A Picture is worth a thousand data points

Data generated during scientific experiments is commonly and effectively represented in the form of a graph or figure. In this exercise students explore how to convert data to graphs and figures.

Grades: 6-8

Subjects: Science

Objectives:
1.To understand how to categorize data by data type and create an organized data table.
2.To understand how to represent data pictorially.
3.To understand the appropriateness of graph/figure type for data type.

Oregon Science Standards:
7.3S.2 Organize, display, and analyze relevant data, construct an evidence-based explanation of the results of an investigation, and communicate the conclusions including possible sources of error.

Materials:
Part A
-Attached graph and figure examples
Part B
-Attached data table, alternatively: classroom generated data table
-Graphing materials, including but not limited to: graph paper, colored pencils, and rulers.
Part C
-Attached data table, alternatively: classroom generated data table
-Attached graph and figure examples
-Graphing materials, including but not limited to: graph paper, colored pencils, and rulers.

Time Considerations
Preparation and explanation: 20 minutes
Activity: 30 minutes

Preparation
Part A
Set-up: Assemble accompanying example graphs and figures. These are most effectively shared via transparencies or a PowerPoint.
Part B
Set-up: Pass out graph paper, rulers, and colored pencils.
Part C
Set-up: Conduct experiments to gather data, or pass out attached data sets. Assign groups of 3 and pass out graph paper, rulers, and colored pencils.

Related Activities
Chemical conundrum: A separations activity
Mass Spec activity

Background

Research generates data that can be used to understand a system and to predict or interpret future events. The product of scientific research is generated to explain scientific observations and to share those observations with others in a meaningful way. This most commonly is in the form of a published research journal article, such as those found in Nature, Science, or Cell. It is necessary to portray data in a manner that can be easily understood and interpreted by others as data is the currency of science. (graph author resource). In this exercise, students will: be introduced to the concept and examples of graphs and figures (Part A), produce a graph or figure as a class (Part B), analyze data and generate a graph or figure in small groups (Part C). Part B and C can utilize data generated from experiments outlined in this packet or can be based upon the included supplemental data sets.

Part A – Understanding Graphs and Figures

Data is often represented pictorially in the form of a figure or graph. Graphing is a representation containing an X and Y axis. In this activity we will consider scatter plots, line graphs, and bar graphs.
-Scatter plots are best used to describe trend data, such weight and size. Different objects of the same size can have different weights. Inversely, different sized objects can have the same weight.
-Line graphs are for data sets in which one axis can constantly be related to the other. For example, distance traveled and time make axes to define the data points of speed. This relationship can be used to find unknown information.
-Bar graphs are used to break data into categories with one variable. For example, quantity of each category.

Figures are data represented by means such as tables or pictures. In this activity we will consider pie graphs and Venn diagrams as figures.
-Pie graphs are circular images that represent data categories as a percentage of the whole data set. For example, students with blonde hair out of all students in a class.
-Venn diagrams are representations of data that lie in more than one category.

Guiding Questions: Do data come in different types? If so, should the different types be represented in different ways?

Part B – Assembling a Graph

In this activity students are exposed to hands on graph making. The attached data table can be used for this exercise. Alternatively, a data table generated by the Chemical Conundrum separations experiment or the Mass Spectrometry experiment may be graphed. Using a blackboard, overhead projector, or easel, the concept of data table as a means to organize scientific data is explored. From there, the best fit figure is selected from those presented in the previous section (Line graph, bar graph, scatter plot, Venn diagram, or pie chart). The teacher then leads by example as the whole class draws the graph or figure.

Guiding Questions: How are data generated during an experiment organized and converted into a graph or figure?

Part C – Independent Graph Making

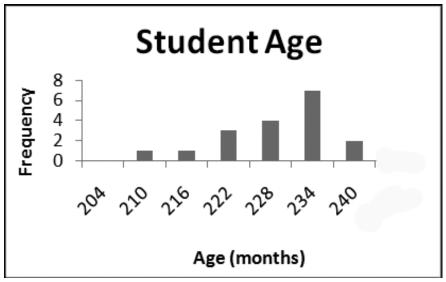
In this exercise students are allowed to utilize critical thinking skills by accessing the appropriateness of graph or figure type choice for the data to be represented. The students are divided into groups of three and provided with graphing materials including, but not limited to, graph paper, colored pencils and rulers. A data table will be provided to the students from the attached materials. Alternatively, data tables can be generated from the Chemical Conundrum separations experiment or the Mass Spec simulation experiment.

