# Chapter 154

Density Plots

Introduction

When analyzing data, you often need to study the characteristics of a single group of numbers, observations, or measurements. You might want to know the center and the spread about this central value. You might want to investigate extreme values (referred to as outliers) or study the distribution or pattern of the data values. Several plots are available to allow you to study the distribution. One such plot is the density plot.

Density Plot of SepalLength
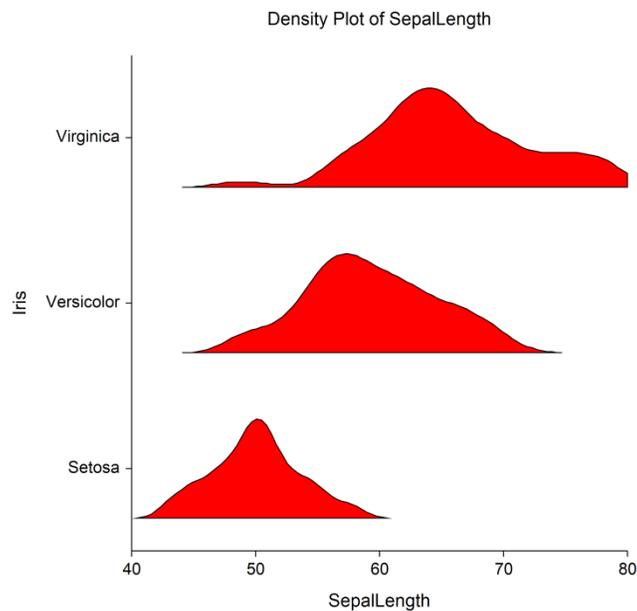
Density Plot

The *Density Plot* shows the smoothed distribution of the points along the numeric axis. The peaks of the density plot are at the locations where there is the highest concentration of points.

Data Structure

A density plot 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.

Procedure Options

This section describes the options available in this procedure.

Variables Tab

This panel specifies which columns are used to create the density plot.

Variables – Simple Density Plots

Data Variable(s)

This option lets you designate which variables are plotted. If more than one Data Variable is designated and no Horizontal (Group) Variable is selected, a set of density plots will be displayed on a single chart, one box for each variable. If more than one variable is designated and a Horizontal (Group) Variable is selected, a separate density plot will be drawn for each variable.

Horizontal (Group) Variable

Designates an optional variable used to separate the observations into groups. The groups defined by this variable will all appear on the same plot.

Frequency Variable

Specify an optional frequency (count) variable. This variable contains integers that represent the number of observations (frequency) associated with each observation. If left blank, each observation has a frequency of one.

Data Label Variable

A data label is text that is displayed beside each outlier point. This option designates the variable containing the data labels. The values may be text or numeric.

Break Variable

Select an optional break (categorical) variable. A separate plot will be generated for each unique value of this variable. If you specify more than one break variable, a separate plot will be generated for each unique combination of the values of the variables specified.

Variables – Two-Factor Box Plots

Data Variable(s)

This option lets you designate which variables are plotted. If more than one Data Variable is designated and no Legend (Subgroup) Variable is selected, the variables will become the legend (subgroup) levels. If more than one variable is designated and a Legend (Subgroup) Variable is selected, a separate plot will be drawn for each variable.

Horizontal (Group) Variable

Designates an optional variable used to separate the observations into groups. An individual density will be displayed for each unique combination of this variable with the Legend (Subgroup) Variable.

Legend (Subgroup) Variable

Designates an optional variable used to separate the observations into subgroups. An individual density will be displayed for each combination of this variable with the Horizontal (Group) Variable. The levels of this variable will be shown in the legend.

Frequency Variable

Specify an optional frequency (count) variable. This variable contains integers that represent the number of observations (frequency) associated with each observation. If left blank, each observation has a frequency of one.

Data Label Variable

A data label is text that is displayed beside each outlier point. This option designates the variable containing the data labels. The values may be text or numeric.

Break Variable

Select an optional break (categorical) variable. A separate plot will be generated for each unique value of this variable. If you specify more than one break variable, a separate plot will be generated for each unique combination of the values of the variables specified.

Format Options

Variable Names

This option specifies whether the column names or column labels are used on the chart.

Value Labels

This option specifies whether the actual values or the value labels are used to label the group level labels of the plot.

Density Plot Format

Format

Click the format button to change the plot settings (see Density Plot Window Options below).

Edit During Run

Checking this option will cause the plot format window to appear when the procedure is run. This allows you to modify the format of the graph with the actual data.

