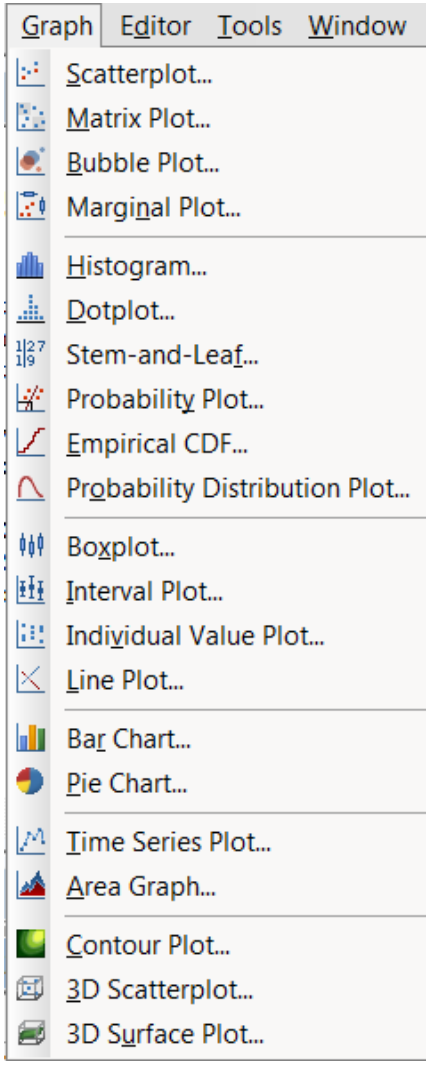
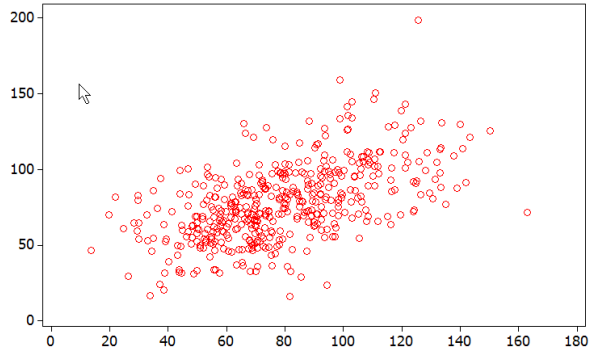
# Graphing

# Graphing

- ***Minitab*** *can create many different types of graphs.*
  - *Most accessed by clicking* ***Graph*** *on the menu bar.*



# Scatter Plots




- The most common plot made during any analysis is the simple X-Y scatter plot.
- The X-Y scatter plot examines the relationship of two quantitative variables.

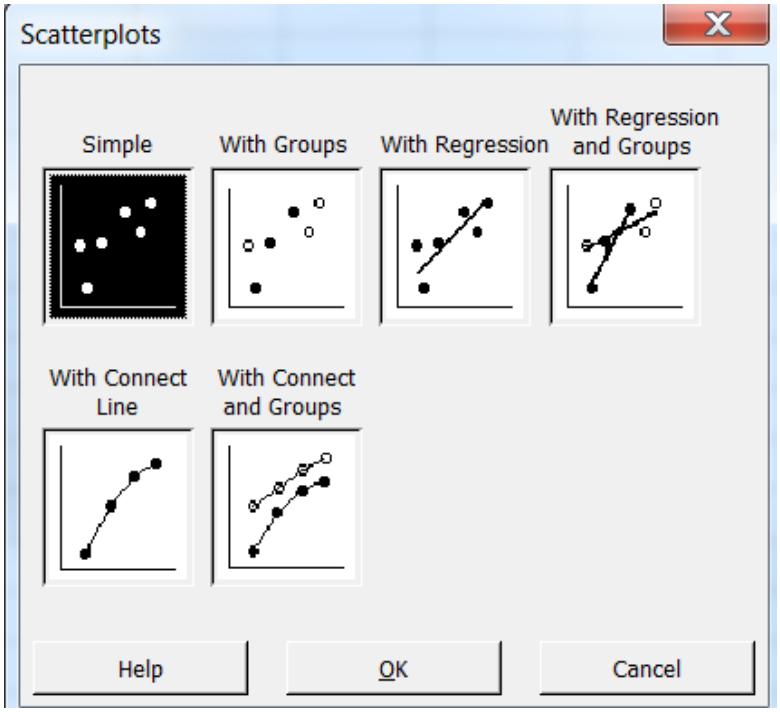
# Scatter Plots

- Close **Minitab** and restart it to clear the old project.
  - Recommended by Minitab to start a new project.
- Import the Excel file ***Minitab Dataset 2***.
  - Hint: use **File** → **Open Worksheet**.
- This is a dataset of basis weight and moisture for a 4167 grade.

# Scatter Plots

- To create a scatterplot click on **Graph** → **Scatterplot…** on the Minitab menu bar.

The following window will appear.



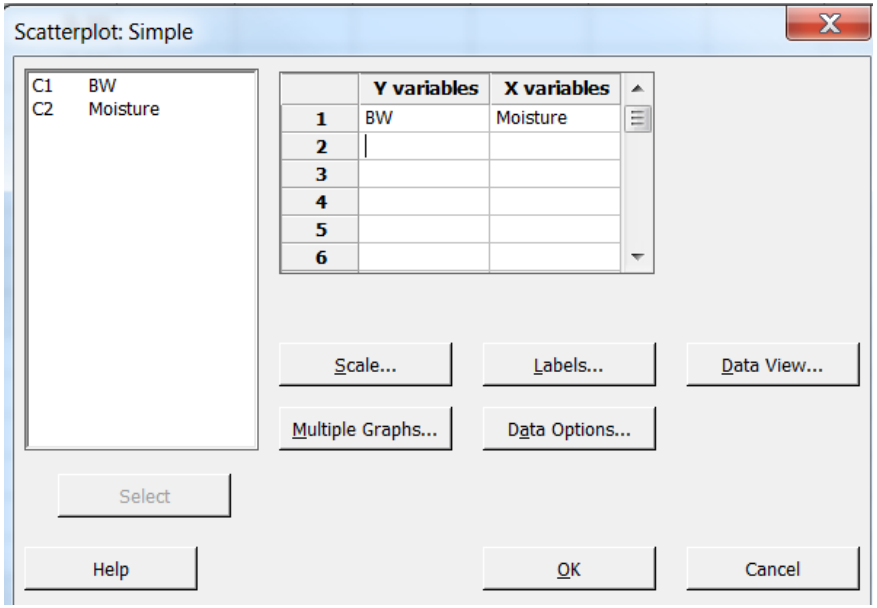
# Scatter Plots

- Choose the **Simple** by clicking in the box and making sure that box is highlighted (all black) and click **OK**.
- Select ***BW*** as the **Y Variable**.
- Select ***Moisture*** as the **X Variable**.

# Scatter Plots

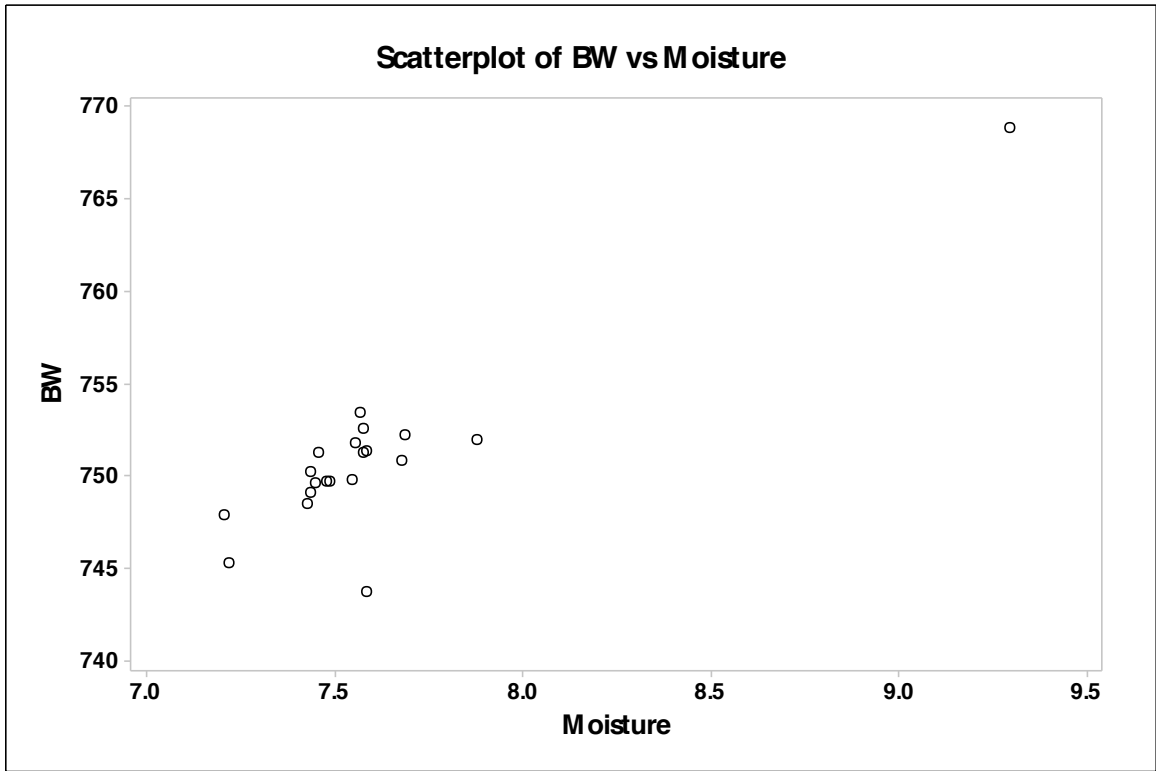
- Your Scatterplot window should look like the example below.

