

## Chapter 253

# Pareto Charts

---

### Introduction

An Italian economist, Vilfredo Pareto (1848-1923), noticed a great inequality in the distribution of wealth. A few people owned most of the wealth. J. M. Juran found that this same phenomenon of the “vital few and the trivial many” applied to many areas of SPC. He is credited with coining the terms “Pareto chart” and “Pareto analysis” to represent this phenomenon.

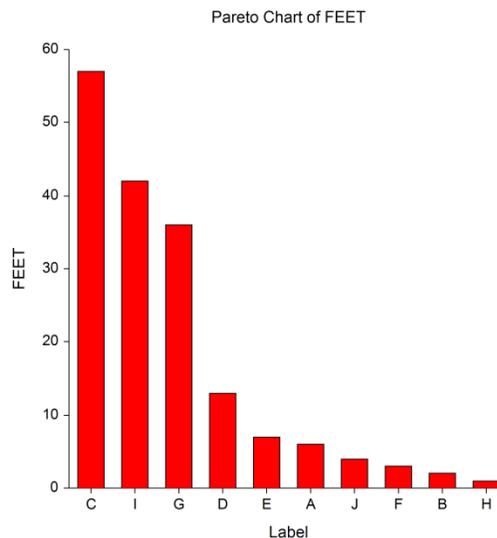
In quality control, Pareto analysis refers to the tendency for the bulk of the quality problems to be due to a few of the possible sources. Hence, by isolating and correcting the major problem areas, you obtain the greatest increase in quality. The Pareto chart is a graphic display that emphasizes the Pareto principle using a bar graph in which the bars are arranged in decreasing magnitude.

NCSS provides two Pareto chart styles as well as a numerical report.

---

### Pareto Charts

The following plot shows a Pareto chart depicting the number of defective board-feet (in 100's) from ten different mills. Notice that three mills account for almost 80% of the defective product. Obviously, efforts should be concentrated on correcting defects in these three mills.



---

### Data Structure

The table below shows the data for the above Pareto chart. It gives the number of defective board-feet (in 100's) from ten different mills (labeled A - J). These data are contained on the QATest dataset.

## Pareto Charts

## QATest dataset (subset)

Label	Feet
A	6
B	2
C	57
D	13
E	7
F	3
G	36
H	1
I	42
J	4

---

## 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.

---

### Data and Label Variables

#### Data Variables

This (required) option specifies which variables on the database contain the actual data values. If more than one variable is specified, the number of charts generated depends on the status of the Data Item option (described below).

Note that all data must be positive for the Pareto Chart. Negative values are ignored.

#### Row Label Variable

An optional variable containing labels for the individual data values may be entered here. Note that this option is only used when a single variable is analyzed. You can use dates (like Jan-23-95) as labels. Here is how. First, enter your dates using the standard date format (like 06/20/93). In the Variable Info screen, change the format of the date variable to something like *mmm-dd-yyyy* or *mm-dd-yy*. The labels will be displayed as labels. Without changing the variable format, the dates will be displayed as long integer values.

#### Chart Arrangement

This option specifies the way in which the data are to be arranged on the Pareto chart.

- **Each Row**

This option causes a separate chart to be generated for each of the Data Variables specified (see above). Labels may be set using the Label Variable option (see above). Each row of data becomes a bar on the Pareto chart. Note that the Category Variable is ignored when this option is used.

- **Total By Variable**

This option causes one chart to be constructed using all of the Data Variables specified (see above). The total for each variable becomes a bar on the Pareto chart. Note that the Category Variable and the Label Variable are ignored when this option is used.

## Pareto Charts

- **Average By Variable**  
This option causes one chart to be constructed using all of the Data Variables specified (see above). The average for each variable becomes a bar on the Pareto chart. Note that the Category Variable and the Label Variable are ignored when this option is used.
- **Total By Category**  
This option causes one chart to be constructed for each of the Data Variables specified (see above). The average of the Data Variable for each unique value of the Category Variable becomes a bar on the Pareto chart. Note that the Label Variable is ignored when this option is used.
- **Average By Category**  
This option causes one chart to be constructed for each of the Data Variables specified (see above). The average of the Data Variable for each unique value of the Category Variable becomes a bar on the Pareto chart. Note that the Label Variable is ignored when this option is used.

---

### Category Specification

#### Category Variable

An optional categorical (grouping) variable may be specified. If it is used, the Data Variable variable will be summed (or averaged) by the values of this variable. Hence, this must be a discrete variable.