Density Plot Window Options

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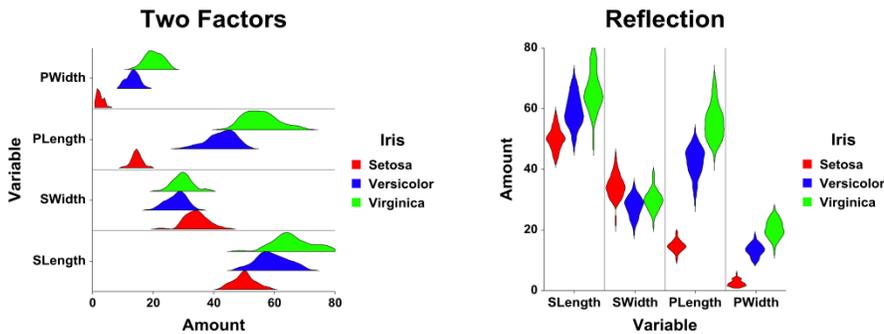
Density Plot Tab

Density Section

You can modify various density calculation parameters, the color, and the direction of the density faces.
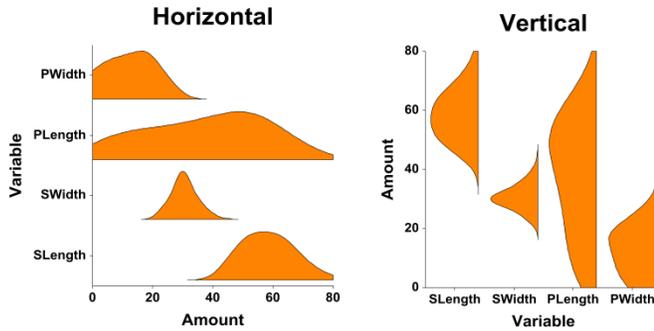
One Factor
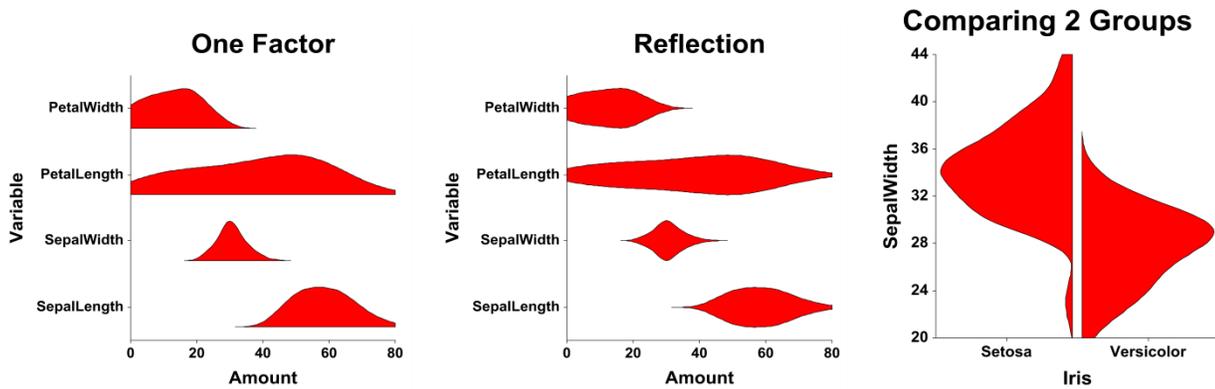
Two Factors
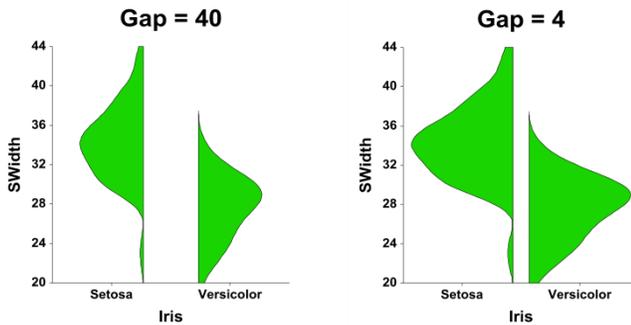
Layout Tab

Orientation Section

You can orient the density plot horizontally or vertically.

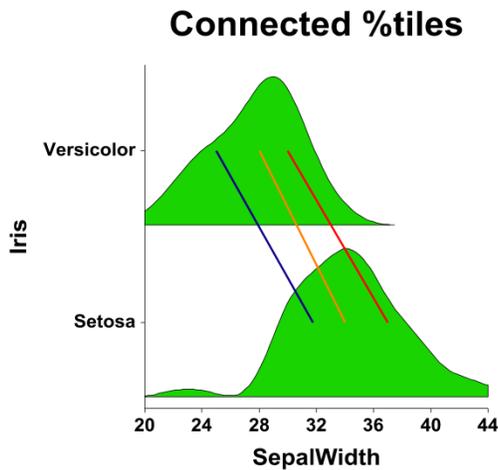## Object Spacing and Size Section

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

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 Density Plot

This section presents an example of how to generate a density plot. The data used are from the Fisher dataset. We will create density plots of the *SepalLength* variable, breaking on the type of iris.

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 Density Plots window.

