

AP Statistics – Summer Project

Name _____

Summer 2020

Date Received _____



FOR MORE INFORMATION, please contact:

Mr. Sean Powers

Director of Science, Technology, Engineering, and Mathematics (STEM)

spowers@carteretschools.org

732-541-8960, ext 6032

Dear Student,

The AP Stats summer project has three parts. Part One is an online tutorial. Part Two is a series of questions that are related to the tutorial. Part Three is a problem set. Please note that the file is somewhat large but that a significant portion of the file contains notes and images. I also used a large font. You may want to complete the project in a notebook and print out only the pages that you need.

If you have trouble obtaining the file, then you must email me no later than July 10 at the email address given below. The expectation is that you will know the concepts and the vocabulary that are presented in the tutorial and in this packet. Please do your best to utilize the online graphing calculator. I have put detailed notes about the calculator at the end of the packet. You may want to read through the packet before you begin working on the project. You must complete this packet. It will serve as your study guide. You will be tested on this material soon after the start of the school year. If you have any questions or concerns, then please email me at lmasse@carteretschools.org. I have established a Google Classroom so that we can communicate about this project over the summer.

AP Statistics Summer Project Fun - 2020

Class code: falayjj

Have a great summer.

Dr. L. Masse

You will need to be able to access the online graphing calculator. It is located on the NJSLA website.

You can access the online graphing calculator at the link given below. You will need to start the test for Algebra II. The calculator icon will be at the top of the screen. You do not log into the test.


<https://nj.testnav.com/client/index.html#login?username=LGN688322866&password=3SJ3PDKX>

***Indicates that I want you to use the TI-84 calculator, **in addition to doing the calculations or graphs without the aid of a calculator**. The instructions for the calculator are at the end of this packet.

Part One: You are to complete the tutorial for AP Stats that is available at www.stattrek.com. The link is:

<http://stattrek.com/tutorials/ap-statistics-tutorial.aspx>

AP Statistics Tutorial



The image shows a screenshot of the AP Statistics Tutorial navigation menu. The menu is titled "AP Statistics Tutorial" in a blue header. Below the header, the menu is organized into three main sections: "Exploring Data", "Charts and graphs", and "Regression". Each section is preceded by a small downward-pointing triangle icon. The "Exploring Data" section includes "The basics", "Charts and graphs", and "Regression". The "Charts and graphs" section includes "Patterns in data", "Dotplots", "Histograms", "Stemplots", "Boxplots", "Cumulative plots", "Scatterplots", and "Comparing data sets". The "Regression" section is currently collapsed.

AP Statistics Tutorial	
Exploring Data	
▾ The basics	
Variables	
Population vs sample	
Central tendency	
Variability	
Position	
▾ Charts and graphs	
Patterns in data	
Dotplots	
Histograms	
Stemplots	
Boxplots	
Cumulative plots	
Scatterplots	
Comparing data sets	
▾ Regression	

You can stop after the section on comparing data sets.

- **Sample problems.** Most of the lessons include sample problems. The sample problems help you test your knowledge. They also illustrate shortcuts and solutions to common statistics problems. You are responsible for the content that is presented in the tutorial prior to the regression unit.

Part Two: Please provide a detailed definition and an example for the terms that are given below. You may have to search beyond the assigned sections of the tutorial to find some of the definitions.

There is a statistics dictionary on stattrek.com. The link is below.

<http://stattrek.com/statistics/dictionary.aspx>

1. Categorical Variables

Example:

2. Quantitative Variables

Example:

3. Univariate Data

4. Bivariate Data

5. Median

6. Mean

7. Population

Example:

8. Sample

Example:

9. Center

10. Spread

11. Symmetry

12. Unimodal and Bimodal

13. Skewness

Sketch Skewed left:

Sketch Skewed right:

14. Uniform

15. Gaps

16. Outliers

17. Dotplots

18. Difference between bar chart and histogram

19. Stemplots

20. Boxplots

21. Quartiles

22. Range

23. Interquartile Range

24. Parallel boxplots

25. Parameter

26. Statistic

27. Two- Way Table

28. Marginal Distribution

29. Conditional Distribution

30. Standard Deviation

31. Standard Score

32. Z-score

33. SOCS – Shape, Outliers, Center, and Spread – You describe the distribution of a set of data by looking at these key elements of a graph or picture of a distribution. You can also use SOCS to compare two or more distributions.