#### Minimum Value

Values on the Pareto chart less than or equal to this value are lumped together into one category. This combined category is labeled using the Other Category Label.

#### Below Minimum Category Label

This option specifies the label to be displayed for the combined value when the Minimum Value is used.

---

### Specify Rows

#### Row Specification Method

This option specifies how the rows that are used in the calculations and displayed on the charts are specified.

- **All Rows**  
All rows are used.
- **First Row - Last Row**  
The first and last row is specified.
- **First N Rows**  
The first N rows on the dataset are used. The value of N is specified below.
- **Last N Rows**  
The last N rows on the dataset are used. The value of N is specified below.

---

### Specify Rows – First Row / Last Row Details

#### First Row

This option designates the first row to be used. Rows before this row are ignored. This option is only used when Row Specification Method is set to First Row - Last Row.

## Pareto Charts

### Last Row

This option designates the last row to be used. Rows after this row are ignored. This option is only used when Row Specification Method is set to First Row - Last Row.

---

### Specify Rows – N Details

#### N

This option designates the value of N. This option is only used when Row Specification Method is set to First N Rows or Last N Rows.

---

## Reports Tab

The options on this panel control the format of the reports.

---

### Select Reports

#### Numeric Report

This option specifies whether to display the numeric report.

---

### Report Options

#### Precision

This option specifies the precision of numbers in the report. A single-precision number will show seven-place accuracy, while a double-precision number will show thirteen-place accuracy. Note that the reports are formatted for single precision. If you select double precision, some numbers may run into others. Also note that all calculations are performed in double precision regardless of which option you select here. This is for reporting purposes only.

#### Variable Names

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

---

### Page Title

#### Page Title

This option specifies a title to appear at the top of each page.

---

## Plots Tab

This panel sets the options used to define the appearance of the chart.

---

### Pareto Chart

#### Pareto Chart

This option specifies whether to display the regular chart.

#### Format Button

Click the plot format button to change the plot settings (see Pareto Chart Window Options below).

## Pareto Charts

### Edit During Run

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

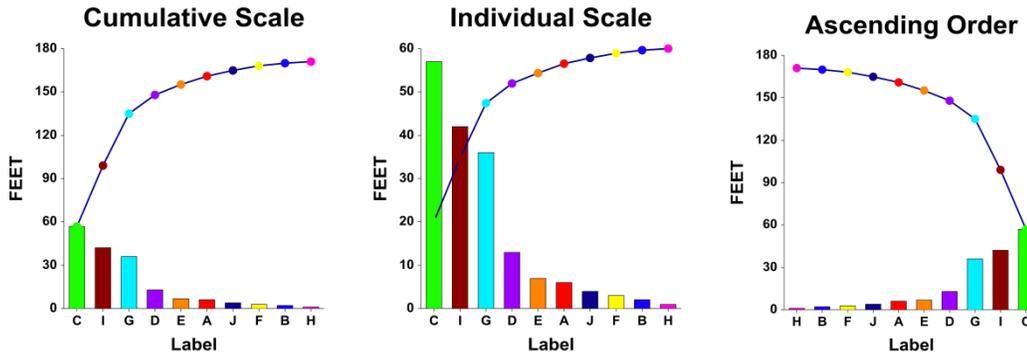
## Pareto Chart Window Options

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

### Pareto Chart Tab

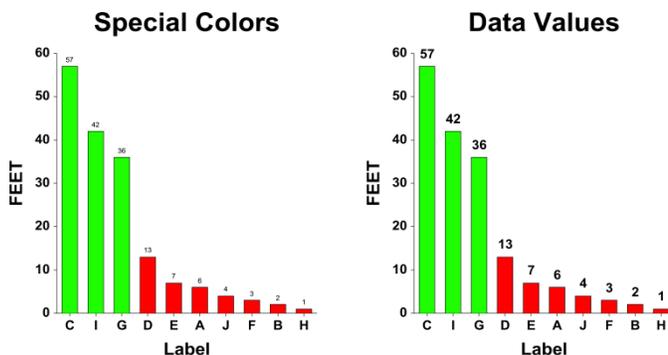
#### General Section

You can specify whether you want individual or cumulative scaling. You can also specify how you want the bars to be sorted.



