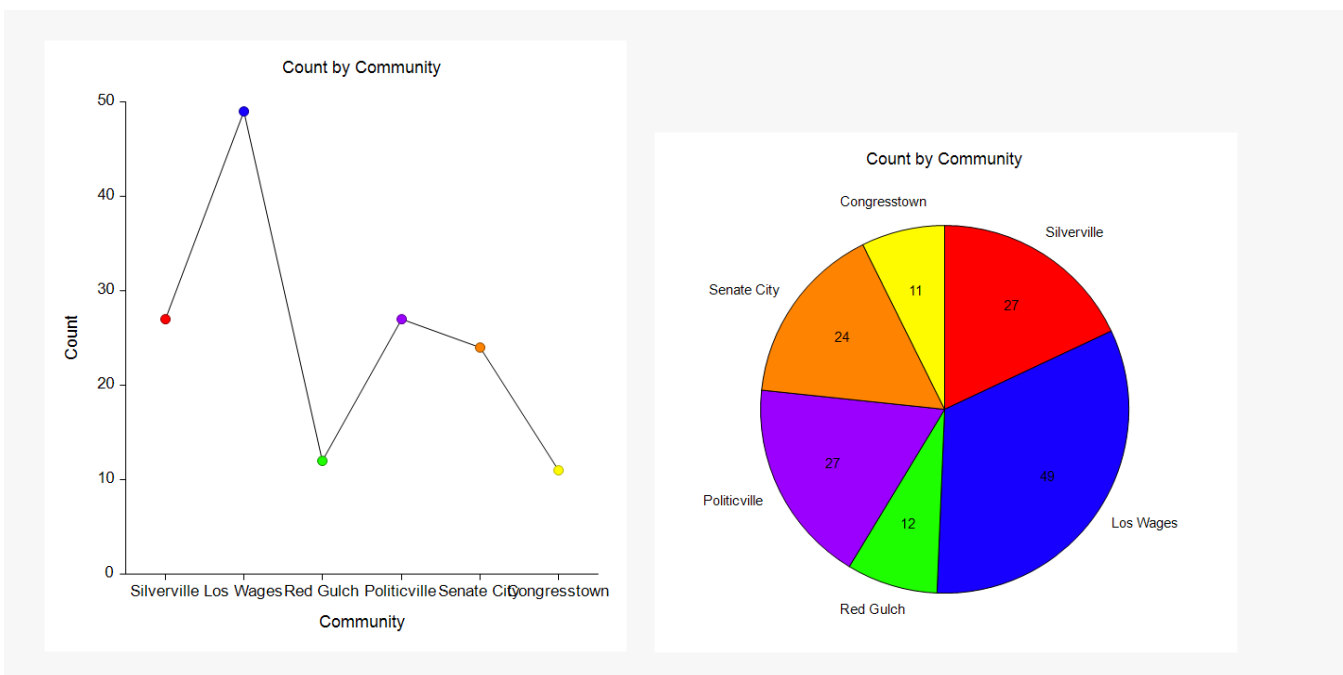
Frequency Distribution of Community

| Community | Count | Cumulative Count | Percent | Cumulative Percent | Graph of Percent |
|--------------|-------|------------------|---------|--------------------|------------------|
| Silverville | 27 | 27 | 18.00% | 18.00% | |
| Los Wages | 49 | 76 | 32.67% | 50.67% | |
| Red Gulch | 12 | 88 | 8.00% | 58.67% | |
| Politicville | 27 | 115 | 18.00% | 76.67% | |
| Senate City | 24 | 139 | 16.00% | 92.67% | |
| Congresstown | 11 | 150 | 7.33% | 100.00% | |

Frequency Distribution Charts



Frequency Tables



This report presents the counts (the frequencies), percentages, and a rough bar graph of the data. Note that the bar graph is constructed so that each “|” is worth 2.5 percentage points. By default, distribution bar charts, line charts, and pie charts are also output.

The frequency table and plots can be modified using the settings below to display the values in descending or ascending order, sorted by counts.