Click **OK** to create the scatterplot



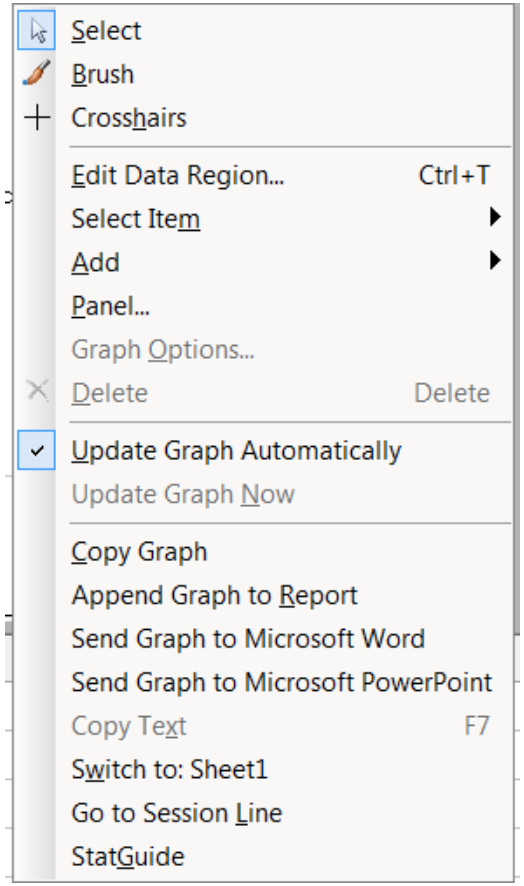
# Scatter Plots

The scatterplot created should look like this.



# Scatter Plots

- *Right mouse click* anywhere on the graph to bring up the graph editing menu.



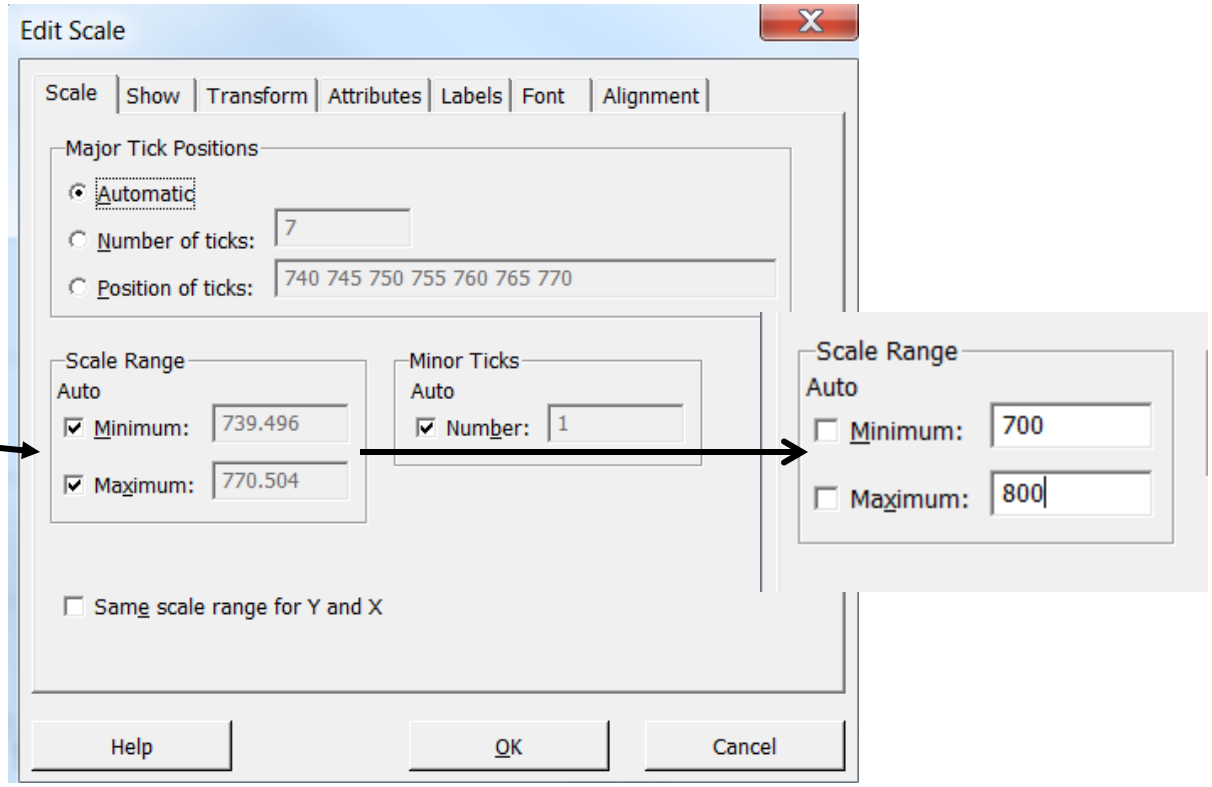
# Scatter Plots

- The **Add** menu entry is the most commonly used.
- **Brush**, **Edit Figure Region…** and **Copy Graph** are also useful features that are used less frequently.
- Change the graph features by *right-clicking* or *double-clicking* on them.
- Let's try changing the y-axis scale.

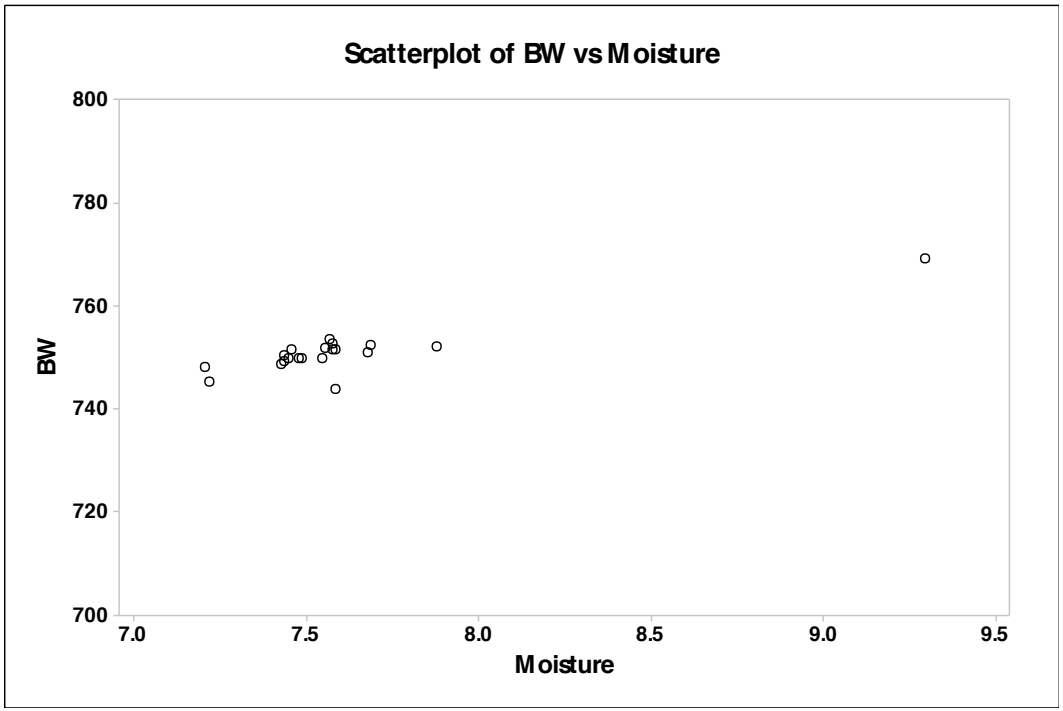
# Scatter Plots

- Double-click on any axis to change its scale. Eg, double-click the y axis.

