Statistics is the study of the collection, organization, analysis, and interpretation of data. It also deals with surveys and experiments.

The study of statistics can be divided into two main areas:

1. Descriptive statistics has to do with collecting, organizing, summarizing, and presenting data (information).
2. Inferential statistics has to do with drawing inferences or conclusions, about populations based on information from samples.

Terminology –

Population: includes "All" items of interest

Sample: includes "some" items of interest

Raw data: information collected but not organized/processed

Quantitative data: numerical data - numbers

Qualitative data: non-numerical data - names, words

Quantitative data: The number of siblings in ten different families: 3, 1, 2, 1, 5, 4, 3, 3, 8, 2

Qualitative data: The makes of six different automobiles: Toyota, Ford, Nissan, Toyota, Chevrolet, Honda

In this section, we are going to look at seven different ways to visually display data:

Frequency Distribution
Grouped Frequency Distribution
Histogram
Bar Graphs
Line Graphs
Stem and Leaf Displays
Circle Graphs

EXAMPLE: Twenty-five students in a math class were polled or asked how many siblings were in their individual families. Here were their responses:

2, 3, 1, 3, 3, 5, 2, 3, 3, 1, 1, 4, 2, 4, 2, 5, 4, 3, 6, 5, 1, 6, 2, 2, 2

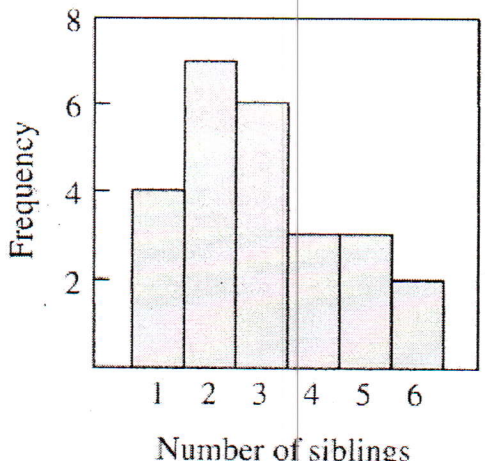
The data ranges from 1 to 6 with many of the same numbers repeated. So we can view the frequency of the data in a Frequency Distribution table.

Frequency f is counting. Counting how many times specific data is repeated.
Relative Frequency f/n is the fraction or percentage of the data represented by each item.

Number x	Frequency f	Relative Frequency $\frac{f}{n}$
1	4	$\frac{4}{25} = 16\%$
2	7	$\frac{7}{25} = 28\%$
3	6	$\frac{6}{25} = 24\%$
4	3	$\frac{3}{25} = 12\%$
5	3	$\frac{3}{25} = 12\%$
6	2	$\frac{2}{25} = 8\%$

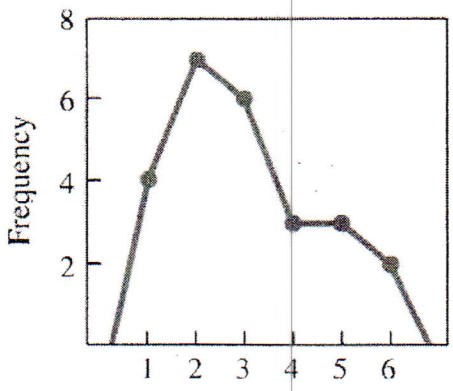
$\frac{frequency}{total\ number}(100) = \%$

The data can also be interpreted using a Histogram, which uses a series of rectangles whose height represents the frequency. All rectangles are vertical and next to each other: touching, no spaces.



Histogram

The same data can also be interpreted using a Line Graph, which is also called a Frequency Polygon, when the lines come down to touch the x – axis. Each dot is placed at the number representing the frequency and line segments connect all the dots.



Frequency polygon

EXAMPLE: Forty students were asked to estimate the number of hours they spend studying per week. Here were their responses:

18	60	72	58	20	15	12	26	16	29
26	41	45	25	32	24	22	55	30	31
55	39	29	44	29	14	40	31	45	62
36	52	47	38	36	23	33	44	17	24

The data here ranges from 12 to 72. Since there are a large number of responses then we can view the frequency of the data in a Grouped Frequency Distribution table. But first we must look at how we want to group the data.

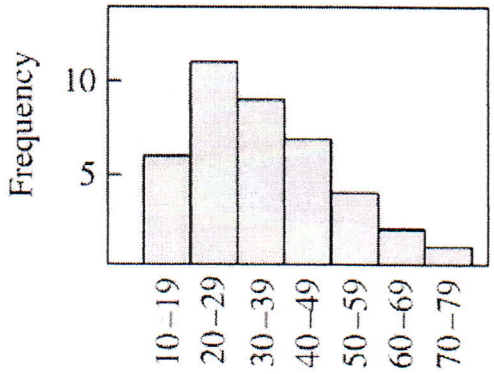
Groups of numbers are called class widths. Each class must be the same. We can use classes or groups of 10 for this data.

