**GRAPHING PACKET**

Graphs are used to show numerical information in a useful format. Graphs help us understand and interpret data. They can be powerful tools for examining cycles and showing trends in our data. The three main types of graph are circle graphs, bar graphs, and line graphs.

**Circle Graphs:**

A circle graph, often referred to as a pie chart, shows different parts of the data in relation to all of the data. Each part of the circle represents a different part of the data; the entire circle represents all of the data. For example, a biologist studying a plot of land in a hardwood forest in Wisconsin found that it contained five different types of trees. The data table below summarizes the biologist’s findings.

***Wisconsin Hardwood Trees***

|  |  |  |
| --- | --- | --- |
| **Type of tree** | **Number found** | **Percentage of total** |
| Oak | 539 | 17 |
| Maple | 758 | 24 |
| Beech | 319 | 10 |
| Birch | 1327 | 42 |
| Hickory | 222 | 7 |
| **Total** | **3165** | **100** |

To aid in the interpretation of the data and to compare it with other sets of data, the biologist constructed a pie chart of the data.



**The following are important characteristics of a pie chart.**

1. Sections of the pie are accurately calculated and representative in size.
2. The sections are clearly labeled.
3. Labels, colors, or shading are used to represent each section of the circle.

**Line Graphs**:

While circle graphs are useful for depicting percentages, line graphs are used most often to demonstrate continuous change. Most scientific graphs are line graphs. Examine the following data:

***Population of the United States 1880-1990***

|  |  |  |  |
| --- | --- | --- | --- |
| **Year** | **Population (in millions)** | **Year** | **Population (in millions)** |
| 1881 | 50.2 | 1940 | 131.7 |
| 1890 | 62.9 | 1950 | 151.3 |
| 1900 | 76 | 1960 | 179.2 |
| 1910 | 92 | 1970 | 203.2 |
| 1920 | 105.7 | 1980 | 226.5 |
| 1930 | 122.8 | 1990 | 251.4 |

In the example given above, both the year and the population are variables. The factor which is *changed* or *manipulated*, in this case the year, is called the **independent variable (IV)**. The *measured* *effect* of the IV is called the **dependent variable** **(DV)**. The population is determined by the year; therefore the population is the dependent variable. Another way to think about the independent and dependent variables is to think about the amount of sleep you get. You know that how alert or tired you feel often depends on the number of hours of sleep you got the night before. The amount of sleep is the independent variable; your alertness is the dependent variable. Throughout your year of studying biology you will be asked to identify variables in many different investigations.

The following are **important** **characteristics of a line graph**:

1. The title of the graph accurately reflects the data – “The effect of IV on DV” is appropriate.
2. The independent variable is put on the (X) axis; the dependent variable is put on the (Y) axis.
3. An appropriate scale and interval are used on each axis. Interval refers to the spacing between the tick marks on each axis, and must be consistent for the entire length of the axes. For example, if 1cm represents an interval of 10 years, 30 years must be represented by 3cm.
4. Scale your graph according to your data set. To do this, for each variable subtract the lowest data value from the highest data value. The divide the difference by the number of lines (tick marks) on your graph. This spread should utilize MOST of your graph space.
5. Label each axis (including units).
6. The data pairs are plotted as accurately as possible. If more than one set of data is presented on a graph, each line must be differentiated (solid, dotted, dashes, etc.) and a legend/key provided.

**Using line graphs to summarize trends in the data:**

The purpose of a line graph is to summarize a trend in the data. If the data points are widely scattered or do not follow any general pattern, no correlation exists and you cannot draw a line. However, if the data points seem somewhat clustered and sloped in a straight upward or downward direction, a line is drawn that best approximates the pattern; it is not appropriate to “connect-the-dots”.

The purpose of the graph is to visually display relationships which may not be apparent from data tables. Experimental errors which are always present may obscure the relationships. If a linear relationship exists between the variables, a **best-fit line**, also called a linear regression, averages out the errors. There are ways of calculating a best-fit line, but most of the time a close estimation is fine. In the absence of graphing software, to estimate a line of best-fit, simple try to incorporate as many data points above and below the line, ignoring any obvious outliers. Do not try to force the line to go through any particular data points. The graph at right displays a best fit for the data points shown.

 **1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990**

**Using line graphs to make predictions**:

Graphs show trends in the data that may not be obvious from data tables. Examine the graph of the population of the U.S. from 1880-1990. Do you notice any trends? Can you predict what the population of the U.S. was in the year 2000? The process of going beyond the data points in a graph is called **extrapolation**. We can also use graphs to find data points between two sets of plotted data pairs. For example, we can read the graph to determine that the population of the United States in 1905 was approximately 84 million people. Determining data points between two sets of data pairs is called **interpolating**.