**Uncheck** to change the **Min** or **Max** value of the selected axis. Type in new values you want.



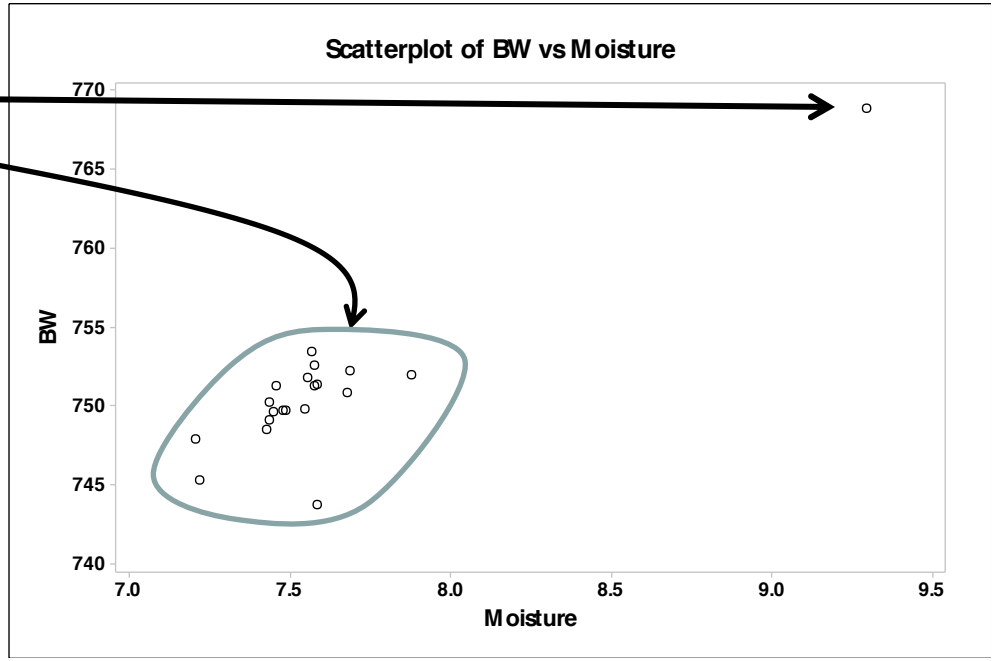
# Scatter Plots




Now try putting the default y-axis scale back. Hint: put the checks back.

# Scatter Plots

- Next, let's **double-click** on one <u>data symbol</u> to edit their **attributes**, **grouping** and add **jitter**.

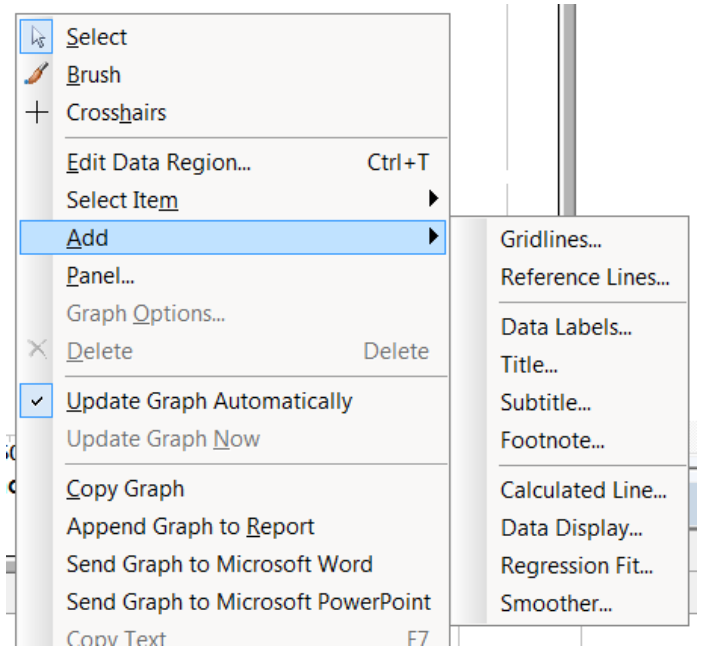


# Scatter Plots

- **Attributes –** changes the type, color and size of the data symbol.
- **Groups –** adds a grouping variable which changes the symbol to identify the group.
- **Jitter –** spreads out data points in the **x** and/or **y** direction for better visualization.

# Scatter Plots

- You can **Add** many features to a graph.
- Right-click on the graph select the **Add** menu entry.



# Scatter Plots

- The **Add** menu:
  - **Gridlines…** adds a grid to the plot.
  - **Reference Lines…** adds lines at user specified locations in the x or y direction.
  - **Data Labels…** adds numeric or text labels to y values on the plot.
  - **Title…** , **Subtitle…** , **Footnote…** are self-explanatory.

  - Practice on your plot.

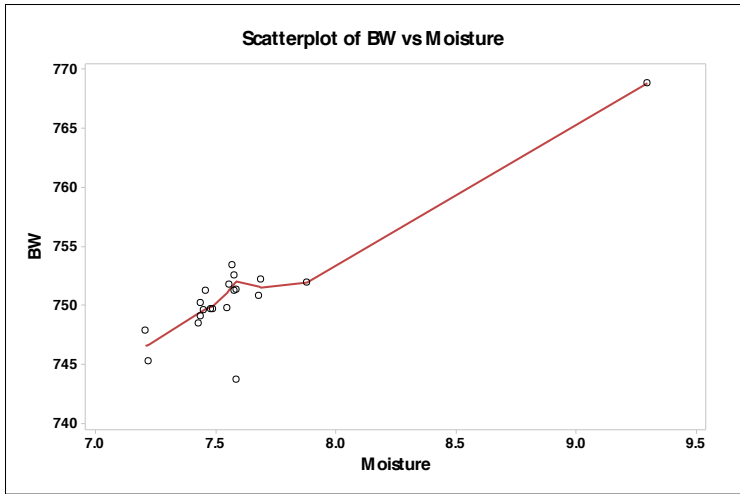
# Scatter Plots

- More **Add** menu:
  - **Calculated Line…** adds a line of paired data that is in your data set.
  - **Data Display…** adds symbols, connect lines, projected lines and shaded areas.
  - **Regression Fit…** adds a linear, quadratic or cubic regression line to the plot.
  - **Smoother…** adds a Lowess smoothed line thru the data in the plot.

  - Practice on your plot.

# Scatter Plots

- Let's add a **Lowess** smoothed line.
- **Right click** on the graph → select **Add** → select **Smoother…** → click **OK**

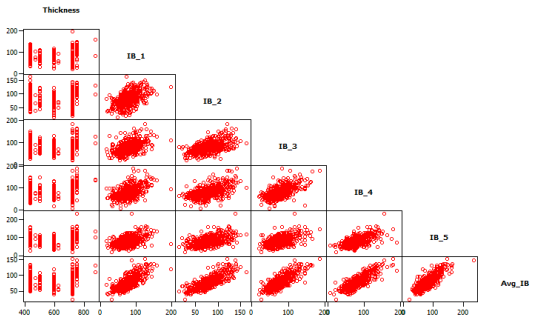


# EXERCISE

- Practice "**Add**"-ing things on your plot.

- Next, create a scatter plot with the data in the file ***Minitab Dataset 10.xls***

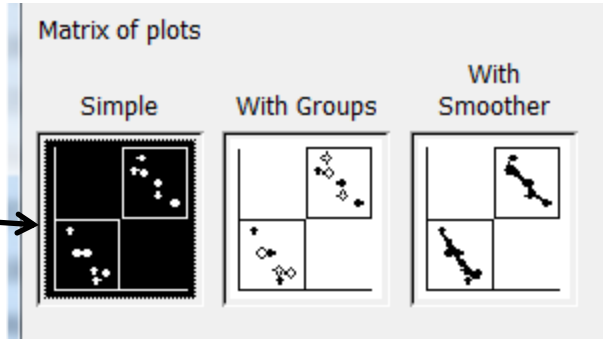
# Matrix Plot




- A good method of displaying many quantitative variables is the *matrix scatterplot*.
  - Quickly examine each variables relation to the others.
  - Helps to spot unusual data points.

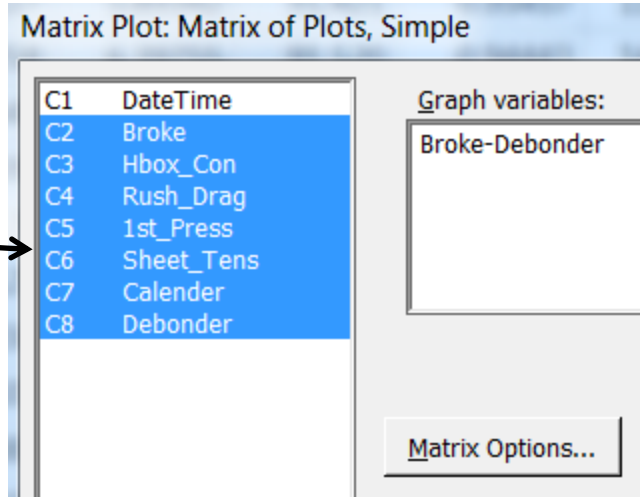
# Matrix Plot

- Close and restart Minitab.
- Open worksheet ***Minitab Dataset 3.xls***.
- Select **Graph** → **Matrix Plot…** from the menu.
- The most commonly used option is **Matrix of Plots Simple.**



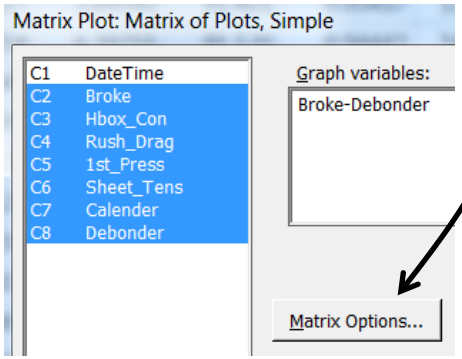
# Matrix Plot

- Select **Matrix of Plots Simple** and click **OK**.
- Select ***Broke*** thru ***Debonder***



# Matrix Plot

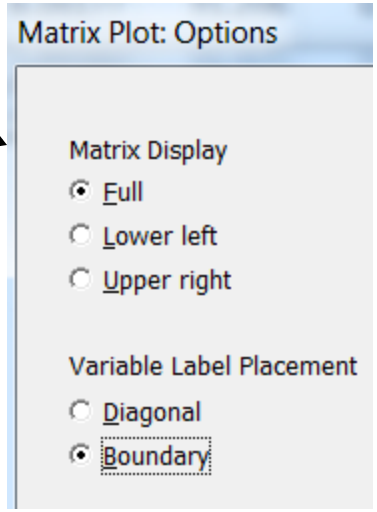
- Click the **Matrix Options…** button.




