

Measures of Central Tendency & Dispersion

Measures that indicate the approximate center of a distribution are called **measures of central tendency**. Measures that describe the spread of the data are **measures of dispersion**. These measures include the mean, median, mode, range, upper and lower quartiles, variance, and standard deviation.

A. Finding the Mean

The **mean** of a set of data is the sum of all values in a data set divided by the number of values in the set. It is also often referred to as an **arithmetic average**. The Greek letter (“mu”) is used as the symbol for population mean and the symbol \bar{x} is used to represent the mean of a sample. To determine the mean of a data set:

1. Add together all of the data values.
2. Divide the sum from Step 1 by the number of data values in the set.

Formula:

$$\frac{\Sigma}{n}$$

Example:

Consider the data set: 17, 10, 9, 14, 13, 17, 12, 20, 14

$$\frac{\Sigma}{n}$$

The mean of this data set is **14**.

B. Finding the Median

The **median** of a set of data is the “middle element” when the data is arranged in ascending order. To determine the median:

1. Put the data in order from smallest to largest.
2. Determine the number in the exact center.
 - i. If there are an odd number of data points, the median will be the number in the absolute middle.
 - ii. If there is an even number of data points, the median is the mean of the two center data points, meaning the two center values should be added together and divided by 2.

Example:

Consider the data set: 17, 10, 9, 14, 13, 17, 12, 20, 14

Step 1: Put the data in order from smallest to largest.

9, 10, 12, 13, 14, 14, 17, 17, 20

Step 2: Determine the absolute middle of the data.

9, 10, 12, 13, **14**, 14, 17, 17, 20

Note: Since the number of data points is odd choose the one in the very middle.

The median of this data set is **14**.

C. Finding the Mode

The **mode** is the most frequently occurring measurement in a data set. There may be one mode; multiple modes, if more than one number occurs most frequently; or no mode at all, if every number occurs only once. To determine the mode:

1. Put the data in order from smallest to largest, as you did to find your median.
2. Look for any value that occurs more than once.
3. Determine which of the values from Step 2 occurs most frequently.

Example:

Consider the data set: 17, 10, 9, 14, 13, 17, 12, 20, 14

Step 1: Put the data in order from smallest to largest.

9, 10, 12, 13, 14, 14, 17, 17, 20

Step 2: Look for any number that occurs more than once.

9, 10, 12, 13, **14, 14**, **17, 17**, 20

Step 3: Determine which of those occur most frequently.

14 and **17** both occur twice.

The modes of this data set are **14** and **17**.

D. Finding the Upper and Lower Quartiles

The **quartiles** of a group of data are the medians of the upper and lower halves of that set. The **lower quartile**, Q_1 , is the median of the lower half, while the **upper quartile**, Q_3 , is the median of the upper half. If your data set has an odd number of data points, you do not consider your median when finding these values, but if your data set contains an even number of data points, you will consider both middle values that you used to find your median as parts of the upper and lower halves.

1. Put the data in order from smallest to largest.
2. Identify the upper and lower halves of your data.
3. Using the lower half, find Q_1 by finding the median of that half.
4. Using the upper half, find Q_3 by finding the median of that half.

Example:

Consider the data set: 17, 10, 9, 14, 13, 17, 12, 20, 14

Step 1: Put the data in order from smallest to largest.

9, 10, 12, 13, 14, 14, 17, 17, 20

Step 2: Identify the lower half of your data.

9, 10, 12, 13, 14, 14, 17, 17, 20

Step 3: Identify the upper half of your data.

9, 10, 12, 13, 14, **14, 17, 17, 20**

Step 4: For the lower half, find the median.

9, **10, 12**, 13

Since there are an even number of data points in this half, you will find the median by summing the two in the center and dividing by two. This is Q_1 .

—

Step 5: For the upper half, find the median.

14, **17, 17**, 20

Since there are an even number of data points in this half, you will find the median by summing the two in the center and dividing by two. This is Q_3 .

—

Q_1 of this data set is **11** and Q_3 of this data set is **17**.

E. Finding the Range

The **range** is the difference between the lowest and highest values in a data set. To determine the range:

1. Identify the largest value in your data set. This is called the **maximum**.
2. Identify the lowest value in your data set. This is called the **minimum**.
3. Subtract the minimum from the maximum.

Example:

Consider the data set: 17, 10, 9, 14, 13, 17, 12, 20, 14

Step 1: Put the data in order from smallest to largest.

