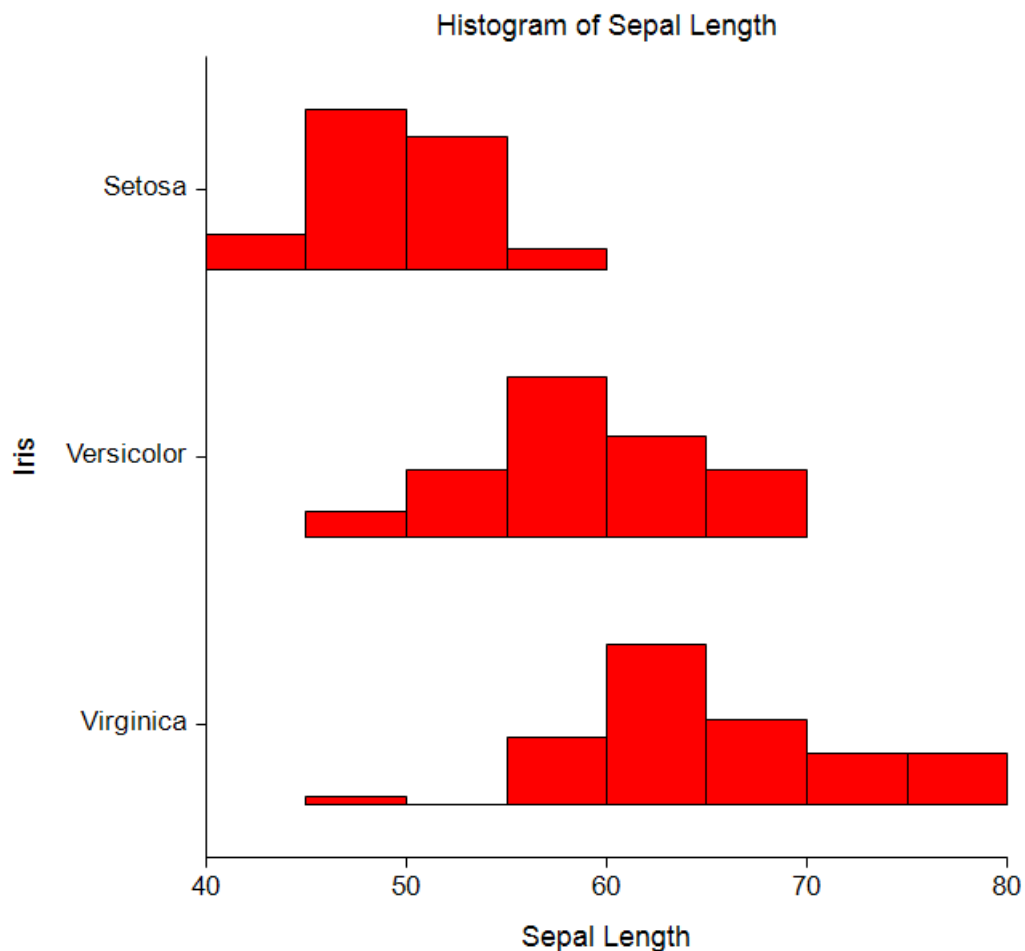Chapter 151

Histograms – Comparative

Introduction

A histogram displays the frequency distribution of a set of data values. These procedures facilitate the visual comparison of the distributions of two or more groups through comparing side-by-side histograms.