#### Bars Section

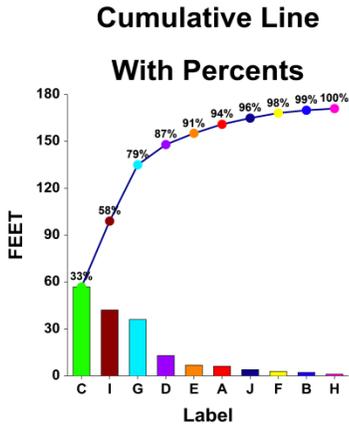
You can specify the format of the bars using the options in this section. You can add a connecting line and/or the data values to the plot.



## Pareto Charts

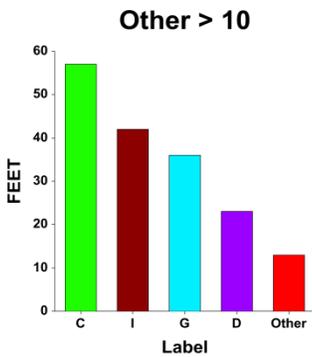
### Cumulative Line Section

You can add specify the format of the cumulative line using the options in this section. You can add cumulative percentages above the symbols.



### Other Category Section

You can combine several low values into and Other category.

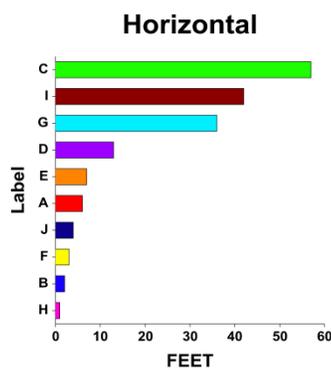
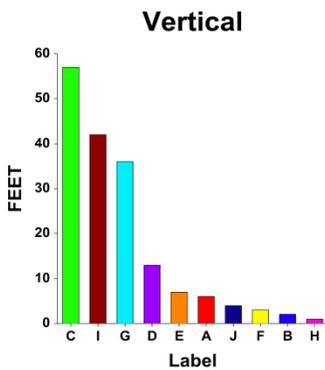



---

## Layout Tab

### Orientation Section

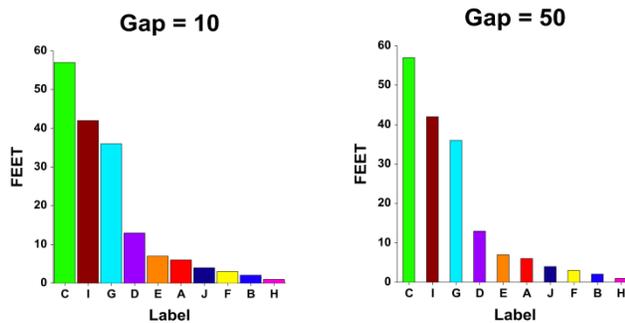
You can orient the chart vertically or horizontally.



## Pareto Charts

## Object Spacing and Size Section

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



## 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 Pareto Chart

This section presents an example of how to generate a Pareto chart. The data used are shown in the table at the beginning of the chapter and are found in the QATest 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 Pareto Charts window.

#### 1 Open the QATest dataset.

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

#### 2 Open the Pareto Charts window.

- On the menus, select **Graphics**, then **Quality Control**, then **Pareto Charts**. The Pareto Charts procedure will be displayed.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