Part Three – Problem Set

P1.

Please explain. Do not simply label as C or Q. Give a reason or examples.

CATEGORICAL OR QUANTITATIVE

Determine if the variables listed below are *quantitative* or *categorical*.

1. Time it takes to get to school
2. Number of people under 18 living in a household
3. Hair color
4. Temperature of a cup of coffee
5. Teacher salaries
6. Gender
7. Smoking
8. Height
9. Amount of oil spilled
10. Age of Oscar winners
11. Type of Depression medication
12. Jellybean flavors
13. Country of origin
14. type of meat
15. number of shoes owned

P2.

STATISTIC – WHAT IS THAT?

A statistic is a number calculated from data. Quantitative data has many different statistics that can be calculated. Determine the given statistics from the data below on the number of homeruns Mark McGuire has hit in each season from 1982 – 2001.

70	52	22	49	3	32	58	39
39	65	42	29	9	32	9	33

Mean	
Minimum	
Maximum	
Median	
Q1	
Q3	
Range	
IQR	

***Put the data into L1 and calculate One-Variable Statistics. You may also want to practice sorting the data.

P3.

IT'S A TWISTA

The data below gives the number of hurricanes that happened each year from 1944 through 2000 as reported by *Science* magazine.

3	2	1	4	3	7	2	3	3	2	5	2	2	4	2	2	6	0	2	5	1	3	1	0
3	2	1	0	1	2	3	2	1	2	2	2	3	1	1	1	3	0	1	3	2	1	2	1
1	0	5	6	1	3	5	3																

Make a dotplot to display these data. Make sure you include appropriate labels, title, and scale.

P4.

SHOPPING SPREE!

A marketing consultant observed 50 consecutive shoppers at a supermarket. One variable of interest was how much each shopper spent in the store. Here are the data (round to the nearest dollar), arranged in increasing order:

3	9	9	11	13	14	15	16	17	17
18	18	19	20	20	20	21	22	23	24
25	25	26	26	28	28	28	28	32	35
36	39	39	41	43	44	45	45	47	49
50	53	55	59	61	70	83	86	86	93

Make a stemplot using tens of dollars as the stem and dollars as the leaves. Make sure you include appropriate labels, title and key.



P5.

SSHA SCORES

Here are the scores on the Survey of Study Habits and Attitudes (SSHA) for 18 first-year college women:

154 109 137 115 152 140 154 178 101 103 126 126 137 165 165 129 200 148

and for 20 first-year college men:

108 140 114 91 180 115 126 92 169 146 109 132 75 88 113 151 70 115 187 104

- a. Put the data values in order for each gender. Compute numeral summaries for each gender.

Women		Men	
Mean		Mean	
Minimum		Minimum	
Q1		Q1	
Median		Median	
Q3		Q3	
Maximum		Maximum	
Range		Range	
IQR		IQR	

- b. Using the minimum, Q1, Median, Q3, and Maximum from each gender, make parallel boxplots to compare the distributions.

***You can sort data that are entered into a list. You can calculate the One Variable Statistics that are needed for a box plot (box and whisker). You also create a box plot using the TI-84. You can create stacked box plots as well. Again, you will have to set your window.

Please review the definitions and examples on www.stat Trek.com for the following seven sampling methods or designs:

P6.

The 7 types of sampling designs are:

- A. voluntary response B. convenience C. simple random D. stratified
E. cluster F. multistage G. systematic

1. The Maryland division of Weight Watchers is doing research to determine how many people on the Weight Watchers diet cheat at least once a week. They decide that anonymous surveys will give them an accurate representation but do not have time to get responses from ALL the Maryland Weight Watchers people.