4 Modify the Category Order

- The settings for this example are listed below and are stored in the **Example 1b** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

```

Reports Tab
-----
Category Order (under Frequency Distribution Table Report) ..... Sort by Count (Descending)
    
```

5 Run the procedure

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

Output

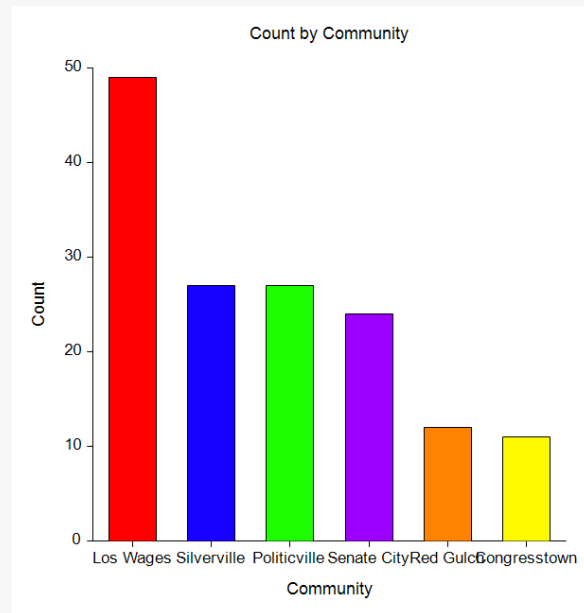
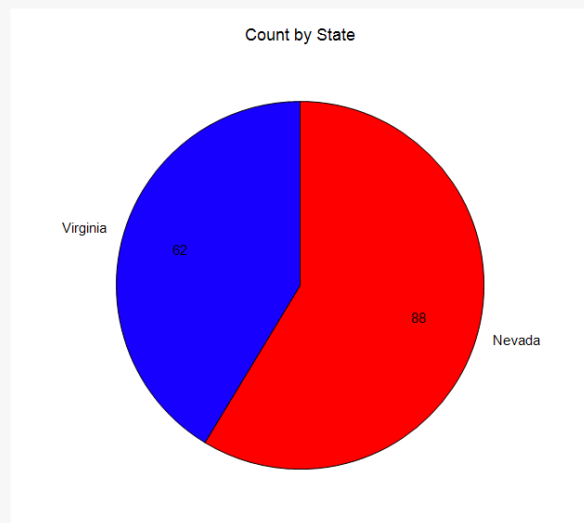
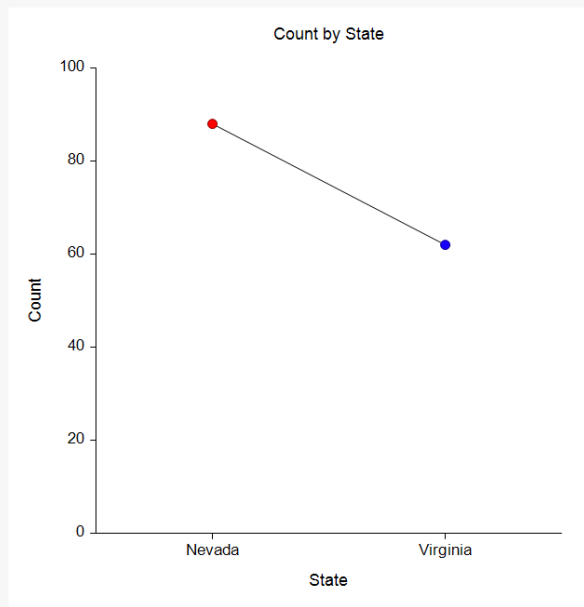
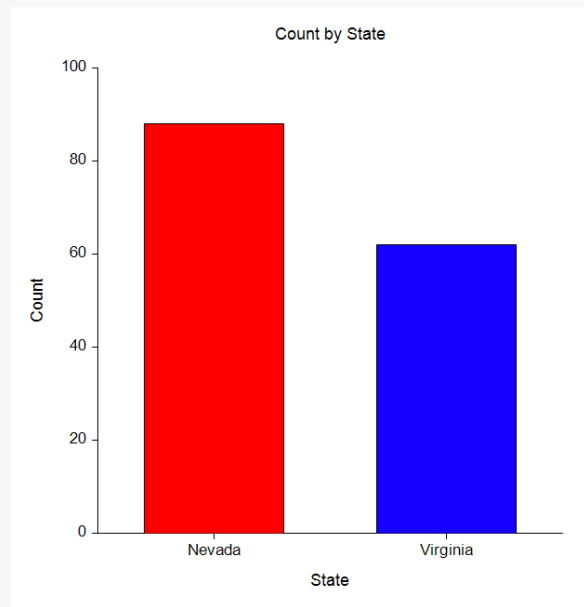
| Frequency Distribution of State | | | | | |
|---------------------------------|-------|------------------|---------|--------------------|------------------|
| State | Count | Cumulative Count | Percent | Cumulative Percent | Graph of Percent |
| Nevada | 88 | 88 | 58.67% | 58.67% | |
| Virginia | 62 | 150 | 41.33% | 100.00% | |