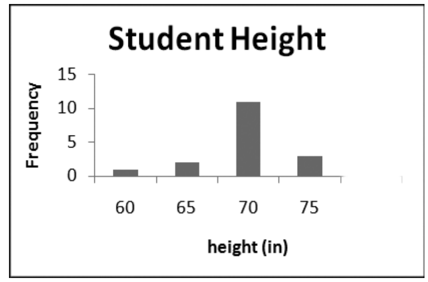
If students are having difficulty with the concept of graph choice, the instructor may restrict the graph possibilities or guide the students through the graph style choice process. The groups will act independently to construct graphs or figures. The instructor should be available for questions or to offer assistance if students are struggling with the concepts. If the students are grasping the concepts and selected graph style on their own, a follow up comparison and contrast of graph choice by group and appropriateness to the data can be a student lead discussion and exploration.

Guiding Questions: What makes a good graph? What does a bad graph look like? What factors need to be considered when making a graph style choice?

HISTOGRAMS

Histograms summarize discreet data or continuous data that are measured on an interval scale.





Histograms are a popular tool for illustrating the major features of a data set. Instead of visually presenting every data point, data is sorted into categories and the population of each category is plotted.

Constructing a Histogram

1. A histogram is constructed by grouping data into classes or groups of convenient size, and representing the population of each group.
2. For each group, construct a rectangle with a height proportional to the population of that category.
3. If your data is categorical, or if the size of each division is constant, than all rectangle on your histogram should be of equal width. Although uncommon, histograms with rectangles of varying widths are sometimes seen, where the width of the rectangle indicates the width of the category.

Normal Distribution

Skewed Distribution

Bimodal Distribution

Random Distribution

<u>Data Table:</u>

ft	in	yr	mo
5	5	18	2
5	8	20	0
5	6	19	1
5	6	18	4
5	6	19	4
5	5	19	2
6	2	18	6
6	0	18	11
5	8	17	4
5	8	18	3
5	8	19	5
5	0	19	1
5	7	19	1
5	8	19	7
6	1	17	8
5	9	18	9
5	4	18	10
5	7	19	4
5	7	18	9

<u>Gather your own data:</u>

The histogram is also well suited to displaying the results of the Separations activity included in this packed. Students can construct histograms for their data set as well as one for the class as a whole and examine for differences in population distribution. A histogram is preferred over a pie chart if there are more than five categories of data, of if the categories are of similar size.

The height and age data used here was gathered by a quick survey of Oregon State University freshman chemistry students. A similar survey of your students for any categorical data (music taste, favorite tv shows, etc) could be used.

This data set and Power Point slides of the figures for classroom presentation can be downloaded at:

PIE CHARTS

Pie charts are used to display categorical data or percentage distributions as portions of a whole.

Data Table:

Height and age data for sample students:

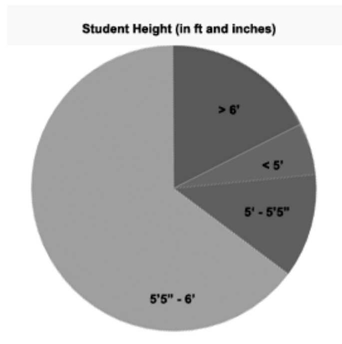
ft	in	yr	mo
5	5	18	2
5	8	20	0
5	6	19	1
5	6	18	4
5	6	19	4
5	5	19	2
6	2	18	6
6	0	18	11
5	8	17	4
5	8	18	3
5	8	19	5
5	0	19	1
5	7	19	1
5	8	19	7
6	1	17	8
5	9	18	9
5	4	18	10
5	7	19	4
5	7	18	9

Gather your own data:

The pie chart is well suited to displaying the results of the ***Separations activity*** included in this packed. Students can construct pie charts for their data set as well as one for the class as a whole and examine for differences in population distribution.

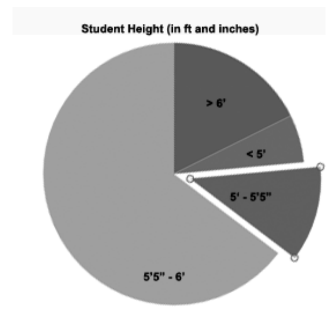
The height and age data used here was gathered by a quick survey of Oregon State University freshman chemistry students. A similar survey of your students for any categorical data (music taste, favorite tv shows, etc) could be used.

