

# Statistics Study Guide

## The basics of using our calculator:

### How to enter data into my calculator:

1. **STAT**
2. Choose 1:Edit (the default)
3. Arrow over to the list (L1-L6) you want
4. Type a number, hit **ENTER**; repeat

```

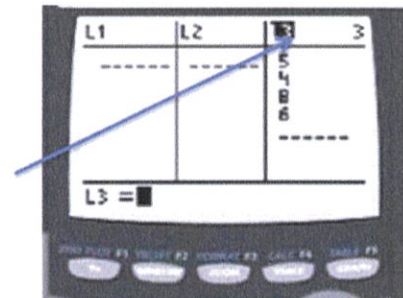
3:0000 CALC TESTS
1:Edit...
2:SortA(
3:SortD(
4:ClrList
5:SetUpEditor

L1  L2  L3  1
-----
L1(1) =
    
```



### How to clear an old list:

1. **STAT**
2. Choose 1:Edit (the default)
3. Arrow over to the list you want to clear
4. Arrow up to highlight the list name
5. **CLEAR**
6. HIT DOWN ARROW



### How to find the mean, median, standard deviation, and box plot data:

1. **STAT**
2. Arrow over to CALC
3. Choose 1: 1-Var Stats
4. **ENTER**
5. **ENTER**

```

3:0000 CALC TESTS
1:Edit...
2:SortA(
3:SortD(
4:ClrList
5:SetUpEditor
    
```

```

EDIT 3:0000 TESTS
1:1-Var Stats
2:2-Var Stats
3:Med-Med
4:LinReg(ax+b)
5:QuadReg
6:CubicReg
7:QuartReg
    
```

← to find the equation of the line of best fit + the "r" (correlation coefficient)

```

1-Var Stats
x̄=44.46666667
Σx=667
Σx²=38315
Sx=24.86496866
σx=24.02184191
n=15
    
```

mean →  $\bar{x}$   
 sum →  $\Sigma x$   
 sample standard deviation →  $S_x$   
 population standard deviation →  $\sigma_x$   
 # of terms →  $n$

```

1-Var Stats
n=15
minX=4
Q1=30
Med=35
Q3=55
maxX=110
    
```

minimum value → minX  
 1<sup>st</sup> quartile → Q1  
 2<sup>nd</sup> quartile: Median → Med  
 3<sup>rd</sup> quartile → Q3  
 maximum value → maxX

\* To turn on the "r" **mode** + go to STAT Diagnostics + turn it **on**

## Statistics Vocabulary

**Mean-** The **average** of the numbers: a calculated “central” value of a set of numbers

**Median-** Is the “**middle**” value in the list of ordered numbers

**Mode-**The number that appears **most** often in a set of numbers

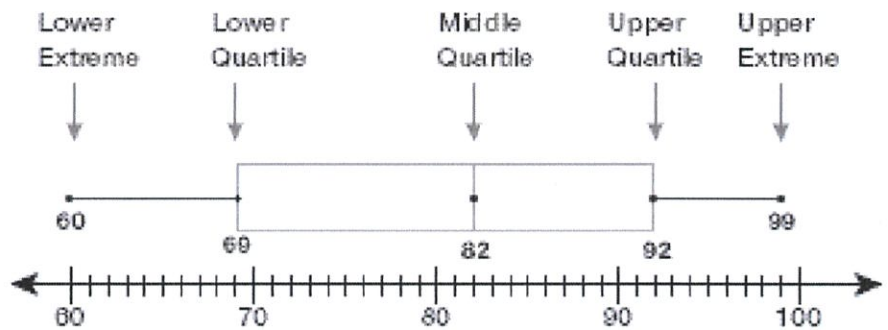
**Interquartile Range-** is a measure of variability where you find the difference between the upper quartile (Q3) and the lower quartile (Q1)

**Standard Deviation-**is a measure of variability that describes how closely a set of data clusters about the mean

**Correlation Coefficient-** the correlation coefficient, denoted  $r$ , is the measure of how well a collection of data points can be modeled by a line. Correlation coefficients range between  $-1$  and  $1$ . The closer the correlation coefficient is to  $1$  or  $-1$ , the less scattered the points are and the stronger the relationship. Only data points on the scatter plot which are a perfectly straight line can have  $r = -1$  or  $r = 1$

### Plot the 5 values from the Statistical Summary to Create a Box Plot:

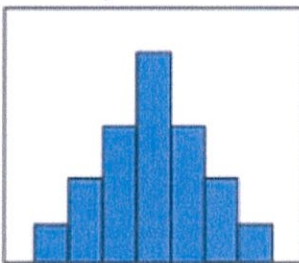
Min X = 60  
Q1 = 69  
Med = 82  
Q3 = 92  
Max X = 99



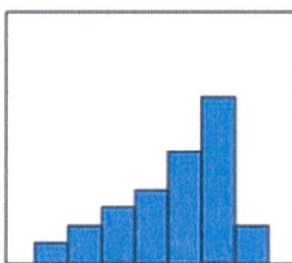
### Interquartile Range (IQR): $Q3 - Q1$

Ex from the right:  $92 - 69 = 23$

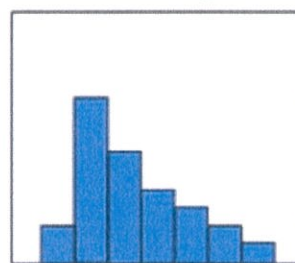
Symmetric



Skewed Left



Skewed Right



When the data is symmetrical- the **mean** would be a typical value to use.

When the data is skewed left/skewed right- the **median** would be a typical value to use since there is an outlier

## Linear Regressions and Correlation Coefficients

To get the linear regression equation you must type the data into the calculator by using STAT- EDIT- L1 and L2. Then hit STAT-CALC-4. Copy down the a and b value and plug into the given equation ( $y = ax + b$ )

\*\*If the question asks “**what is the correlation coefficient**” and gives you a set of points (points in a table) then you **MUST** make sure **STAT DIAGNOSTICS** is turned **ON**. You need this on to see the “**r**” value appear when you pull up your linear regression equation.

click

mode

Scroll down to the bottom until you see STAT DIAGNOSTICS. Then use the arrow to highlight over ON and hit enter.