



SCIENCE FAIR SERIES: DATA HANDLING

Playing with Excel & deciding which graph to use to display your data.



Prepare Excel for Playing

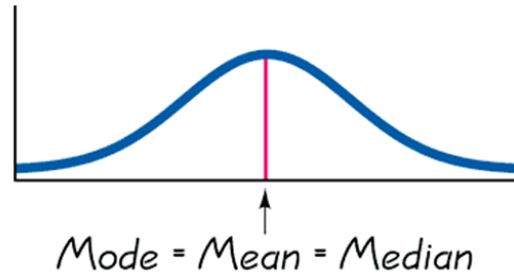
- Instructions:
 - Click “File”
 - Click “Options”
 - Click “Add-ins”
 - Click “Analysis ToolPak”
 - At Manage Excel Add-ins”, Click “Go”
 - In pop-up, Click “OK”

<http://sciencefair.math.iit.edu/analysis/>

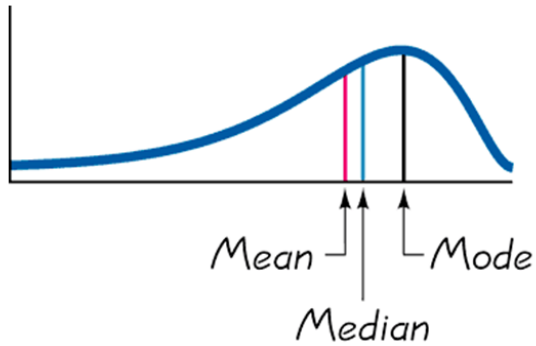
Let's Go to Excel

- And play...

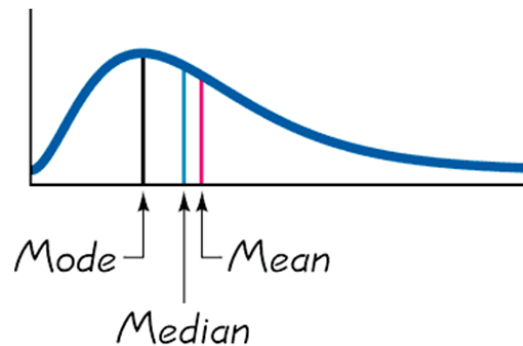
Histogram and Measure of Skew



(b) Symmetric



(a) Skewed to the Left
(Negatively)



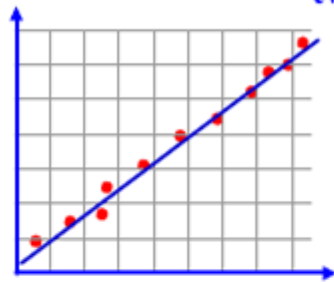
(c) Skewed to the Right
(Positively)

Histogram is really a bar graph for quantitative data (which is data that can be measured like gallons of water or number of darts)

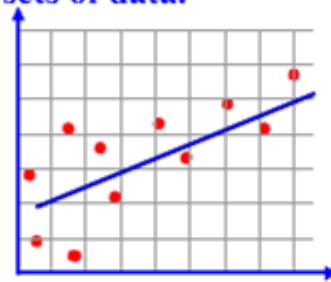
Scatter Plots & Correlation

SCATTERPLOTS & CORRELATION

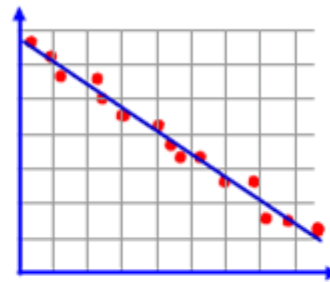
Correlation - indicates a relationship (connection) between two sets of data.