Frequency Tables

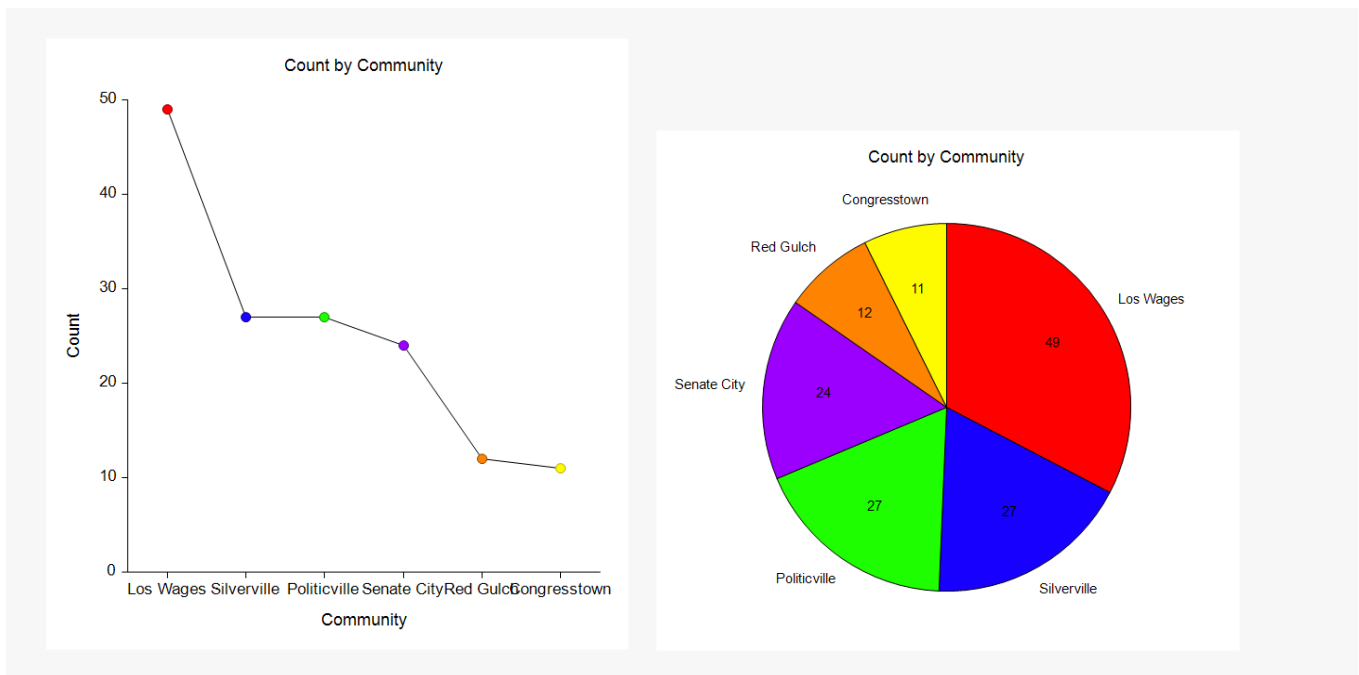
Frequency Distribution of Community

| Community | Count | Cumulative Count | Percent | Cumulative Percent | Graph of Percent |
|--------------|-------|------------------|---------|--------------------|------------------|
| Los Wages | 49 | 49 | 32.67% | 32.67% | |
| Silverville | 27 | 76 | 18.00% | 50.67% | |
| Politicville | 27 | 103 | 18.00% | 68.67% | |
| Senate City | 24 | 127 | 16.00% | 84.67% | |
| Red Gulch | 12 | 139 | 8.00% | 92.67% | |
| Congresstown | 11 | 150 | 7.33% | 100.00% | |

Frequency Distribution Charts



Frequency Tables



The reports and plots are now sorted by count, high to low. Note that the plots follow the table because of the Category Order setting on the Plots tab. This setting can be changed to make the plots different from the reports.

Example 2 – Multinomial Test

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 Frequency Tables procedure options

- Find and open the **Frequency Tables** 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

Type of Data Input **Column(s) in the Database**
 Categorical Variables..... **State**

Reports Tab

Frequency Distribution Table Report **Checked**
 Multinomial Test Report..... **Checked**
 Expected Values..... **60 40**

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.

Output

| State | Count | Cumulative Count | Percent | Cumulative Percent | Graph of Percent |
|----------|-------|------------------|---------|--------------------|------------------|
| Nevada | 88 | 88 | 58.67% | 58.67% | |
| Virginia | 62 | 150 | 41.33% | 100.00% | |

Frequency Tables

Multinomial Test of State

| State | Count | Expected Count | Actual Percent | Expected Percent | Chi-Square Amount |
|----------|-------|----------------|----------------|------------------|-------------------|