Histogram of Sepal Length

Data Structure

A histogram is constructed from a numeric variable. A second variable may be used to divide the first variable into groups (e.g., age group or gender). In the two-factor procedure, a third variable may be used to divide the groups into subgroups.
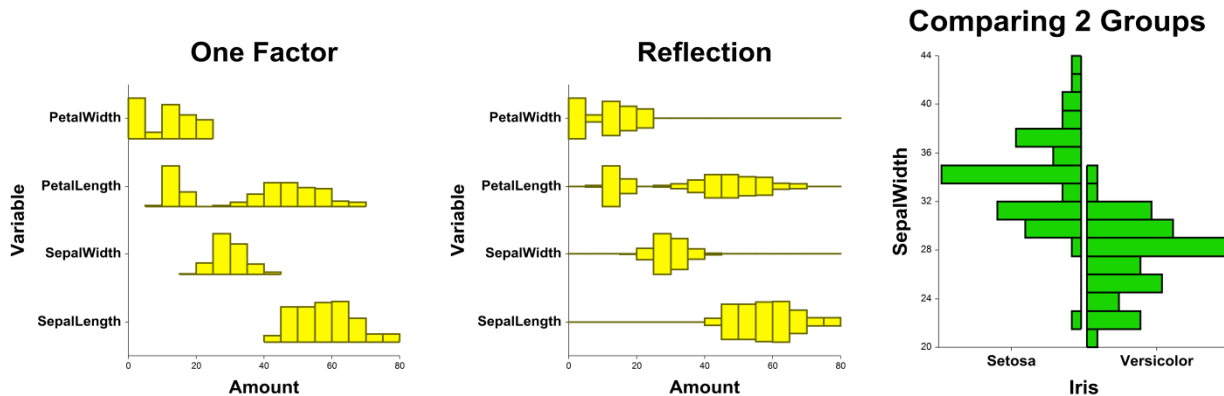
Histogram Window Options

This section describes the specific options available on the Histogram window, which is displayed when the Histogram button is clicked. Common options, such as axes, labels, legends, and titles are documented in the Graphics Components chapter.
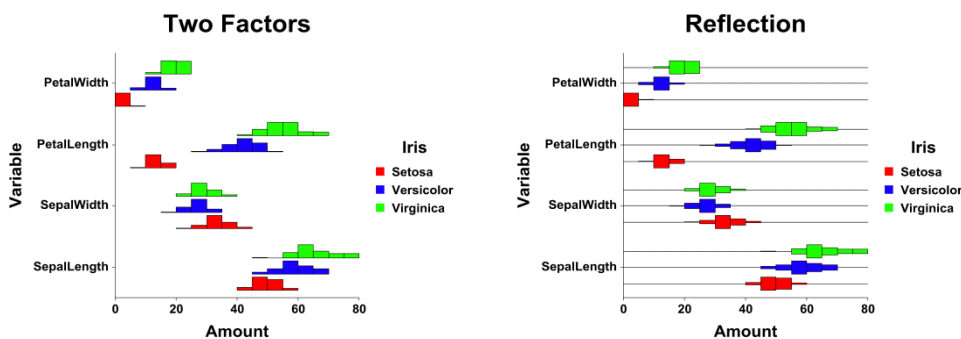
Histogram Tab

Format Section

You can modify the color of the histogram and its outline using the options in this section. You can also change the direction and extent of the bars.
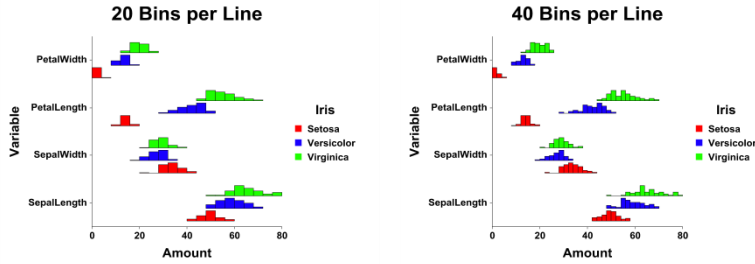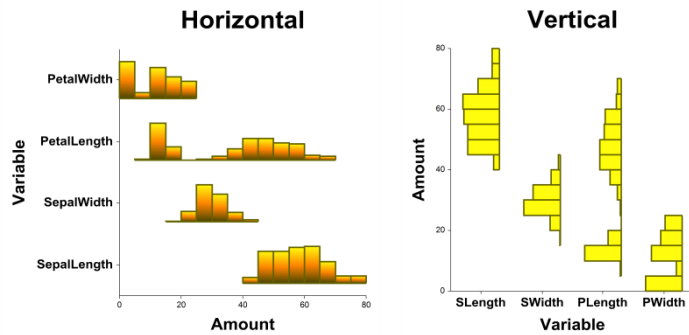
One Factor