The lower number in each class or group is called the lower class limit

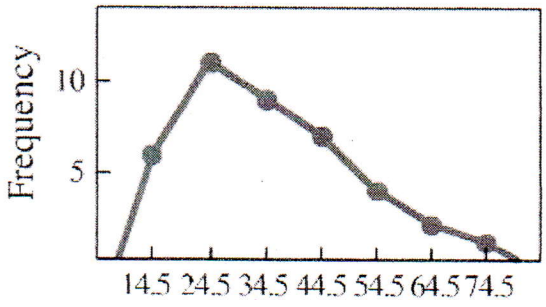
The higher number in each class or group is called the upper class limit

Class Limits	Tally	Frequency f	Relative Frequency $\frac{f}{n}$
10–19	𝍸 |	6	$\frac{6}{40} = 15.0\%$
20–29	𝍸 𝍸 |	11	$\frac{11}{40} = 27.5\%$
30–39	𝍸 ||||	9	$\frac{9}{40} = 22.5\%$
40–49	𝍸 ||	7	$\frac{7}{40} = 17.5\%$
50–59	||||	4	$\frac{4}{40} = 10.0\%$
60–69	||	2	$\frac{2}{40} = 5.0\%$
70–79	|	1	$\frac{1}{40} = 2.5\%$
	Total:	$n = 40$	

This same data can also be interpreted using a **Histogram** and a **Line Graph (frequency polygon).**



Grouped frequency histogram



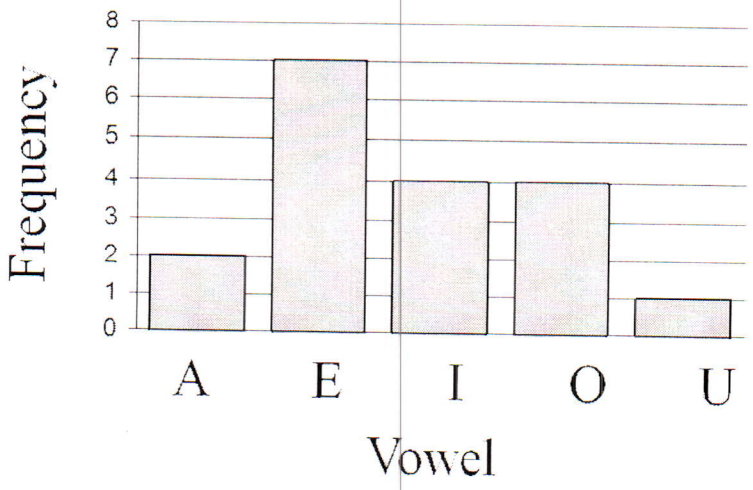
Grouped frequency polygon

Another way we can interpret this same data of the number of hours spent studying, is using a stem-and-leaf. Using the raw data, the numbers are put into groups where the tens digits are on the left representing the "stem" and the ones digits are on the right representing the "leaves".

tens	***ones***										
1	8	5	2	6	4	7					
2	0	6	9	6	5	4	2	9	9	3	4
3	2	0	1	9	1	6	8	6	3		
4	1	5	4	0	5	7	4				
5	8	5	5	2							
6	0	2									
7	2										

EXAMPLE: Bar Graphs are similar to a histogram, except, the rectangles in a bar graph do not touch. A bar graph can have rectangles that are vertical, like the example, or they can be horizontal.

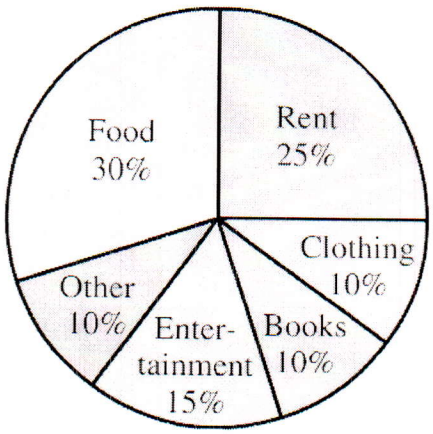
"A bar graph is given for the occurrence of vowels in this sentence."



EXAMPLE: Another alternative for displaying data is a Circle Graph or Pie Chart which uses a circle to represent the total. The circle is divided into sections or wedges that look like pie pieces. Each section is labeled with a percent of the total. Remember, percents are multiplied.

The circle graph below shows the expenses of a college student her freshman year. Suppose her total expenses for her freshman year were $12,000. How much did she spend for each category?

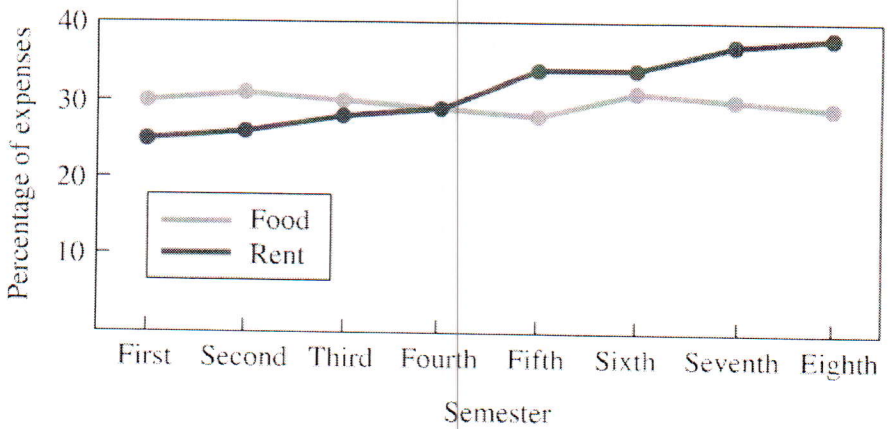
- Rent $0.25(12{,}000) = \$3{,}000$
- Food $0.30(12{,}000) = \$3{,}600$
- Entertainment $0.15(12{,}000) = \$1800$
- Books $0.10(12{,}000) = \$1200$
- Clothing $\$1200$
- Other? $\$1200$



Expense categories

Another interpretation of college expenses can be used in a _____. Line graphs have dots that can show frequency or magnitude (how much or how many). Lines connecting the dots can show rise and fall of data over time.

The line graph below shows the expenses of a college student during her four years in college.



Comparison line graph