| Nevada | 88 | 90.00 | 58.7% | 60.0% | 0.0444 |
| Virginia | 62 | 60.00 | 41.3% | 40.0% | 0.0667 |

Chi-Square = 0.1111 with df = 1 Probability Level = 0.738883

The first table is the standard frequency table. The second table presents the results of the multinomial test. Note that in this case, the test is not significant.

Count

The number of observations (rows) for each variable. This is the value of O_i .

Expected Count

The number of observations (rows) that would be obtained if the hypothesized proportions were followed exactly. This is the value of E_i .

Actual Percent

The percent that this category is of the total.

Expected Percent

The percentage that this category would have if the hypothesized proportions were followed exactly.

Chi-Square Amount

The amount that this line contributes to the Chi-square statistic. This is equal to

$$CS_i = \frac{(O_i - E_i)^2}{E_i}$$

where O_i is the actual count and E_i is the expected count of the i^{th} category.

Chi-Square

This is the value of the chi-square test statistic.

$$\chi_{k-1}^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

df

The degrees of freedom of the above test statistic. This is equal to the number of categories minus one.

Probability Level

This is the significance level of the multinomial test. If you are testing at an alpha of 0.05, you would reject the null hypothesis that the hypothesized proportions are true if this value is less than 0.05.

Example 3 – Frequency Tables and Multinomial Tests from Summary Data

Using the same variables and data as in Example 2, an alternate way to enter the data for this procedure is to enter the counts for each category directly. This is done by setting **Type of Data Input** to **Table of Counts**.

