**Data Analysis & Probability**

**Vocabulary:**

* **Circle graph**: is a diagram that uses sectors of a circle to display data.
* **Line graph**: is a graph that displays data by using points joined by line segments.
* **Bar graph**: is a graph that displays data using horizontal or vertical bars.
* **Double bar graph**: is a graph that displays two sets of data.
* **Pictograph**: is a graph that uses a symbol to represent a certain number and repetitions of the symbol illustrates the data.
* **Discrete data**: data that can be counted.
* **Outcome**: is a possible result of an experiment or a possible answer to a survey question.
* **Event**: is any set of outcomes of an experiment.
* **Independent events**: are two events where the result of one event does not depend on the result of the other event.
* **Probability**: is the likelihood of a particular outcome; the number of times a particular outcome occurs, written as a fraction of the total number of outcomes.
* **Probability of an event**: is the likelihood of a particular set of outcomes occurring when conducting an experiment.

**Types of Graphs**

1. **Circle Graphs** – wonderful way of visualizing data that needs to be quickly compared but it is hard to draw them accurately and you cannot make quick calculations using these graphs.

Pros:

* shows parts of the whole
* shows amounts as percents
* sectors easily show data using relative size

Cons:

* does not show the actual number of things
* cannot be used to calculate total amounts
* difficult to draw accurately
1. **Line Graphs:** great way to represent data that occurs over time but a poor way of displaying relative amounts of things.

Pros:

* great way to show data that changes over time
* easy to draw and read
* can use a zigzag symbol on the vertical axis when data starts at a larger number
* can estimate values between and beyond data points

Cons:

* not good for comparing amounts at a set time
* Can become confusing if showing more than one set of data
1. **Bar Graphs:** represent different amounts of various things and to compare relative amounts but they do not show percent.

Pros:

* the heights of the bars can be used to compare responses
* the vertical axis uses a simple scale used to calculate each response and total amounts
* easy to draw

Cons:

* can be difficult to read accurately
* does not show percents
1. **Double Bar Graph:** can represent and compare two sets of discrete data, but you have to be careful when labeling your scale or they will be hard to read accurately.

Pros:

* displays two sets of data that can be counted
* lengths of bars are easily used to represent and compare data
* easy to draw and read
* great way to visualize differing amounts

Cons:

* can only be used to show discrete data
* can be difficult to read accurately depending on the scale used
1. **Pictograms/Pictographs:** visually appealing but can be hard to read and draw accurately.

Pros:

* the length of each row symbols shows an immediate comparison of responses
* visually appealing
* key is used to easily calculate responses

Cons:

* lots of symbols can make it hard to read
* difficult to draw all the symbols
* does not show percents
1. **Double Line Graphs:** uses two lines to represent changes of two sets of data over time.
2. **Trends:** We can think of a trend as the general direction that a set of data is going; like the increasing use of social media over the past decade.

Line graphs & other types of graphs can display trends.

Example: we predict the trend would be increasing. The temperature will continue to rise over the next few days.

1. **Intervals:** this is a spread between the smallest and the largest number in a range of numbers.

Example:

|  |  |  |
| --- | --- | --- |
| Interval | Tally | Frequency |
| 120-129cm130-139cm140-149cm150-159cm | IIIII IIIIIIIIII IIIII IIIIIIII | 101341 |

The interval would be 10 for each data range.

These compare sets of data not individual pieces of data.

1. **Distortion:** to distort something means to change the appearance of that thing or twist the meaning of something in a way that is misleading.

Example: two graphs displaying the exact same data but using different scales.



Even though the data is exactly the same, the way it is shown (due to the scale used) can make data look more dramatic than they really are or less impressive.

Watch out for this kind of data representation when reading about things like nutritional data of foods, types of investments to make, how long a product will last, etc. Advertises skew scales to make you more likely to buy their stuff!