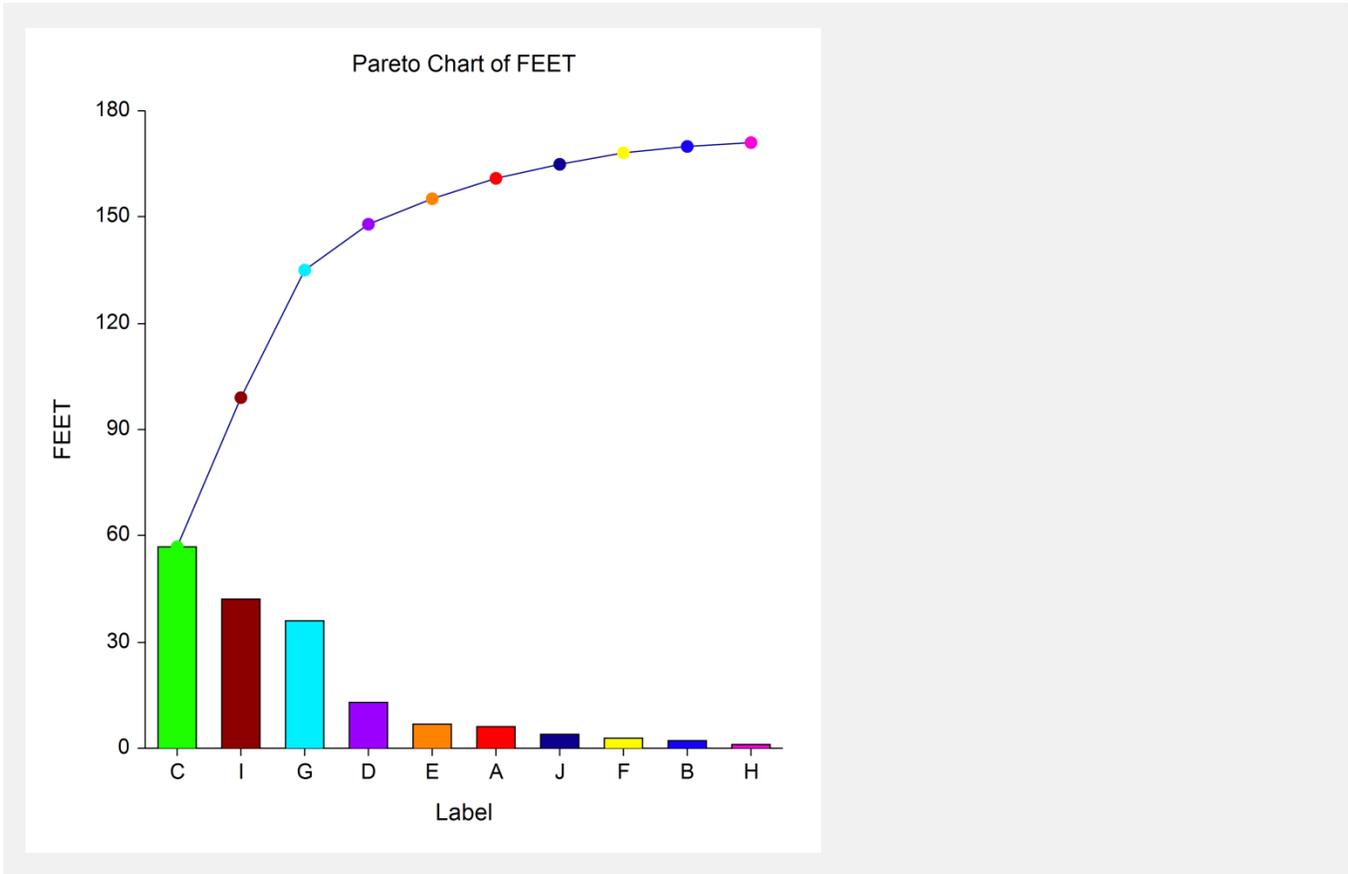
#### 3 Specify the variables.

- On the Pareto Charts window, select the **Variables tab**.
- Double-click in the Data Variables text box. This will bring up the variable selection window.
- Select **FEET** from the list of variables and then click **Ok**. “FEET” will appear in the Data Variables box.
- Double-click in the Row Label Variable text box. This will bring up the variable selection window.
- Select **Label** from the list of variables and then click **Ok**. “Label” will appear in the Label Variable box.

#### 4 Run the procedure.

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

## Pareto and Cumulative Pareto Chart



This section displays the cumulative Pareto chart. Note that this version of the Pareto chart combines the bar chart version with a line representing the cumulative total for each bar. The cumulative percentage is displayed above the plotting symbol. For example, we see that 79% of the defects are caused by mills C, I, and G.

**Pareto Numeric Report**

Label	FEET	Cumulative FEET	Percent	Cumulative Percent
A	6	6	3.51	3.51
B	2	8	1.17	4.68
C	57	65	33.33	38.01
D	13	78	7.60	45.61
E	7	85	4.09	49.71
F	3	88	1.75	51.46
G	36	124	21.05	72.51
H	1	125	.58	73.10
I	42	167	24.56	97.66
J	4	171	2.34	100.00

This report gives the numerical details of the analysis.

---

## Example 2 – Using Several Variables

This section presents an example of how to generate a Pareto chart of the total for several variables. The data used are the values of S1 - S5 found in the QATest dataset. Suppose, for the moment, that these five variables represent the fifty daily numbers of defects produced by each of five shifts.