- Most people are used to seeing the variable names on the axis and you only need one-half of the Matrix Plot.

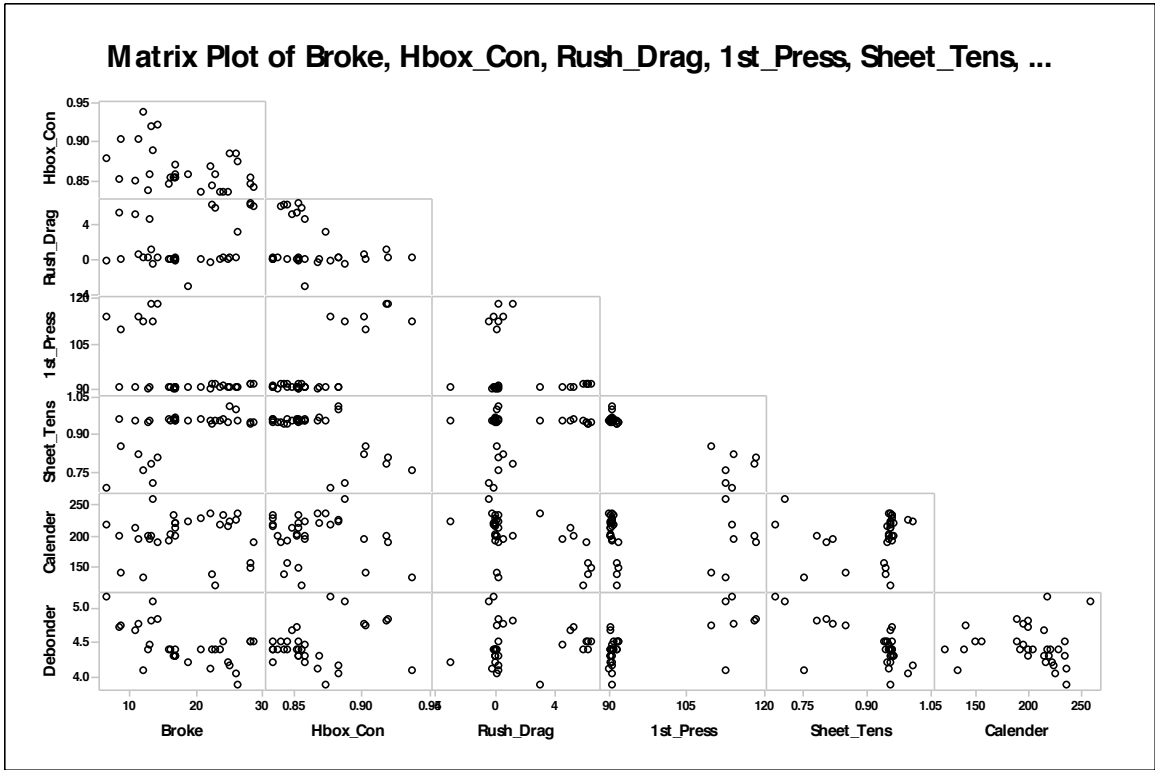
# Matrix Plot

- Click the **Lower Left** radio button and the **Boundary** for the Variable Label Placement.




- Now click **OK** twice.

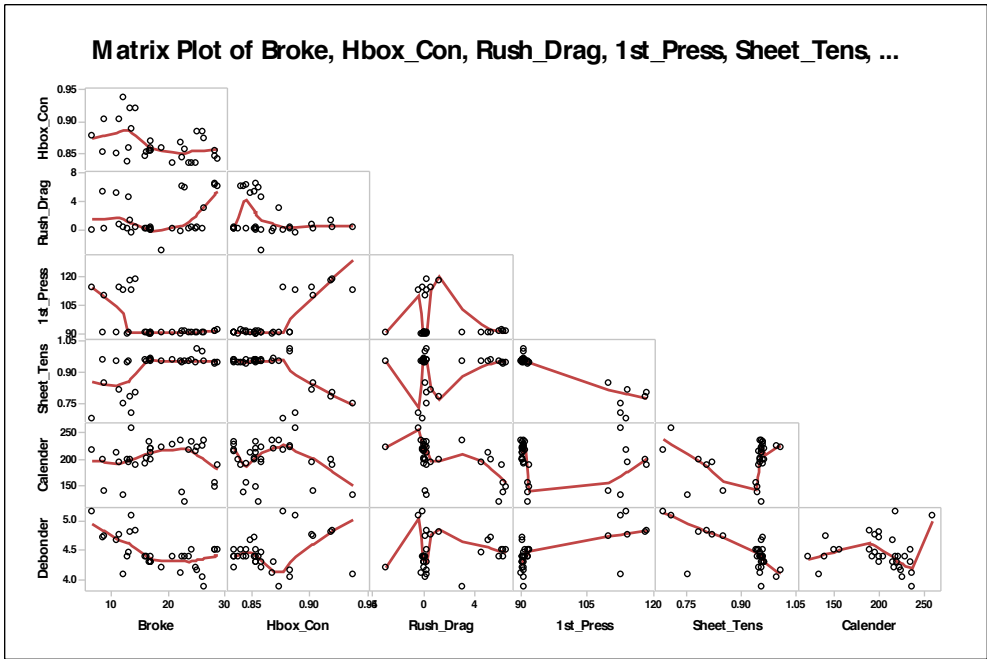
# Matrix Plot




Sometimes adding a **Lowess** smoothed line is useful. Let's do it.

# Matrix Plot

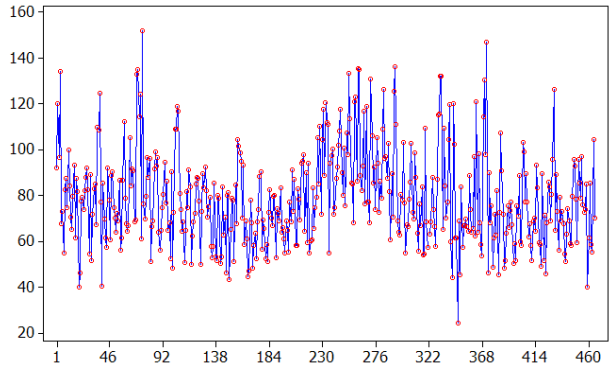
- **Right click** on plot → select **Add** → **Smoother…** → click **OK**



# EXERCISE

- Open the file ***Exercise 9.xlsx*** and create a Matrix plot of all the continuous data.

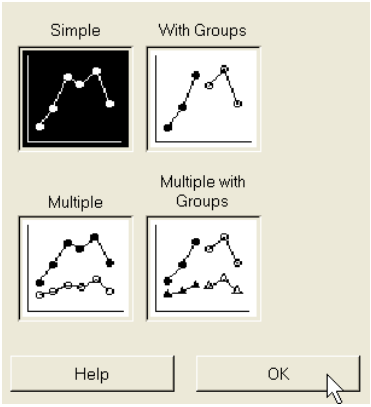
# Time Series Plot




- Many times it is useful to examine a dataset over time.
- Minitab provides the *time series plot* to easily build this type of graph.

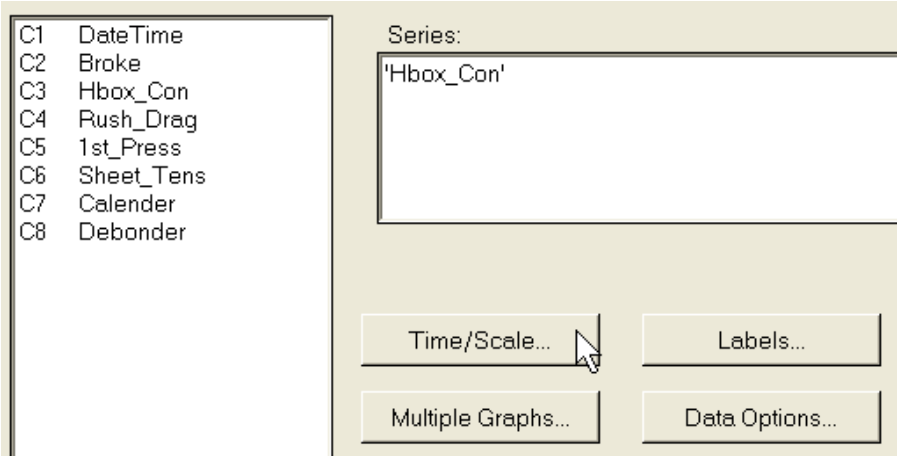
# Time Series Plot

- Let's use ***Minitab Dataset 3.xls*** dataset.
- To create a *time series plot*, select **Graph** → **Time Series Plot…**
- The **Simple** and **Multiple** time series plots are the most frequently used.
- Select the **Simple** box → click **OK**.