**Bar Graphs**:

Sometimes it is not appropriate to use a pie chart or line graph to represent data. Consider the following information:

***Precipitation in Hartford, Connecticut***

***April 1-15, 2006***

|  |  |  |  |
| --- | --- | --- | --- |
| **Date** | **Precipitation (cm)** | **Date** | **Precipitation (cm)** |
| April 1 | 0.5 | April 9 | 0.25 |
| April 2 | 1.25 | April 10 | 0.0 |
| April 3 | 0.0 | April 11 | 1.0 |
| April 4 | 0.0 | April 12 | 0.0 |
| April 5 | 0.0 | April 13 | 0.25 |
| April 6 | 0.0 | April 14 | 0.0 |
| April 7 | 0.0 | April 15 | 6.50 |
| April 8 | 1.75 |  |  |

It would be impossible to plot the data as represented on a pie chart. Days on whichthere was no precipitation could not be shown. It is also inappropriate to use a line graph to represent the data for precipitation. Consider, for example, the period from April 13 to April 15. If the points between April 13 and April 15 were connected, the graph would indicate that over 3 cm of precipitation fell on April 14. However, April 14 was a sunny day! A bar graph is appropriate for data that are not continuous. A bar graph is a good indicator of trends if the data are taken over a sufficiently long period of time.

The following bar graph depicts the data presented in the table above.



**The following are important characteristics of a bar graph:**

1. An appropriate scale is used on each axis.
2. Reasonable starting points are used for each axis.
3. The axes are labeled.
4. Data are accurately plotted.
5. The title of the graph accurately reflects the data presented.
6. The graph is easy to understand and interpret.

**Graphing Exercises** Names

**REVIEW:**

Circle Graphs

1. For what kind of data is a circle graph more appropriate than a bar graph?
2. The populations of 100 large African animals in a zoo are listed in the chart below.

Construct a circle graph that indicates the percentage each animal is of the total population.

Paste your graph here

|  |  |
| --- | --- |
| **Animal** | **Number of Animals** |
| Baboon | 25 |
| African Elephant | 35 |
| Leopard | 12 |
| Tiger | 16 |
| Giraffe | 12 |

Line Graphs

Examine the graph of the population of the United States from1880-1990 (page 2 of the reading)

1. Predict what the population of the United States was in the year 2010.
2. Determine the approximate population of the United States in:

 1935 1945 1985

1. What was the approximate population of the United States in 1870?
2. What will be the approximate population of the United States in 2020?

Bar Graphs

1. Explain for which kind of data a bar graph is used versus a line graph?

1. According to the bar graph of precipitation in Hartford, CT (page 3 of the reading) how many centimeters of rain fell in Hartford on April 11?

1. Are you able to see any trends in this data set? If so, what is the trend? If not, why not?
2. Can a bar graph show a trend, even if the data are not continuous? Explain.
3. Can the bar graph be used to predict precipitation in Hartford on April 20? Why or why not?

**PRACTICE**:

1. In 1989, the US. Department of the Interior reported that there were 360 endangered species of plants and animals in the United States. These endangered organisms included 32 species of mammals, 61 species of birds, 8 species of reptiles, 5 species of amphibians, 45 species of fishes, 3 species of snails, 32 species of clams, 8 species of crustaceans, 10 species of insects, 3 species of spiders, and 153 species of plants. Construct a bar graph of the total number of endangered plants and animals in 1989.

Paste your graph here

1. Of the total number of endangered species (continue from #12), 151 were vertebrates (mammals, birds, reptiles, amphibians, fishes). Construct a pie chart to compare the different types of endangered vertebrates with each other.

Paste your pie chart here

|  |  |
| --- | --- |
| **Date****(February, 1996)** | **Number of ill students** |
| 1 | 12 |
| 2 | 18 |
| 3 | 30 |
| 4 | 49 |
| 5 | 115 |
| 6 | 127 |
| 7 | 125 |
| 8 | 107 |
| 9 | 108 |
| 10 | 115 |
| 11 | 117 |
| 12 | 95 |
| 13 | 60 |
| 14 | 52 |

1. After an outbreak of influenza, a student gathered data on the number of students who became ill, until she became sick on the 14th day of her study. The information she gathered is shown below.

Create a line graph of the data by using the insert line under AutoShapes feature. Otherwise make it somewhere and paste it below.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

On what day were the most students ill?

During which period of time did the most students become ill?

What was the greatest number of students who were ill on anyone day?

Estimate the total number of students who were ill on the 15th day.