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

2   Open the Density Plots window.
- Using the Graphics menu or the Procedure Navigator, find and select the **Density Plots** procedure.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

3   Specify the variables.
- On the Density Plots window, select the **Variables tab**.
- Double-click in the **Data Variable(s)** text box. This will bring up the variable selection window.
- Select **SepalLength** from the list of variables and then click **Ok**. "SepalLength" will appear in the Variable(s) box.
- Double-click in the **Horizontal (Group) Variable** text box. This will bring up the variable selection window.
- Select **Iris** from the list of variables and then click **Ok**. "Iris" will appear in the Horizontal (Group) Variable box.
- Set **Value Labels** to **Value Labels**.

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

Density Plot Output

Density Plot of SepalLength
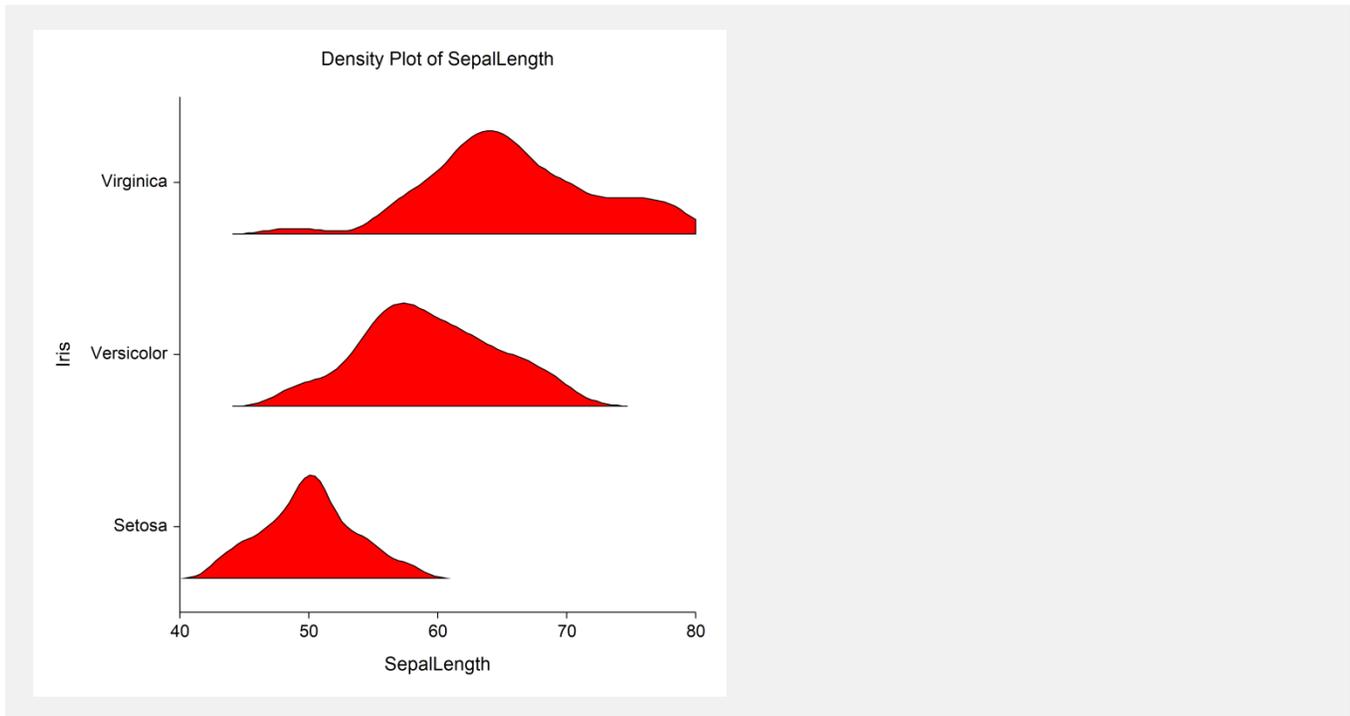
Example 2 – Creating a Density Plot with Subgroups

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

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 Density Plots (2 Factors) window.

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

2   Open the Density Plots (2 Factors) window.
- Using the Graphics menu or the Procedure Navigator, find and select the **Density Plots (2 Factors)** procedure.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

3   Specify the variables.
- On the Density Plots (2 Factors) window, select the **Variables tab**.
- Double-click in the **Data Variable(s)** text box. This will bring up the variable selection window.
- Select **Diameter** from the list of variables and then click **Ok**. "Diameter" will appear in the Variable(s) box.
- Double-click in the **Horizontal (Group) Variable** text box. This will bring up the variable selection window.
- Select **Species** from the list of variables and then click **Ok**. "Species" will appear in the Horizontal (Group) Variable box.

- Double-click in the **Legend (Subgroup) Variable** text box. This will bring up the variable selection window.
- Select **Sunlight** from the list of variables and then click **Ok**. "Sunlight" will appear in the Legend (Subgroup) Variable box.
- Set **Value Labels** to **Value Labels**.

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

Density Plot Output

Density Plot of Diameter