Two Factors

Bins Section

You can specify the number of bins (bars) of the histogram in several ways. In this case, the number of bins is for the whole axis, not just the histogram that is showing. Hence, you need to increase the number of bins to about double what they would normally be.
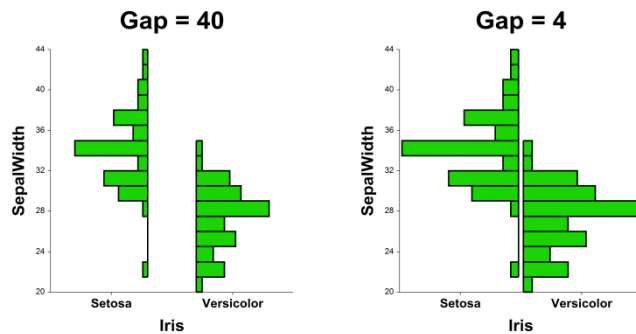

Layout Tab

Orientation Section

You can orient the histogram horizontally or vertically.
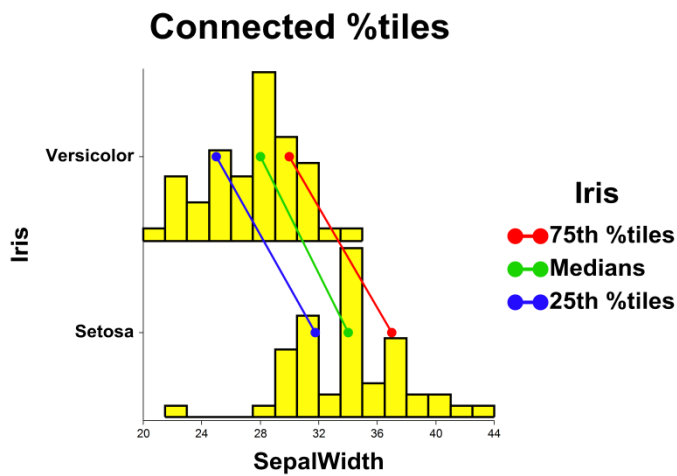

Object Spacing and Size Section

You can change the size of the gap between individual histograms.

Connecting Lines Tab

Connect Between Groups Section

You can add reference lines at group means and percentiles.

Titles, Legend, Numeric Axis, Group Axis, Grid Lines, and Background Tabs

Details on setting the options in these tabs are given in the Graphics Components chapter.

Example 1 – Creating a Comparative Histogram Plot

This section presents an example of how to generate a comparative histogram. The data used are from the Fisher dataset. We will create histograms of the *SepalLength* variable, grouping on the type of iris.

Setup

To run this example, complete the following steps:

1 Open the Fisher example dataset

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

2 Specify the Histograms – Comparative procedure options

- Find and open the **Histograms – Comparative** 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

Data Variable(s)...**SepalLength**
Horizontal (Group) Variable ....**Iris**

Report Options (*in the Toolbar*)