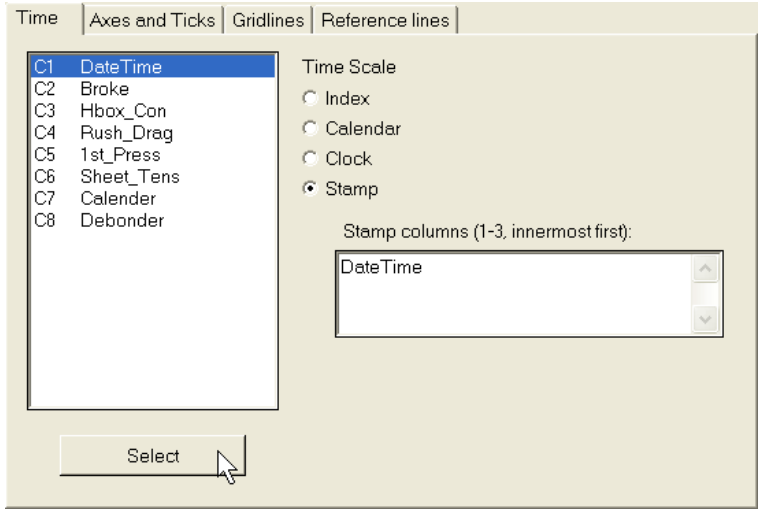
# Time Series Plot

- Select ***Hbox_Con*** as the **Series** variable.
- Click the **Time/Scale…** button.

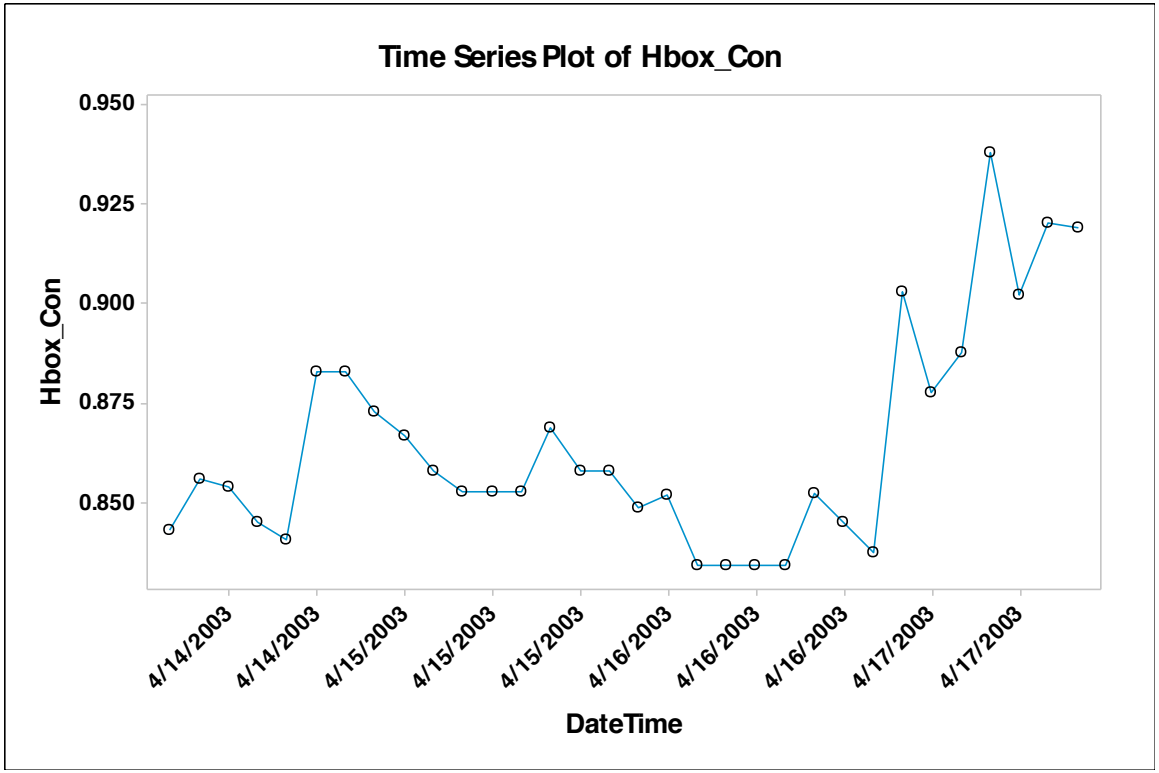


# Time Series Plot

- Click the **Stamp** radio button.
- Click in the **Stamp columns** box and select ***DateTime***
- Click **OK** twice.



# Time Series Plot

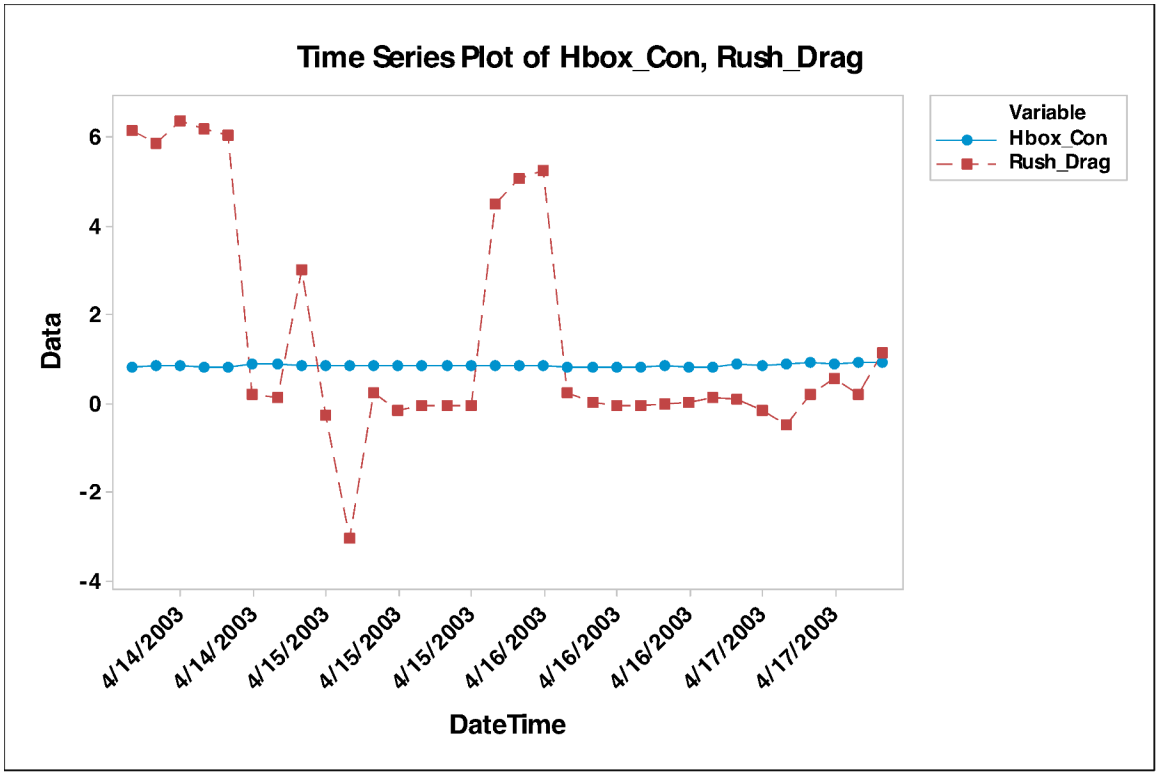



If you have a lot of data, the symbols can hinder the visualization. Delete the symbols.

# Time Series Plot

- Create a Multiple Time Series of ***Hbox_Con*** and ***Rush_Drag***.
- Select **Graph** → **Time Series** → **Plot…** → **Multiple** → **OK**.
- Make sure to put ***DateTime*** on the X-axis.

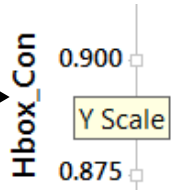
# Time Series Plot



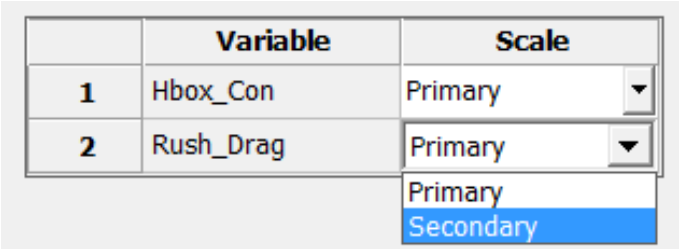

This plot is useless since the scales are so different. We can create a ***secondary axis*** to make it more useful.