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 Pareto Charts window.

### 1 Open the QATest dataset.

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

### 2 Open the Pareto Charts window.

- On the menus, select **Graphics**, then **Quality Control**, then **Pareto Charts**. The Pareto Charts 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.

- On the Pareto Charts window, select the **Variables tab**.
- Double-click in the **Data Variables** text box. This will bring up the variable selection window.
- Select **S1** through **S5** from the list of variables and then click **Ok**. “S1-S5” will appear in the Data Variables box.
- Select **Total By Variable** from the **Chart Arrangement** list box.

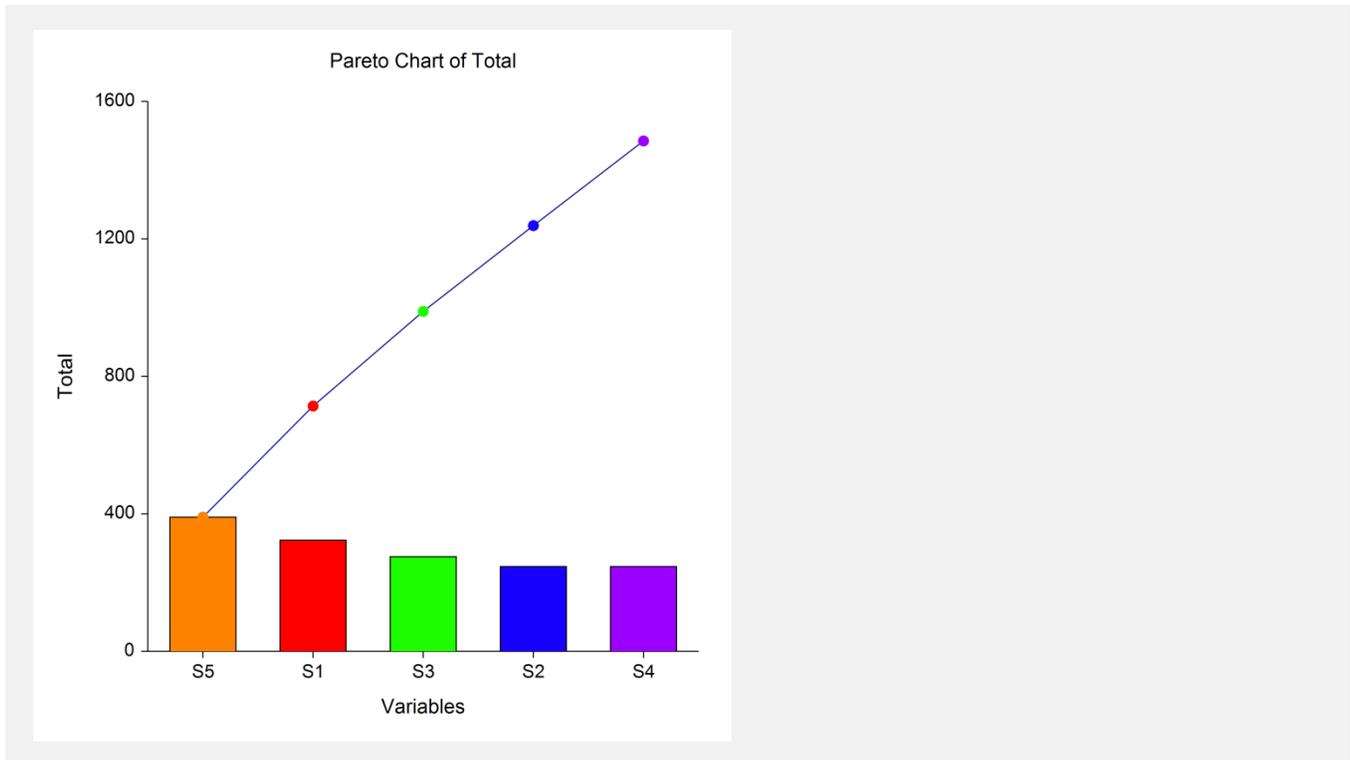
### 4 Specify the reports.

- On the Pareto Charts window, select the **Reports tab**.
- Click **Regular Chart** so that it is **not checked**.
- Click **Numeric Report** so that it is **not checked**.

### 5 Run the procedure.

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

## Pareto Chart



Notice that a somewhat uniform defect rate in the five shifts is reflected here by almost identical bars and by an almost straight cumulative line.