9, 10, 12, 13, 14, 14, 17, 17, 20

Step 2: Identify your maximum.

9, 10, 12, 13, 14, 14, 17, 17, **20**

Step 2: Identify your minimum.

9, 10, 12, 13, 14, 14, 17, 17, 20

Step 3: Subtract the minimum from the maximum.

20–9=**11**

The range of this data set is **11**.

F. Finding the Variance and Standard Deviation

The **variance** and **standard deviation** are a measure based on the distance each data value is from the mean.

1. Find the mean of the data. (if calculating for a population or if using a sample)
2. Subtract the mean (or) from each data value (x_i).
3. Square each calculation from Step 2.
4. Add the values of the squares from Step 3.
5. Find the number of data points in your set, called n .
6. Divide the sum from Step 4 by the number n (if calculating for a population) or $n - 1$ (if using a sample). This will give you the **variance**.
7. To find the **standard deviation**, square root this number.

Formulas:			
Sample Variance, s^2 :		Population Variance, σ^2 :	
$\frac{\sum (x_i - \bar{x})^2}{n - 1}$		$\frac{\sum (x_i - \bar{x})^2}{n}$	
Sample Standard Deviation, s :		Population Standard Deviation, σ :	
$\sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}$		$\sqrt{\frac{\sum (x_i - \bar{x})^2}{n}}$	

Example: Calculate the sample variance and sample standard deviation

Consider the sample data set: 17, 10, 9, 14, 13, 17, 12, 20, 14.

Step 1: The mean of the data is 14, as shown previously in Section A.

Step 2: Subtract the mean from each data value.

$$17-14=3; 10-14=-4; 9-14=-5; 14-14=0$$

$$13-14=-1; 17-14=3; 12-14=-2; 20-14=6; 14-14=0$$

Step 3: Square these values. $3^2 = 9$; $(-4)^2 = 16$; $(-5)^2 = 25$; $0^2 = 0$; $(-1)^2 = 1$; $3^2 = 9$; $(-2)^2 = 4$; $6^2 = 36$

Step 4: Add these values together.

$$9+16+25+0+1+9+4+36=100$$

Step 5: There are 9 values in our set, so we will divide by 9 $-1=8$.

$$\frac{100}{9} = 12.5$$


Note: This is your variance.

Step 6: Square root this number to find your standard deviation.

$$\sqrt{12.5} = 3.536$$

The **variance** is 12.5 and the **standard deviation** is 3.536.


G. Using the TI-84

1. To enter the data values, press the  key. Select Edit...

```
NORMAL FLOAT AUTO a+bL DEGREE MP
EDIT CALC TESTS
1:Edit...
2:SortA(
3:SortD(
4:ClrList
5:SetUpEditor
```

2. Enter the data values in the L1 column.

```
NORMAL FLOAT AUTO a+bL DEGREE MP
L1 L2 L3 L4 L5 1
17
10
9
14
13
17
12
20
14
-----
L1(10)=
```

3. Press the  key. Under CALC, select 1-VarStats

```
NORMAL FLOAT AUTO a+bL DEGREE MP
EDIT CALC TESTS
1:1-Var Stats
2:2-Var Stats
3:Med-Med
4:LinReg(ax+b)
5:QuadReg
6:CubicReg
7:QuartReg
8:LinReg(a+bx)
9:LnReg
```

4. Make sure the List is L1 then select Calculate.

```
NORMAL FLOAT AUTO a+bL DEGREE MP
1-Var Stats
List:L1
FreqList:
Calculate
```

```
NORMAL FLOAT AUTO a+bL DEGREE MP
1-Var Stats
x̄=14
Σx=126
Σx²=1864
Sx=3.535533906
σx=3.333333333
n=9
minX=9
↓Q1=11
```

Mean → $\bar{x}=14$
Sum of all data values → $\Sigma x=126$
Sample Standard Deviation → $Sx=3.535533906$
Population Standard Deviation → $\sigma x=3.333333333$
Number of data values → $n=9$

```
NORMAL FLOAT AUTO a+bL DEGREE MP
1-Var Stats
↑Sx=3.535533906
σx=3.333333333
n=9
minX=9
Q1=11
Med=14
Q3=17
maxX=20
```

Lower Quartile → $Q1=11$
Median → $Med=14$
Upper Quartile → $Q3=17$

These could be subtracted to find the range.
Smallest data value ← $minX=9$
Largest data value ← $maxX=20$