# Time Series Plot

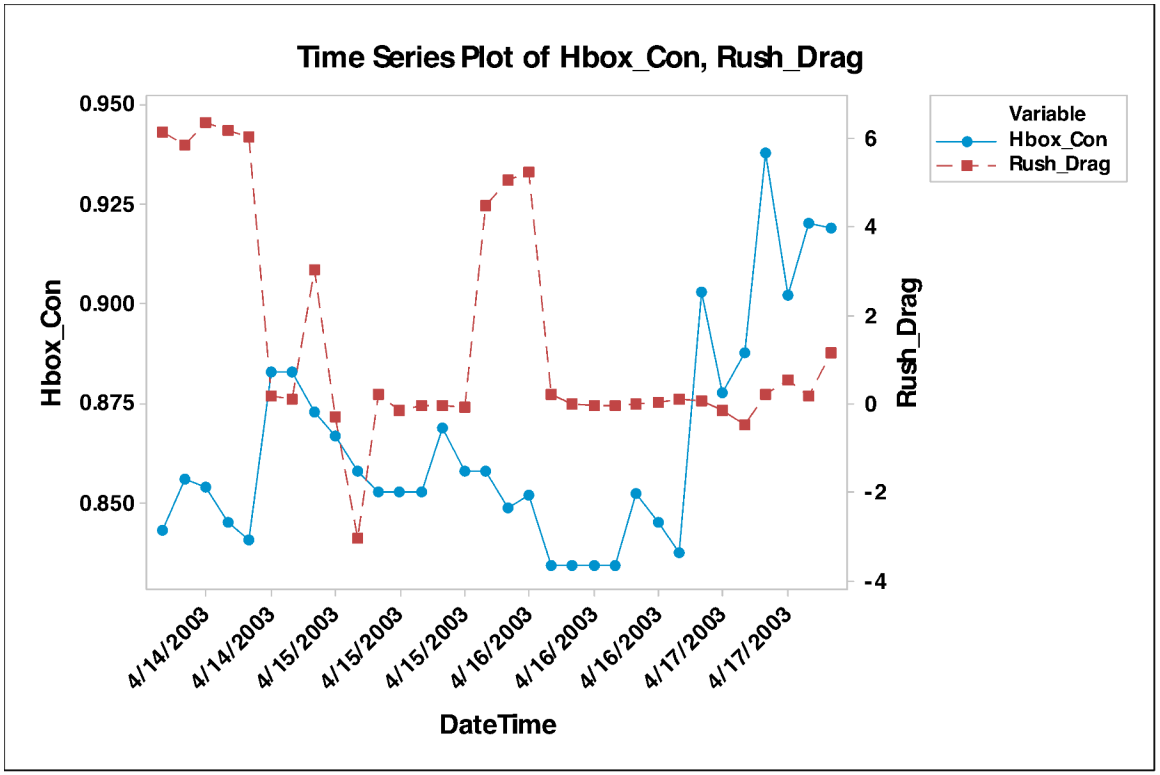
- Double-click on any number on the Y-axis to **Edit Scale**




- Select the **Secondary** tab → click the down arrow in the **Scale** entry for ***Rush_Drag*** and select **Secondary** → click **OK**



# Time Series Plot

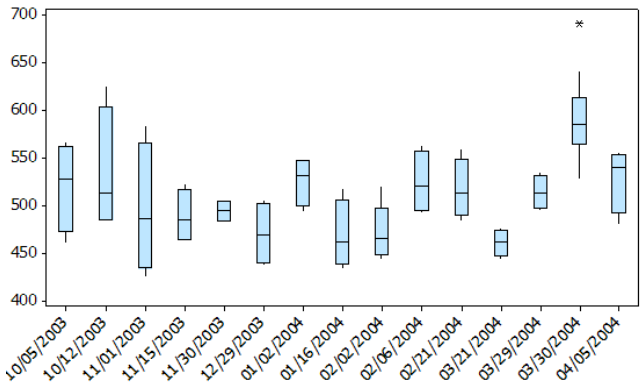



This plot is more useful.

# EXERCISE

- Create Time Series plots of the rest of the continuous dataset that we are using.

# Box Plot




- The **box plot** (or box and whisker plot) is used to examine the distribution of data for a variable.

# Box Plot

The box plot construction:

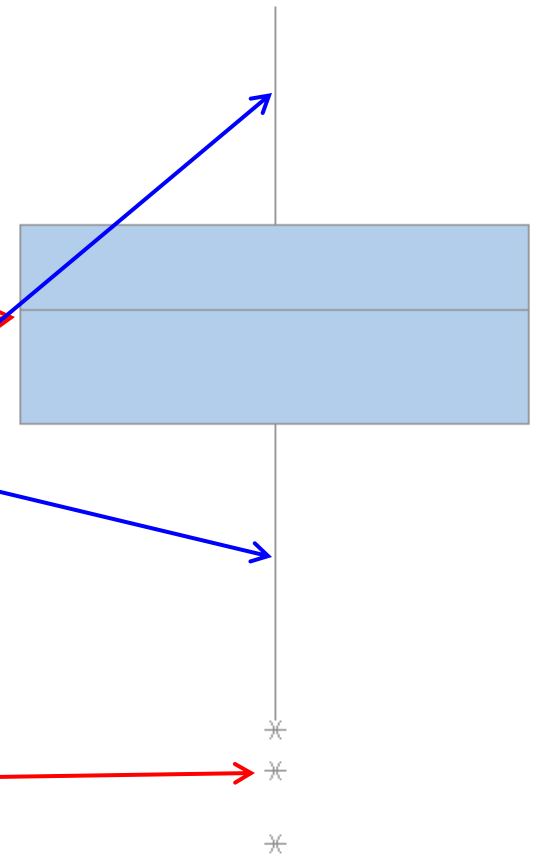
- The box is the ***interquartile range*** (IQR )

  1st quartile is lower box line and 3rd quartile is upper line

- The line in the middle of the box is the ***median*** (Q2 or 50th percentile).
- The whiskers represent ±1.5 x IQR for the upper and lower.

  The IQR is a robust estimate of dispersion and is ~1.33 x standard deviation.

- The asterisk (*) represents an ***outlier*** value.

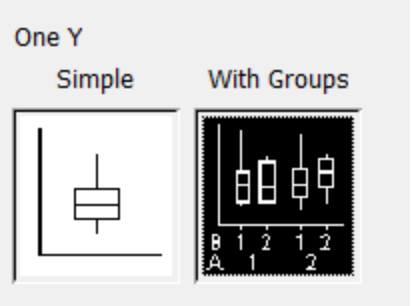


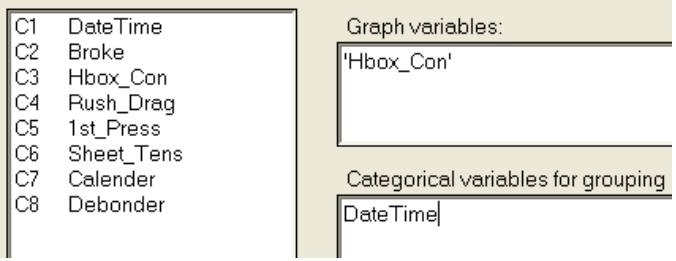
# Box Plot

- Box plots using the **With Groups** option is used to examine the location and dispersion of a variable with several categories.
- We will continue with ***Minitab Dataset 3.xls***.
- Select **Graphs** → **Boxplot…** from the menu.

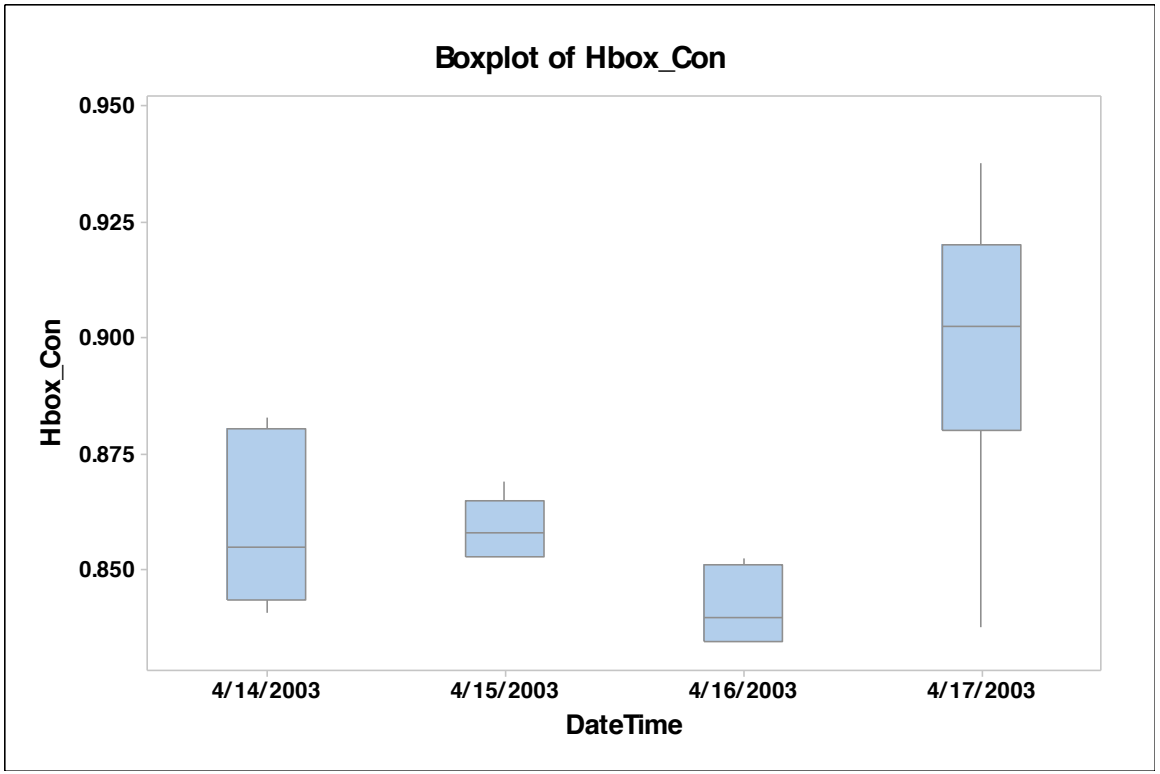
# Box Plot

- Select the One Y **With Groups** box → click **OK**.
- Select ***Hbox_Cons*** as the **Graph Variable** and ***DateTime*** as the **Categorical Variable** → click **OK**.