This data set and Power Point slides of the figures for classroom presentation can be downloaded at:



Pie charts use percentages to display the relative size of the component parts of a whole. The circle provides a visual example of the concept of a whole. Each segment represents a particular category, with an area that is proportional to that segments proportion of the whole. Pie charts are commonly used in mass media due to their simplicity of design and interpretation.

While popular and visually appealing, pie charts do have some limitations. To avoid an overly complex figure, their useage should be limited to fairly simple data sets with no more than six components. Pie charts are also less useful when component values are similar, as it is to difficult to distinguish the resulting difference in area.



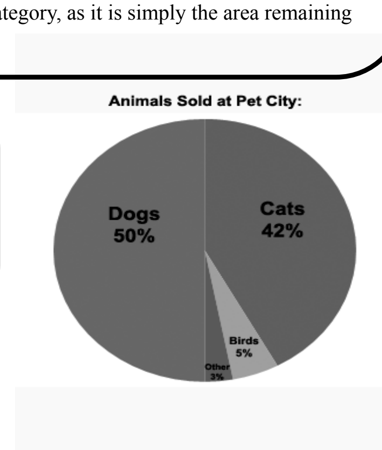
A segment of the chart is sometimes separated from the whole for emphasis of an important category.

Constructing a Pie Chart

A pie chart is constructed by converting the size of each category into a percentage of 360 degrees.

1. Find what percentage of the whole each category represents.
2. To convert the number of members of each category into a fraction, divide the number of members of that category by the total number of individuals.
3. To convert these fractions to a "slice size", multiply each by 360 degrees.
4. Draw a circle with your protractor
5. Starting from the 12 O' Clock position measure an angle equal to the size of the first segment and mark off the slice that is delineated.
6. Repeat for each category, filling in its radius according to its percentage of 360 degrees. Note that it is not necessary to measure the final category, as it is simply the area remaining after the other segments have been measured.

Labeling the segments with percentages and with category labels directly on the segments can aid quick and easy interpretation.



LINE GRAPHS

Line graphs are an easily constructed way to visually compare two variables in x-y space, and reveal data trends.

Data Table:

Sample of complete data set used to create figures at right.

year	month	day	T max	T min	T avg
1995	6	1	79	49	64
1995	6	2	81	45	63
1995	6	3	76	46	61
1995	6	4	75	50	63
1995	6	5	67	46	57
1995	6	6	60	45	53
1995	6	7	60	45	53
1995	6	8	72	44	58
1995	6	9	79	46	63
1995	6	10	76	53	65
1995	6	11	64	38	51
1995	6	12	71	44	58
1995	6	13	66	52	59
1995	6	14	58	51	55
1995	6	15	63	52	58
1995	6	16	67	52	60
1995	6	17	66	52	59
1995	6	18	65	50	58
1995	6	19	57	48	53
1995	6	20	65	50	58
1995	6	21	66	51	59
1995	6	22	69	44	57
1995	6	23	78	54	66
1995	6	24	84	53	69
1995	6	25	83	51	67
1995	6	26	83	51	67
1995	6	27	84	58	71
1995	6	28	88	56	72
1995	6	29	93	66	80
1995	6	30	94	56	75

The climate data used here was collected by the Oregon Climate Service at Oregon State University. (www.ocs.orst.edu)

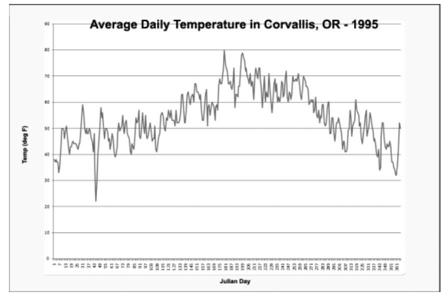
Included in the full dataset is daily max temperature, minimum temperature, average temperature, rainfall and snowfall for Corvallis, OR from 1893 – 2010. The full data set can be used to construct plots of manthly and annual average temperatures to illustrate the concept of temporal scaling of trends.

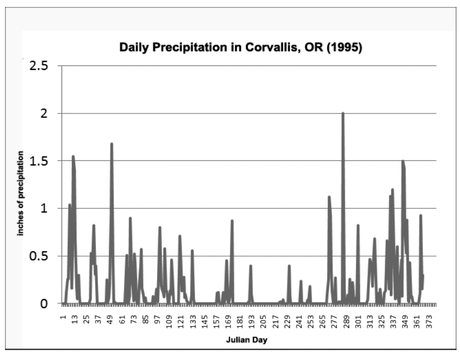
This data set and Power Point slides of the figures for classroom presentation can be downloaded at:

Gather your own data:

Have your students collect temperature, rainfall or other climate data and plot to reveal seasonal trends and compare to historical records.

