

# Statistical Measures

Statistics is an area of math that deals with gathering information (called data). It is often used to make predictions.

## Important terms:

- **Population** – A population is an entire set of objects sharing similar characteristics, such as human beings, automobiles, or measurements, from which data can be collected and analyzed.
- **Sample** – A sample is a portion of a population from which data is collected to estimate the characteristics of the entire population.
- **Statistics** – A statistic is a measure calculated from a sample of data.
- **Parameter** – A parameter is a measure calculated from data for an entire population.

**Sample Problems:** Tell whether the following are samples or populations.

1. The cans of pop dispensed from a vending machine during its existence

*Population – all cans are being considered*

2. 100 freshmen from those enrolled at Bethel for this year

*Sample – a small number is being considered*

3. The automobiles built by Ford in the 1900s

*Population – all Fords from that time are being considered*

4. A school official gathers data about all students at North Central when studying the majors offered

*Population – all students are being considered*

5. A magazine wishes to determine the political viewpoint of college students, so it sends a questionnaire to all students at the U of M

*Sample – U of M students are a subset of all students*

### **Types of Sampling:**

Let's say we have a phone book and are going to get a sample from it. There are 4 types of samples we can get:

- **Random** – A random number generator is used to determine page number, column, and row of each person in the sample.
- **Systematic** – Every 30<sup>th</sup> person listed in the phone book is included in the sample.
- **Convenience** – A questionnaire with a return envelope is mailed out with the phone bill.
- **Cluster** – Everyone whose address indicates that he lives in a particular section of town is included in the sample.

**Sample Problems:** Identify the type of sampling.

1. A teacher who attends a workshop in her hometown

*convenience*

2. Every 5<sup>th</sup> person on the CHAT student roster

*systematic*

3. One of the CHAT classes is polled about the food choices in the café

*cluster*

4. The problem numbers for a test are written on slips of paper. Five are drawn from the slips and assigned on the test.

*random*

## Range

How many people are in your family?

**Definition:** The **range** is the difference between the largest and smallest numbers in a set of data. It measures the variation (spread) of the data.

What is the range in this class of family sizes?

## Measure of Central Tendency

In many sets, it is common for the numbers in the middle of the set to appear more frequently. This is especially true when measuring human characteristics, like height, weight, etc. The most common ways that we pick a representative number for a set of data are the mean, median, and mode. These are called the **measures of central tendency**.

**Definition**: The **mean** is the arithmetic average.

For the data set 4, 3, 7, 5, 8, 5 the **mean** is

$$\frac{4 + 3 + 7 + 5 + 8 + 5}{6} = 5.3$$

**Definition**: The **median** is the middle number in the list if the list is in numerical order. If there is an even number of data, the median is the average of the 2 middle numbers.

For the data set 4, 3, 7, 5, 8, 5 the **median** is the average of the 2 middle numbers.

3   4   5   5   7   8

$$\frac{5 + 5}{2} = 5$$

**Definition**: The **mode** is the number or numbers that appear the most. (There may be more than one mode.)

For the data set 4, 3, 7, 5, 8, 5 the **mode** is 5.

**Sample Problem**: Find the range, mean, median, and mode of the following:

13, 25, 22, 18, 17, 17, 14, 16, 22

*Solution*: In order the numbers are:

13, 14, 16, 17, 17, 18, 22, 22, 25

Range:  $25 - 13 = 12$

Mean:  $\frac{164}{9} = 18.22$

Median: 17

Mode: 17 and 22

**Note**: For the median, take the total number of numbers and divide by 2 to find the middle. If you have an even number of data in your list, then the “half of” number and the next number tell you which data to average for the median. If you have an odd number in your list, then your “half of” number will be a decimal number. You want the next greatest whole number to tell you which term is the median.

**Examples of finding the median:**

1. 1, 3, 4, 5, 5, 6, 8, 9, 10, 11, 11, 11, 12

There are 13 numbers. Take  $13 \div 2 = 6.5$  which tells me I need the 7<sup>th</sup> number in my list for the median. The 7<sup>th</sup> number in the list is 8, so 8 is the median.

2. 1, 4, 7, 9, 14, 15, 15, 16, 17, 20

There are 10 numbers. Take  $10 \div 2 = 5$  which tells me I need the 5<sup>th</sup> and 6<sup>th</sup> numbers in my list to calculate the median. The 5<sup>th</sup> number in the list is 14 and the 6<sup>th</sup> number is 15, so the median is  $\frac{14+15}{2} = 14.5$ .

**Sample Problem:** Find the value of  $x$  so that the mean is 26.

$$24, 29, x$$

Solution: Solve the equation  $\frac{x+24+29}{3} = 26$

$$x = 25$$