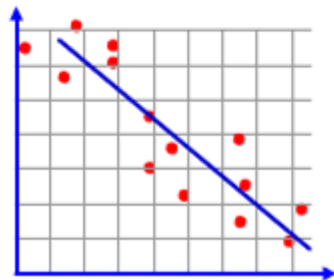
Strong positive correlation



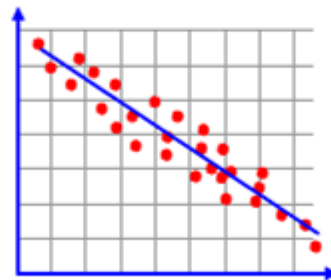
Weak positive correlation



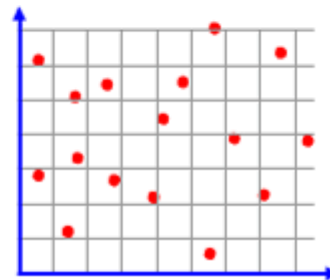
Strong negative correlation



Weak negative correlation



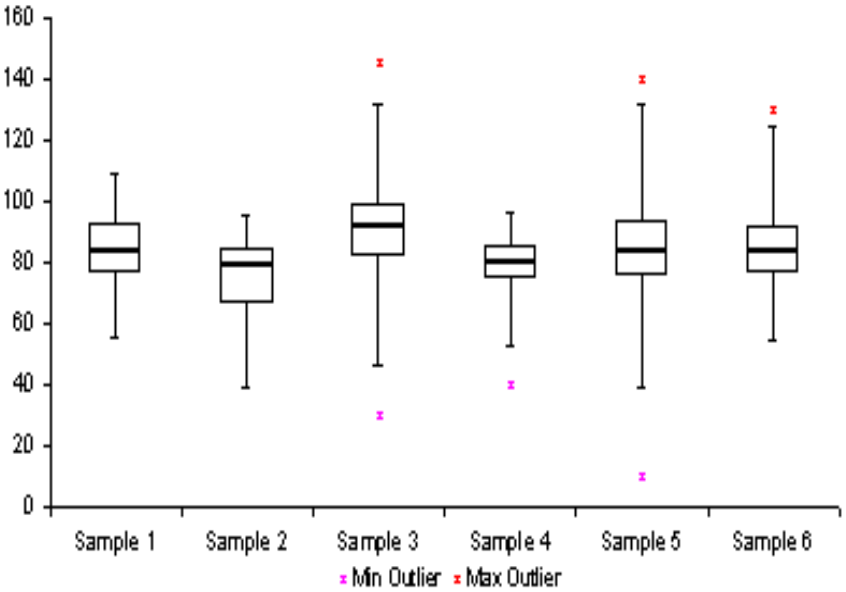
Moderate negative correlation



No correlation

Correlation is a statistical technique used to determine whether there is a statistically significant relationship between 2 variables.

Box and Whisker Plot



- — **OUTLIER** More than 3/2 times of upper quartile
-
- **MAXIMUM** Greatest value, excluding outliers
- **UPPER QUARTILE** 25% of data greater than this value
- **MEDIAN** 50% of data is greater than this value; middle of dataset
- **LOWER QUARTILE** 25% of data less than this value
- **MINIMUM** Least value, excluding outliers
- — **OUTLIER** Less than 3/2 times of lower quartile

Box and Wisker Plot displays minimum, maximum, outliers, ranges, and interquartile ranges, medians and lower and upper quartiles all in one graph.

Line Graphs

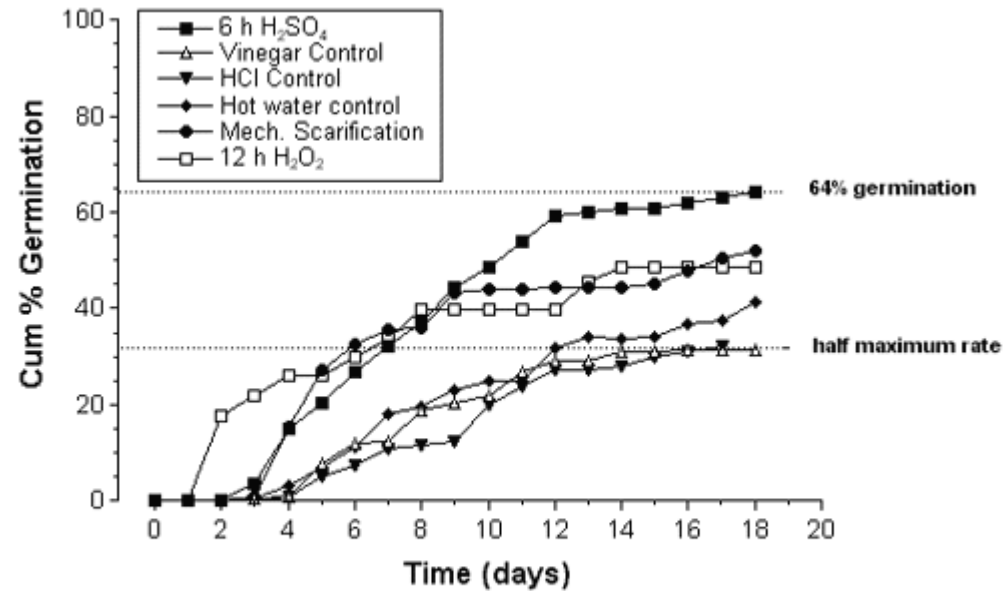
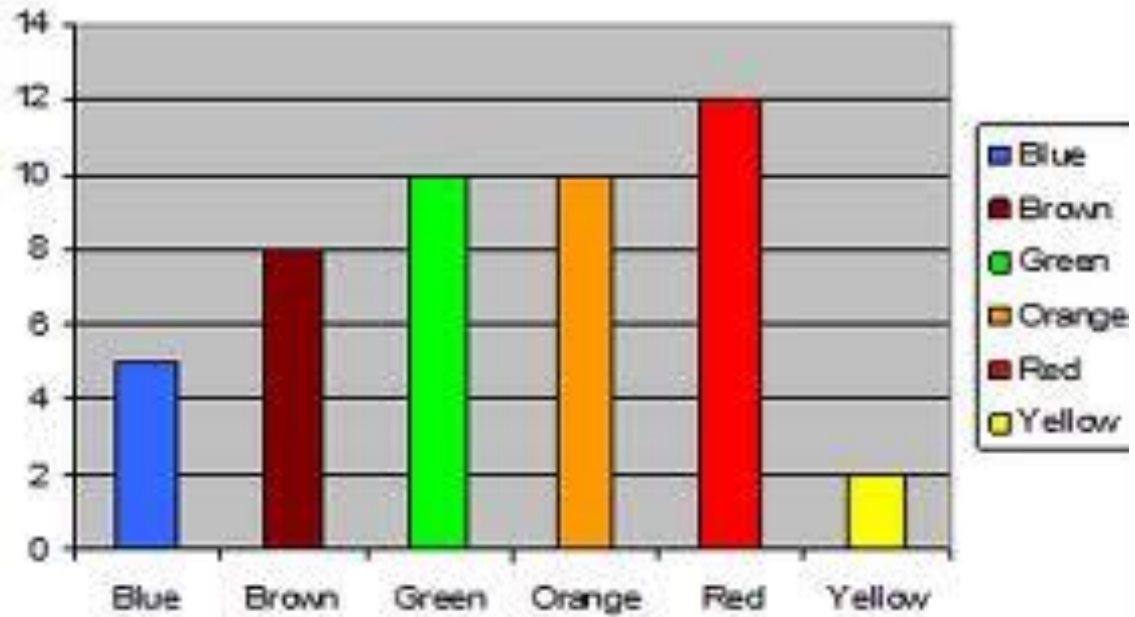