Read the scenarios below and determine which of the 7 sampling methods best describes it.

- _____ I. Randomly select 10 members from each of the branches in the Maryland division.
- _____ II. Use an alphabetical listing of all Maryland division members. Randomly choose a starting person on the list. Then select every 20th person thereafter.
- _____ III. Randomly select 2 or 3 branches of the Maryland division and survey every member of that branch.
- _____ IV. Send out the survey to every member of the Maryland division. Place drop boxes in each WW center. Anyone who returns a survey will be in the sample.
- _____ V. The Maryland regional office is in Baltimore so they survey members at the WW center in Baltimore.
- _____ VI. From a numbered list of all Maryland division members use a computer to randomly select 100 numbers and survey all members with those corresponding numbers.

2. What is the population of interest in the WW situation?

P7.

It is expected that you have a thorough understanding of linear functions.

1. The USDA reported that in 1990 each person in the United States consumed an average of 133 pounds of natural sweeteners. They also claim this amount has decreased by about 0.6 pounds each year.
 - a. Write a linear equation that relates years since 1990 to the average consumption of natural sweeteners. Define your variables.
 - b. What is the slope and what is the y-intercept?
 - c. Predict the average consumption of sweeteners per person for the year 2005.

2. The following equation can be used to predict the average height of boys anywhere between birth and 15 years old: $y = 2.79x + 25.64$, where x is the age (in years) and y is the height (in inches).
 - a. What does the slope represent in this problem? Interpret it in context.
 - b. What does the y-intercept represent in this problem? Interpret it in context.

P8.

1. Restaurants are often ranked using star system (1-5), with five stars being the highest rating. The ratings are based on a number of measures. Give two examples of quantitative measures that might be used to indicate quality. Give two examples of categorical or qualitative measures that might be used to indicate quality.

P9. A school district looked at student absences for the entire year and found the following:

Day	Number of Absences
Mo	125
Tu	59
We	45
Th	65
Fr	134

Present the data in a well-organized bar graph. Would it also be correct to make a pie graph?

Explain. Give some reasons to explain the increase in absences on Mo and Fr.

P10. Have you bowled in the last year? Here are the percentages of people between the ages of 5 and 29 who said that they bowled last year in Flatland based on a survey of all residents.

Age Group	Bowled Last Year
5-9 years	15%
10-14 years	23%
15-19 years	55%
20-24 years	47%
25-29 years	71%

Kara made a circle graph or pie graph of these data. Why is this incorrect?

Create a well-labeled bar graph for the data. Be sure to include labels and a title.

P11.

You are given the following set of data. Create a boxplot using the five-number summary. You will need the min, Q1, Q2, Q3, and the max.

4, 5, 5, 5, 5, 7, 7, 7, 7, 8

What happens to the quartiles when we add the number 30 (outlier) to this data set? What happens to the mean? What happens to the IQR (Q3-Q1)?

Measures that are not greatly affected by outliers are said to be **resistant**. Which of our above measures are resistant?

What would the boxplot, or the box and whisker graph, look like for this new set of data? How would you describe the shape? How does it compare to the original boxplot?

P12.

Outlier: A score X is an outlier if falls more than 1.5 * IQR units from either $Q1$ or $Q3$.

$$X < Q1 - 1.5 * IQR \text{ or}$$

$$X > Q3 + 1.5 * IQR$$

Use this rule to show that 30 is an outlier in the previous question.

P13.

Who sends/receives more text messages? Is the average for females greater than the average for males? Compare the distributions using SOCS. You need to know whether or not the distributions are skewed or symmetric.

The data - Number of texts S/R over a two-day period for a random sample of HS students

Males:

127, 44, 28, 83, 0, 6, 78, 6, 5, 213, 73, 20,
214, 28, 11

Females:

112, 203, 102, 54, 379, 305, 179, 24, 127, 65, 41, 27,
298, 6, 130, 0

You need to plot the data first. What type of plot should you use? What are your choices?