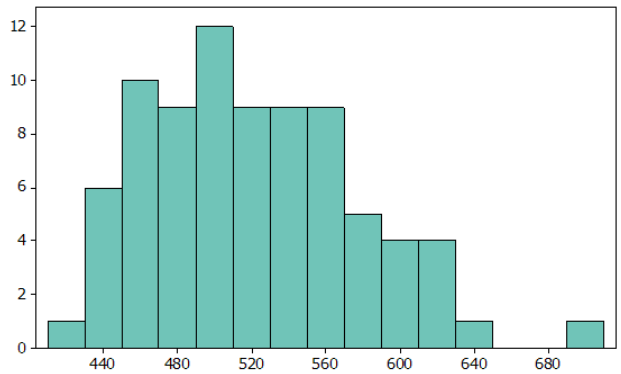
# Box Plot


Boxplot of Hbox_Con
0.950
0.925
0.900
0.875
0.850
Hbox_Con
4/14/2003
4/15/2003
4/16/2003
4/17/2003
DateTime

# EXERCISE

- Open the file ***Minitab Dataset 6.xls***
- Create a **Boxplot With Groups** using ***Moisture*** as the **Graph** variable and ***Roll Position*** as the **Categorical** variable.

# Histogram

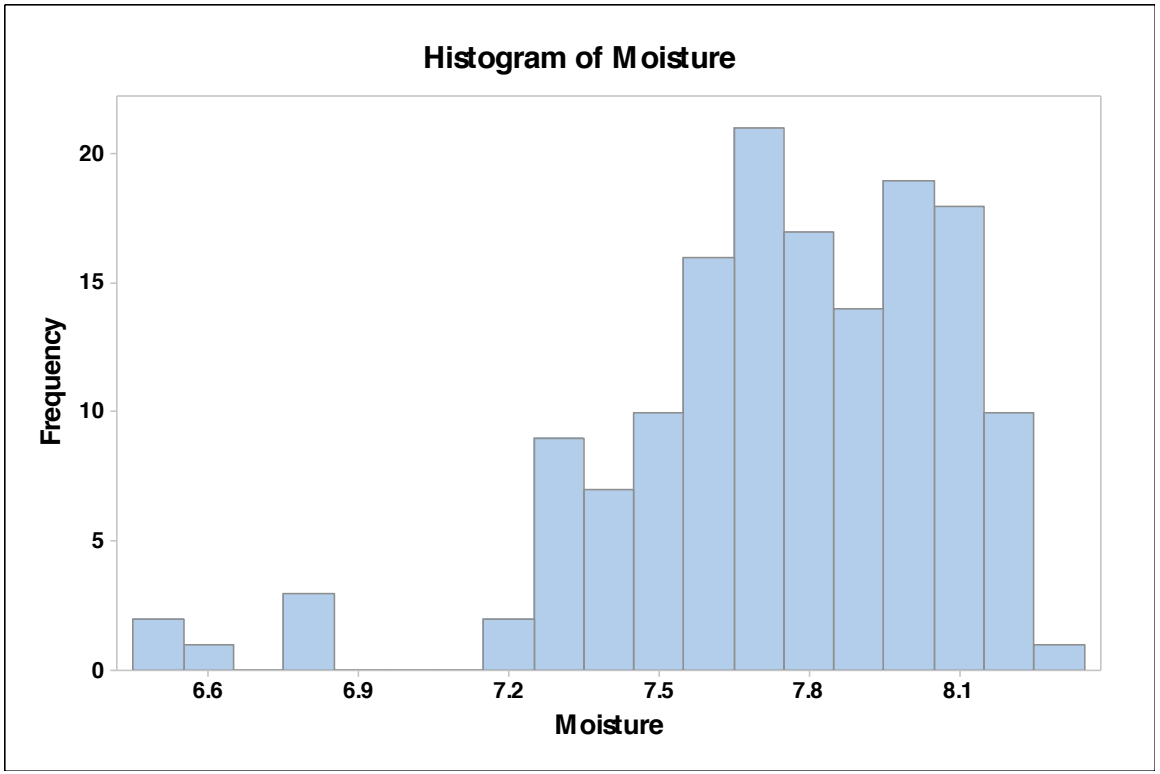



- Another way to graphically summarize a variable is the *histogram*.
- A histogram is used to examine the shape and spread of a variable.

# Histogram

- Close Minitab and re-open Minitab.
- Open Worksheet ***Minitab Dataset 6.xls***.
- To create a *histogram*, select **Graph** → **Histogram…** → **Simple** → click **OK**.
- Select ***Moisture*** for the **Graph Variable** → click **OK**.

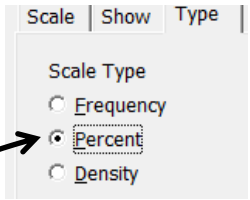
# Histogram

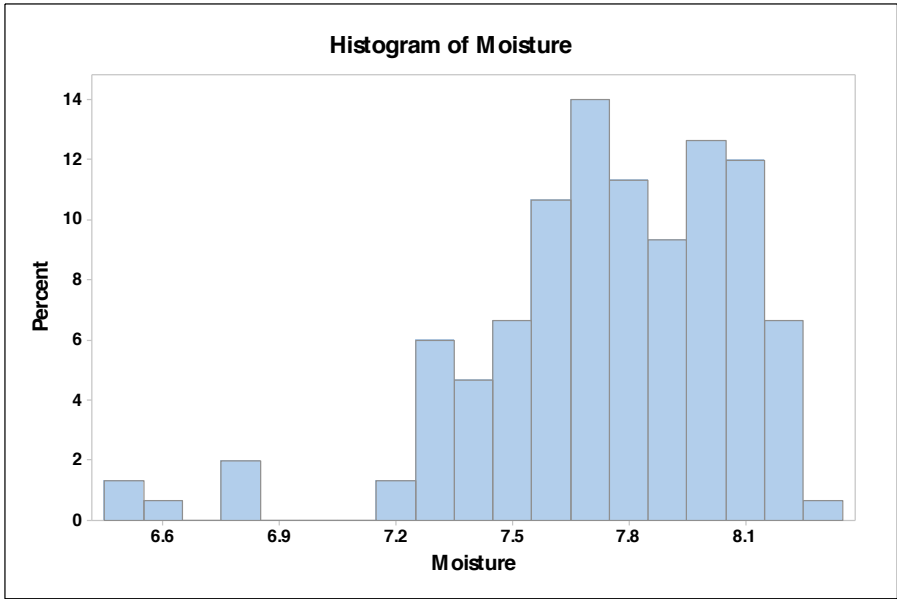



The default is to display the Frequency of each bar. You can change the graph to display the percent for each bar.

# Histogram

- Double-click on the Y-axis (as before) to **Edit Scale** → select the **Type** tab → select **<u>Percent</u>** as the **Scale Type** → click **OK**.



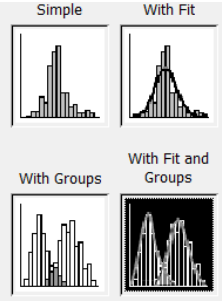



# Histogram

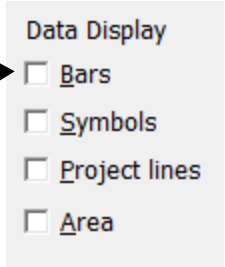
- One way to use the histogram is with the overlaid normal curve fit and categorical grouping option.
  - The mean and standard deviation of each group is used to build the normal curve.
- Open worksheet file ***Minitab Dataset 7.xls***

# Histogram

- Select **Graph** → **Histogram…** → **With Fit and Groups** → **OK**.