Figure 2. Cumulative germination of gourd seeds following various pregermination treatments. n = 100 seeds per trial.

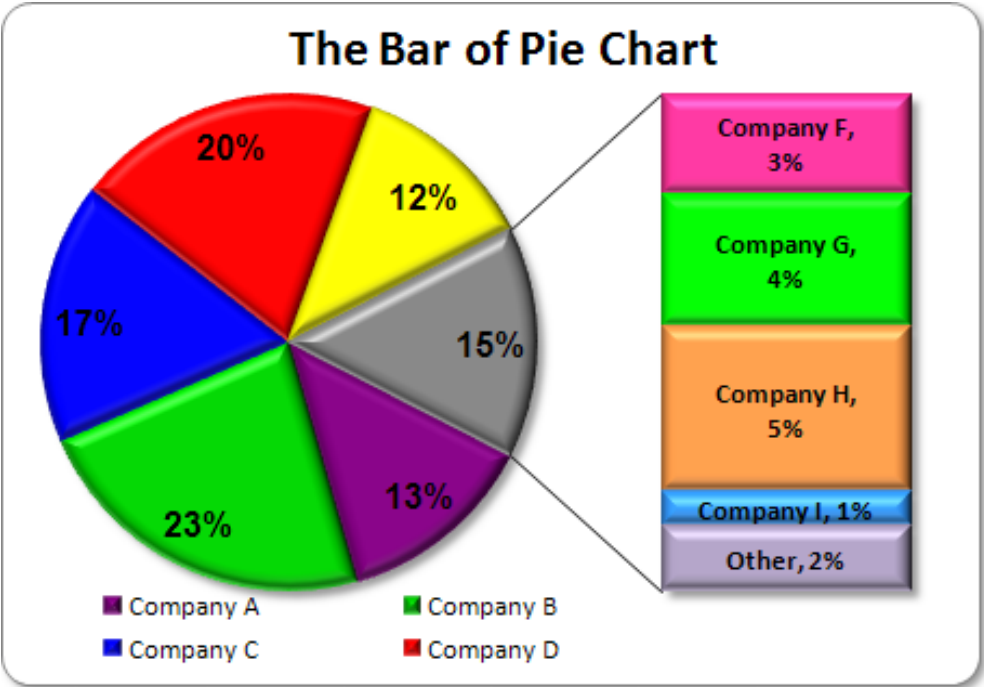
Line graphs display data points connected by line segments often representing change in time.

Bar Graphs



Best to show distribution of data using bars that indicate frequency of a specific category.

Pie Chart & Bar of Pie Chart



Pie charts represent data by dividing a “whole” into sectors by sized that are proportions to the quantity of the data.

What Graph to Use

What to Show	Description	Type of Graph
Distribution of Data	Distribution charts are used to show how variables are distributed over time, helping identify outliers and trends.	Line Histogram or Column Histogram or Scatter Plot
Composition of Data	Composition charts are used to display parts of a whole and change over time	Stacked Column Charts or Pie Charts
Comparison of Data	Comparison charts are used to compare one or more datasets. They can compare items or show differences over time.	Column/Bar Charts or Line Charts
Relationship	Relationship charts are used to show a connection or correlation between two or more variables	Scatter Chart or Bubble Charts

<https://infogram.com/page/choose-the-right-chart-data-visualization>

Join us for
conclusion of this
series

We will talk about how to use
the data to prove your
hypothesis or verify your
engineering design.

Visit **NEOHSTEM Alliance Website**

- For more project information
- <http://neohstem.org/>

Contact **Science & Technology at the Akron-Summit County Public Library**

- 330-643-9075
- stdiv@akronlibrary.org