Setup

To run this example, complete the following steps:

1 Specify the Frequency Tables procedure options

- Find and open the **Frequency Tables** 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

| | |
|---------------------------|------------------------|
| Type of Data Input | Table of Counts |
| Categories (Rows) | 2 |
| Variables (Columns) | 1 |

Table of Counts

| | |
|------------------------|-----------------|
| Variable 1 Label | State |
| Category 1 Label | Nevada |
| Category 2 Label | Virginia |
| Category 1 Count | 88 |
| Category 2 Count | 62 |

Reports Tab

| | |
|---|----------------|
| Frequency Distribution Table Report | Checked |
| Multinomial Test Report | Checked |
| Expected Values | 60 40 |

2 Run the procedure

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

Output

Frequency Distribution of State

| State | Count | Cumulative Count | Percent | Cumulative Percent | Graph of Percent |
|----------|-------|------------------|---------|--------------------|------------------|
| Nevada | 88 | 88 | 58.67% | 58.67% | |
| Virginia | 62 | 150 | 41.33% | 100.00% | |

Multinomial Test of State

| State | Count | Expected Count | Actual Percent | Expected Percent | Chi-Square Amount |
|----------|-------|----------------|----------------|------------------|-------------------|
| Nevada | 88 | 90.00 | 58.7% | 60.0% | 0.0444 |
| Virginia | 62 | 60.00 | 41.3% | 40.0% | 0.0667 |

Chi-Square = 0.1111 with df = 1 Probability Level = 0.738883

The output is exactly the same as in Example 2.

Example 4 – Tables and Two-Factor Bar Charts of Counts and Percentages Across Variables

This example will show how to obtain some of the other table formats that are available from this procedure. The data for this example are found in the Resale dataset.

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 Frequency Tables procedure options

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

Variables Tab

Type of Data Input **Column(s) in the Database**
 Categorical Variables..... **Garage-Fireplace,Brick**

Reports Tab

Frequency Distribution Table Report **Unchecked**
 Counts Table **Checked**
 Row Percentages Table **Checked**
 Column Percentages Table **Checked**
 Table Percentages Table..... **Checked**

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.

Frequency Tables

Output

Counts Table

| Values | Variables | | | Total |
|--------------|-------------|------------|-------------|-------|
| | Garage Size | Fireplaces | Brick Ratio | |
| 0 | 7 | 39 | 51 | 97 |
| 0.5 | 0 | 0 | 47 | 47 |
| 1 | 98 | 78 | 52 | 228 |
| 2 | 43 | 33 | 0 | 76 |
| 3 | 2 | 0 | 0 | 2 |
| Total | 150 | 150 | 150 | 450 |

Row Percentages Table

| Values | Variables | | | Total |
|--------------|-------------|------------|-------------|--------|
| | Garage Size | Fireplaces | Brick Ratio | |
| 0 | 7.2% | 40.2% | 52.6% | 100.0% |
| 0.5 | 0.0% | 0.0% | 100.0% | 100.0% |
| 1 | 43.0% | 34.2% | 22.8% | 100.0% |
| 2 | 56.6% | 43.4% | 0.0% | 100.0% |
| 3 | 100.0% | 0.0% | 0.0% | 100.0% |
| Total | 33.3% | 33.3% | 33.3% | 100.0% |