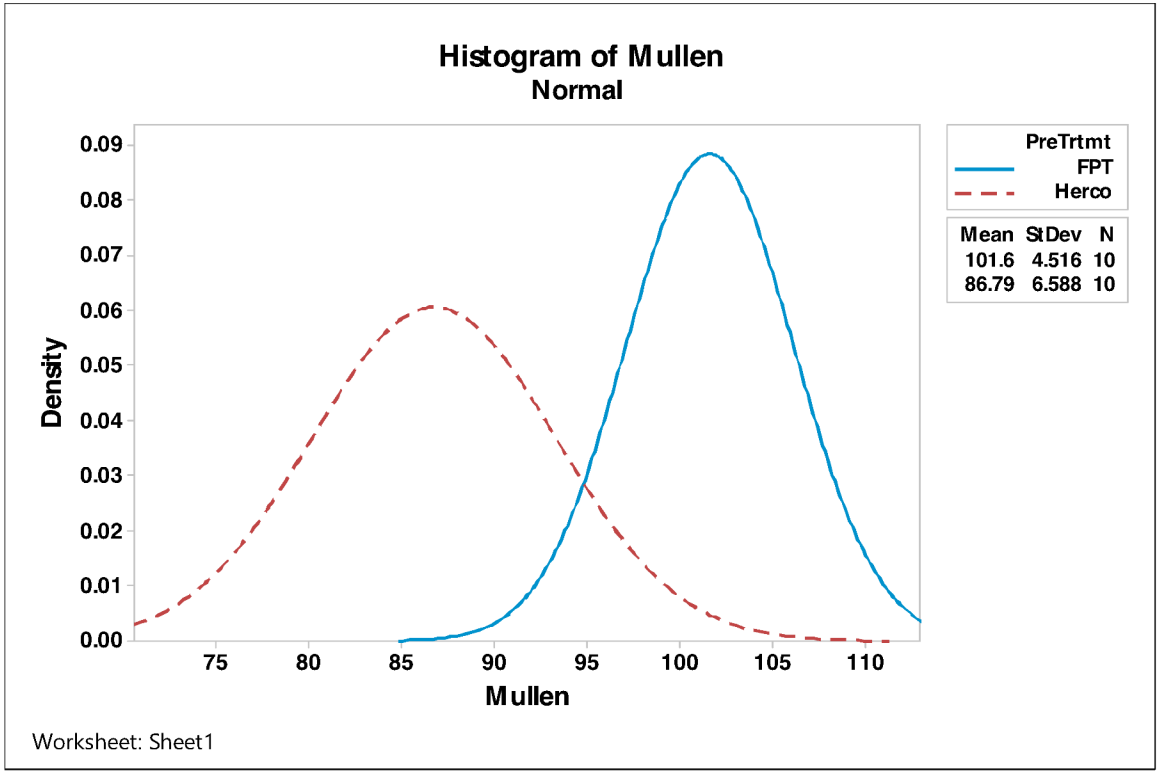



- Select ***Mullen*** for the **Graph** variable, ***PreTrtmt*** for the **Categorical** variable → click the **Data View** button → in the **Data Display** tab, uncheck **Bars**




- Click **OK** twice.

# Histogram



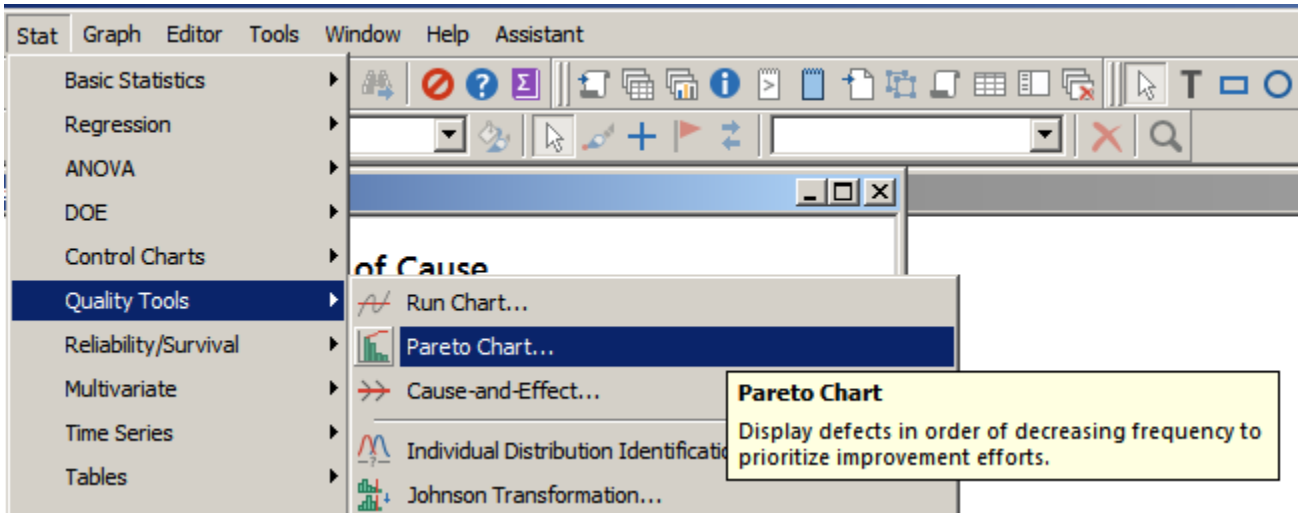

A histogram and summary table of the mean and std dev of **Mullen** for each **PreTrtmt** group is displayed.

# Pareto Chart

- Pareto principle (or **80/20 rule**):
  - **80%** of the problems (defects, upsets, etc.) are the result of **20%** of the causes.
  - Visualized by the Pareto Chart.
    - Y-axis (left side of chart) is the count of problems, and 2nd Y-axis is the percentage.
    - X-axis lists the problem categories.

# Pareto Chart

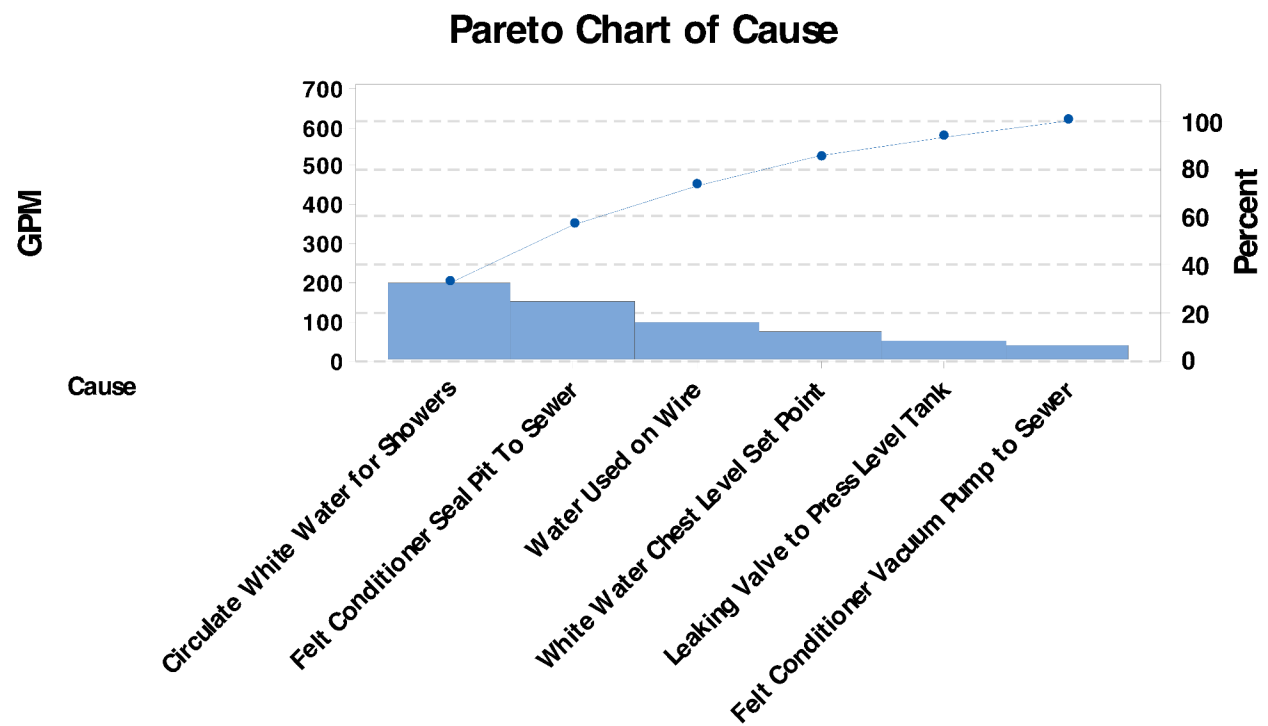
- Use the file "*ParetoExample.xls*".
- To create a Pareto Chart select **Stat** → **Quality Tools** → **Pareto Chart**.



# Pareto Chart

- In the “Defects or attribute data in:” box select **Cause**.
- In the “Frequencies in:” box select **GPM**.
- Click **OK**.

# Pareto Chart




| | | | | | | |
|---|---|---|---|---|---|---|
| GPM | 200 | 150 | 100 | 75 | 50 | 40 |
| Percent | 32.5 | 24.4 | 16.3 | 12.2 | 8.1 | 6.5 |
| Cum % | 32.5 | 56.9 | 73.2 | 85.4 | 93.5 | 100.0 |

# EXERCISE

- A customer has sent info to the pulp mill sales that certain rolls are what they consider “good” or “bad” for their process.
  - A process engineer collected the DCS data associated with the production of these jumbos.
- Use the dataset *Exercise 2.xls*
- Make a histogram or two and use all the other plots you’ve learned about to explore this data.
- Please explain your findings.