Line graphs are especially popular in the field of science because they show data trends and cycles. Line graphs are particularly useful for temporal or time-series data, where the order of the data points is important. Strength of correlation for temporal data often comes from showing patterns over long periods of time. This often means that there are many data points represented in a single graph.





Line Graph Tips

- Choose an appropriate scale for each axis. Your data should occupy most of the chart area, not be squished in the corner!

- If plotting a large number of data points in Excel, do not use large data ticks, but let the connecting line itself indicate the position of the data points.

- Avoid "chartjunk", or unnecessary graphics or colors that add flash without increasing the communications effectiveness of the chart.

Constructing a Line Graph

Line graph activity:

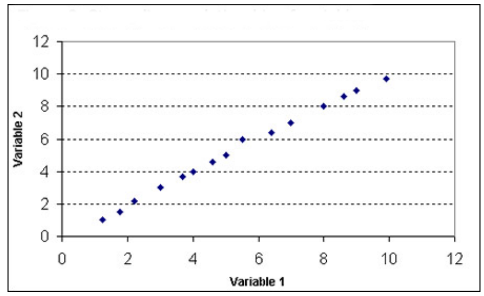
Shooter Activity Tie-In:

SCATTERPLOTS

Scatterplots compare two related variables in x-y space, helping to confirm and describe the nature of the relationship.

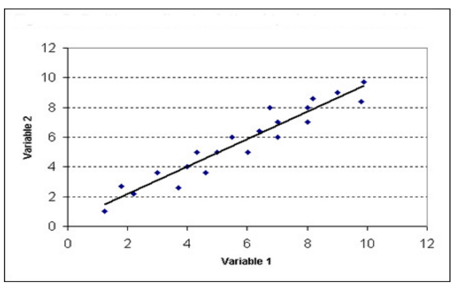
Scatterplots are widely used when measurements are made of two experimental quantities that are known or though to be related. Experimental data is plotted by placing the independent variable (typically varied in a controlled manner during the experiment) on the x-axis, and the dependant variable (typically the response measured during the experiment) on the y-axis.

The closeness of the data points to the plotted line indicates the strength of the relationship between the two variables. The closer the data points lie to the plotted line, the stronger the relationship between the variables represented on the axes.

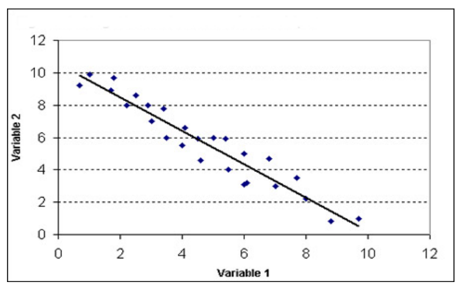


In a scatterplot, the data points are often plotted but not joined, indicating that it is experimental and not continuous data. When plotted, the resulting pattern tells us about the type and strength of the relationship between the two variables. A trendline can be calculated, or even hand drawn with a ruler to find the underlying trend .

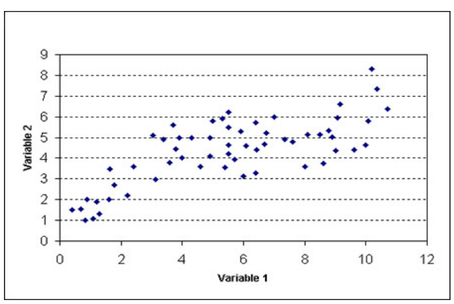
Common Data Relationships



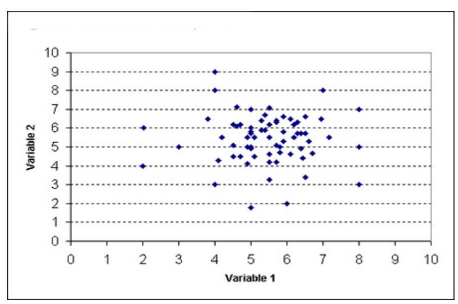
A **positive** or **direct correlation** is indicated when the points cluster around a line that runs from the lower left to upper right. An increase in x is associated with an increase in y.



A **negative** or **inverse correlation** is indicated when the points cluster around a line that runs from the upper left to lower right. An increase in x is associated with a decrease in y.



A **weak correlation** is indicated when the points only show a loose or vague association. A change in x is generally associated with a change in y, but the change is not predictable. A non-linear relationship may be indicated.



Scattered data points indicate no relationship between the two variables.