Column Percentages Table

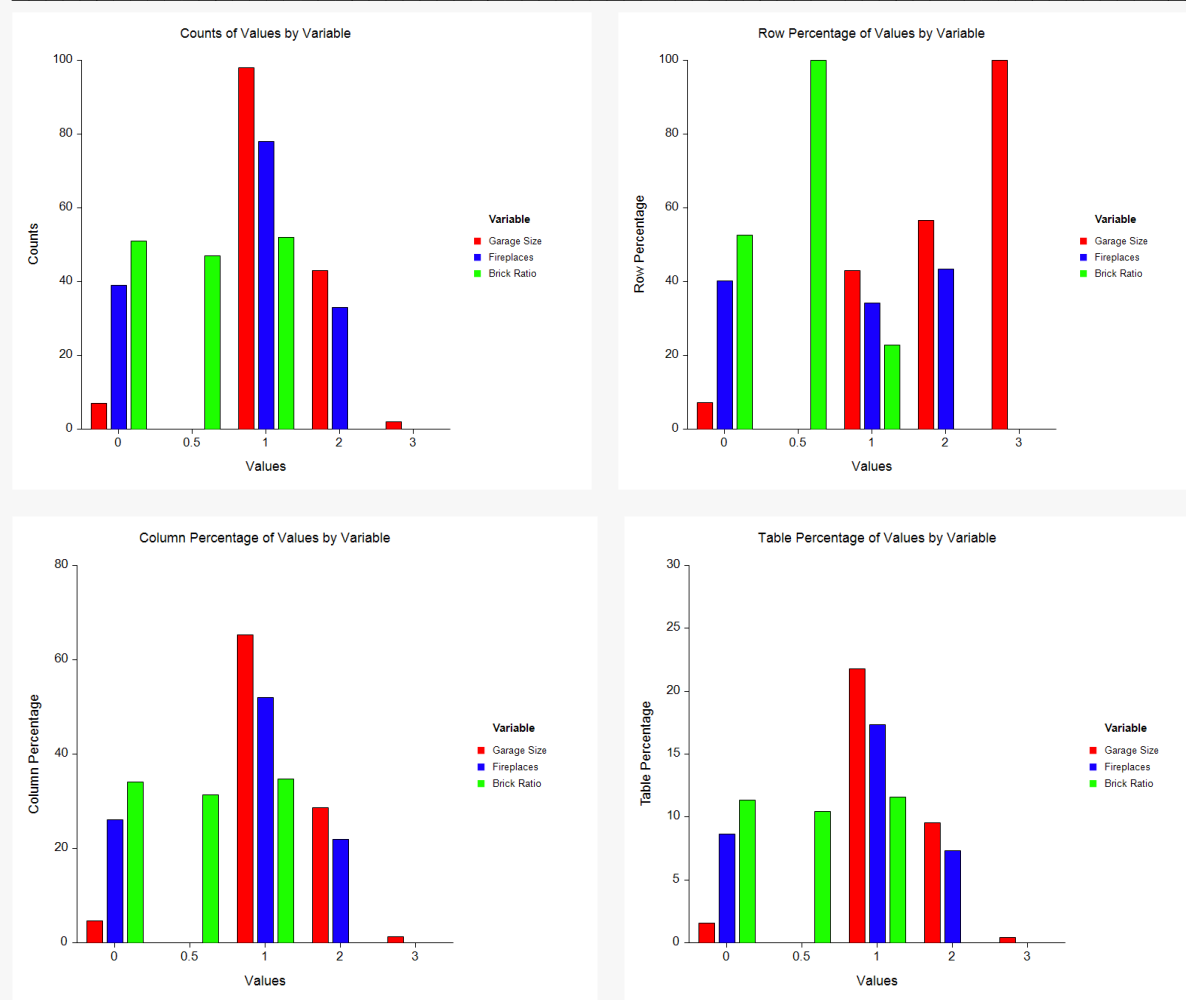
| Values | Variables | | | Total |
|--------------|-------------|------------|-------------|--------|
| | Garage Size | Fireplaces | Brick Ratio | |
| 0 | 4.7% | 26.0% | 34.0% | 21.6% |
| 0.5 | 0.0% | 0.0% | 31.3% | 10.4% |
| 1 | 65.3% | 52.0% | 34.7% | 50.7% |
| 2 | 28.7% | 22.0% | 0.0% | 16.9% |
| 3 | 1.3% | 0.0% | 0.0% | 0.4% |
| Total | 100.0% | 100.0% | 100.0% | 100.0% |

Frequency Tables

Table Percentages Table

| Values | Variables | | | Total |
|--------------|--------------|--------------|--------------|---------------|
| | Garage Size | Fireplaces | Brick Ratio | |
| 0 | 1.6% | 8.7% | 11.3% | 21.6% |
| 0.5 | 0.0% | 0.0% | 10.4% | 10.4% |
| 1 | 21.8% | 17.3% | 11.6% | 50.7% |
| 2 | 9.6% | 7.3% | 0.0% | 16.9% |
| 3 | 0.4% | 0.0% | 0.0% | 0.4% |
| Total | 33.3% | 33.3% | 33.3% | 100.0% |

Two-Factor Charts Across Variables



This report presents tables containing counts and various percentages for all the variables selected. It also provides two-way bar charts of the counts and percentages.