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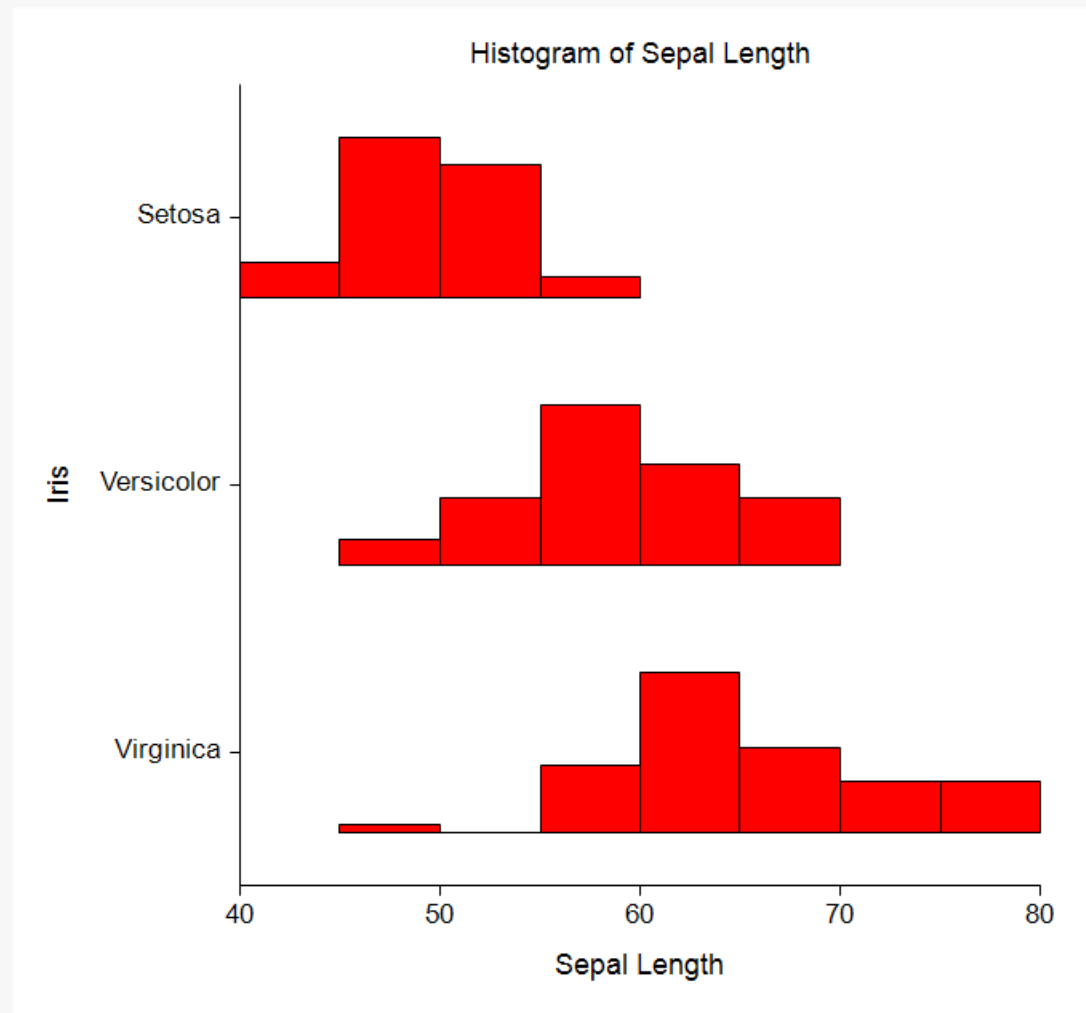Data Labels...**Value Labels**

3 Run the procedure

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

Comparative Histogram Output

Histograms
_____

Example 2 – Creating a Comparative Histogram with Subgroups

This section presents an example of how to generate a histogram plot with subgroups. The data used are from the fictitious Tree dataset. We will create histograms of the *Diameter* variable, grouping on *Species*, with subgroups according to *Sunlight*.

Setup

To run this example, complete the following steps:

1   Open the Tree example dataset

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

2   Specify the Histograms - Comparative (2 Factors) procedure options

- Find and open the **Histograms - Comparative (2 Factors)** 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

  Data Variable(s)...**Diameter**
  Horizontal (Group) Variable ....**Species**
  Legend (Subgroup) Variable............................**Sunlight**

  Report Options (*in the Toolbar*)

  Data Labels...**Value Labels**

3   Run the procedure

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

Comparative Histogram Output (2 Factors)

Histograms