An alternate way to enter this summarized data is to set **Type of Data Input** to **Table of Counts** and enter the titles, labels, and counts into the table on the input window.

Frequency Tables

4 Modify the Data Input Type

- The settings for this example are listed below and are stored in the **Example 4b** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Variables Tab

| | |
|---------------------------|--|
| Type of Data Input | Table of Counts |
| Categories (Rows) | 5 |
| Variables (Columns) | 3 |
| Table of Counts | Enter Labels and Counts from the Counts Table |

5 Run the procedure

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

The output will be exactly the same as that displayed above.

Example 5 – Combined Table

This example will show how to obtain a combined table of various counts and percentages. The data for this example are found in the Resale dataset.

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 Frequency Tables procedure options

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

| | |
|---|----------------------------------|
| Variables Tab | |
| Type of Data Input | Column(s) in the Database |
| Categorical Variables..... | Garage-Fireplace,Brick |
| Reports Tab | |
| Frequency Distribution Table Report | Unchecked |
| Show Combined Table | Checked |
| Counts | Checked |
| Row Percentages | Checked |
| Column Percentages | Checked |
| Table Percentages..... | Checked |
| Report Options (<i>in the Toolbar</i>) | |
| Variable Labels | Column Labels |
| Data Labels..... | Value Labels |

3 Run the procedure

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

Output

| Values | | Variables | | | Total |
|--------------|-----------------|-------------|------------|-------------|--------|
| | | Garage Size | Fireplaces | Brick Ratio | |
| 0 | Count | 7 | 39 | 51 | 97 |
| | % within Row | 7.2% | 40.2% | 52.6% | 100.0% |
| | % within Column | 4.7% | 26.0% | 34.0% | 21.6% |
| | % of Total | 1.6% | 8.7% | 11.3% | 21.6% |
| 0.5 | Count | 0 | 0 | 47 | 47 |
| | % within Row | 0.0% | 0.0% | 100.0% | 100.0% |
| | % within Column | 0.0% | 0.0% | 31.3% | 10.4% |
| | % of Total | 0.0% | 0.0% | 10.4% | 10.4% |
| 1 | Count | 98 | 78 | 52 | 228 |
| | % within Row | 43.0% | 34.2% | 22.8% | 100.0% |
| | % within Column | 65.3% | 52.0% | 34.7% | 50.7% |
| | % of Total | 21.8% | 17.3% | 11.6% | 50.7% |
| 2 | Count | 43 | 33 | 0 | 76 |
| | % within Row | 56.6% | 43.4% | 0.0% | 100.0% |
| | % within Column | 28.7% | 22.0% | 0.0% | 16.9% |
| | % of Total | 9.6% | 7.3% | 0.0% | 16.9% |
| 3 | Count | 2 | 0 | 0 | 2 |
| | % within Row | 100.0% | 0.0% | 0.0% | 100.0% |
| | % within Column | 1.3% | 0.0% | 0.0% | 0.4% |
| | % of Total | 0.4% | 0.0% | 0.0% | 0.4% |
| Total | Count | 150 | 150 | 150 | 450 |
| | % within Row | 33.3% | 33.3% | 33.3% | 100.0% |
| | % within Column | 100.0% | 100.0% | 100.0% | 100.0% |
| | % of Total | 33.3% | 33.3% | 33.3% | 100.0% |

This report presents a single table that combines the counts and percentages of all variables selected.

An alternate way to enter this summarized data is to set **Type of Data Input** to **Table of Counts** and enter the titles, labels, and counts into the table on the input window.

4 Modify the Data Input Type

- The settings for this example are listed below and are stored in the **Example 5b** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

| Variables Tab | |
|---------------------------|--|
| Type of Data Input | Table of Counts |
| Categories (Rows) | 5 |
| Variables (Columns) | 3 |
| Table of Counts | Enter Labels and Counts from the Combined Table |

Frequency Tables

5 Run the procedure

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

The output will be exactly the same as that displayed above.