P13. (partial answer):

Shape: Both distributions are skewed to the right and the mean is greater than the median. The mean is not resistant. The mean is sensitive to outliers.

Center: Females text more than males with a median of 107 messages versus 28 messages. The median value for females is about 4 times the value for males.

We can also note that the median value for females is greater than the third quartile for males (Q3).
What does this mean?

We can say that 75% of males text less than the average female.

Spread: There is much more variability in the data for females than in the data for males. Please look at the lengths of the two plots. The IQR is greater for females and the whiskers are longer as well. The IQR for females is 157 texts and the IQR for males is 77 texts. This “range” is about twice as large.

Specifically, 50% of the data for males are between 6 and 83 and 50% of the data for females are between 34 and 191 (more spread).

Outliers: There are two outliers in the male distribution and no outliers in the female distribution.

Please be very careful to state that you are comparing the medians in this case. You cannot just simply state that you are comparing the averages. Which measure of central tendency did you choose and why did you choose it?

Histograms

- Quantitative variables often take many values. A graph of the distribution may be clearer if nearby values are grouped together.
- The most common graph of the distribution of one quantitative variable is a **histogram**.

How to Make a Histogram

- Divide the range of data into classes of equal width.

- Find the count (*frequency*) or percent (*relative frequency*) of individuals in each class.
- Label and scale your axes and draw the histogram. The height of the bar equals its frequency. Adjacent bars should touch, unless a class contains no individuals.

The following table presents data on the percent of households from each state that have more than three pets.

2.8, 3.9, 4.1, 4.7, 5.5, 5.7, 6.2, 7.9, 8.0, 8.2, 8.6, 8.8, 9.3, 9.6, 10.1, 10.5, 10.7, 11.9, 12.0, 12.1, 12.3, 12.5, 13.1, 13.1, 14.4, 14.9, 15.1, 15.1, 15.5, 15.7, 15.7, 15.9, 16.0, 16.3, 17.4, 17.9, 18.4, 18.8, 19.1, 19.6, 19.9, 20.1, 20.4, 21.5, 21.7, 22.3, 22.9, 23.8, 24.1, 24.3

2, 6, 10, 14, 9, 7, 2 (class counts from above)

- The maximum value is 24.3 and the minimum value is 2.8.
- We can choose 7 classes of width equal to 4.
- Class One: $0 \leq X < 4$
- Class Two: $4 \leq X < 8$
- Class Three: $8 \leq X < 12$
- Class Four: $12 \leq X < 16$
- Class Five: $16 \leq X < 20$
- Class Six: $20 \leq X < 24$
- Class Seven: $24 \leq X < 28$
-

Notice that each class is closed on the left and open on the right.

<https://statistics.laerd.com/statistical-guides/understanding-histograms.php>

Notice that, unlike a bar chart, there are no "gaps" between the bars (although some bars might be "absent" reflecting no frequencies). This is because a histogram represents a continuous data set, and as such, there are no gaps in the data.

To begin, decide how many intervals you would like in your graph. A good rule of thumb is to use the

square root of the number of observations (after rounding). Here, that is the square root of $50 = 7.07$; round and use 7.

The interval width should then be approximately equal to the range divided by the number of intervals. $\text{Range}/\# \text{ Intervals} = 21.5/7$. The width is equal to 3.07. We will use 4 as the width. We can see that using 3 will result in more than 7 classes.

There are a number of rules (one is given above) to help you choose the number of classes and the class widths. However, no rule will work all of the time. You are generally free to choose the number of classes that you want for your histogram. Remember that you want to see the distribution of the data. Choosing too few or too many classes will not allow you to properly see the distribution.

P14.

Complete the frequency table and create the histogram for the above set of data. Tables and graphs must have a title and you must label all axes. Describe the distribution using SOCS.

Frequency Table:

Pets per Household in the United States, 2015

<u>Number of Pets (X)</u>	<u>Frequency</u>
$0 \leq X < 4$	2
$4 \leq X < 8$	6
$24 \leq X < 28$	2

Histogram:

Pets per Household in the United States, 2015

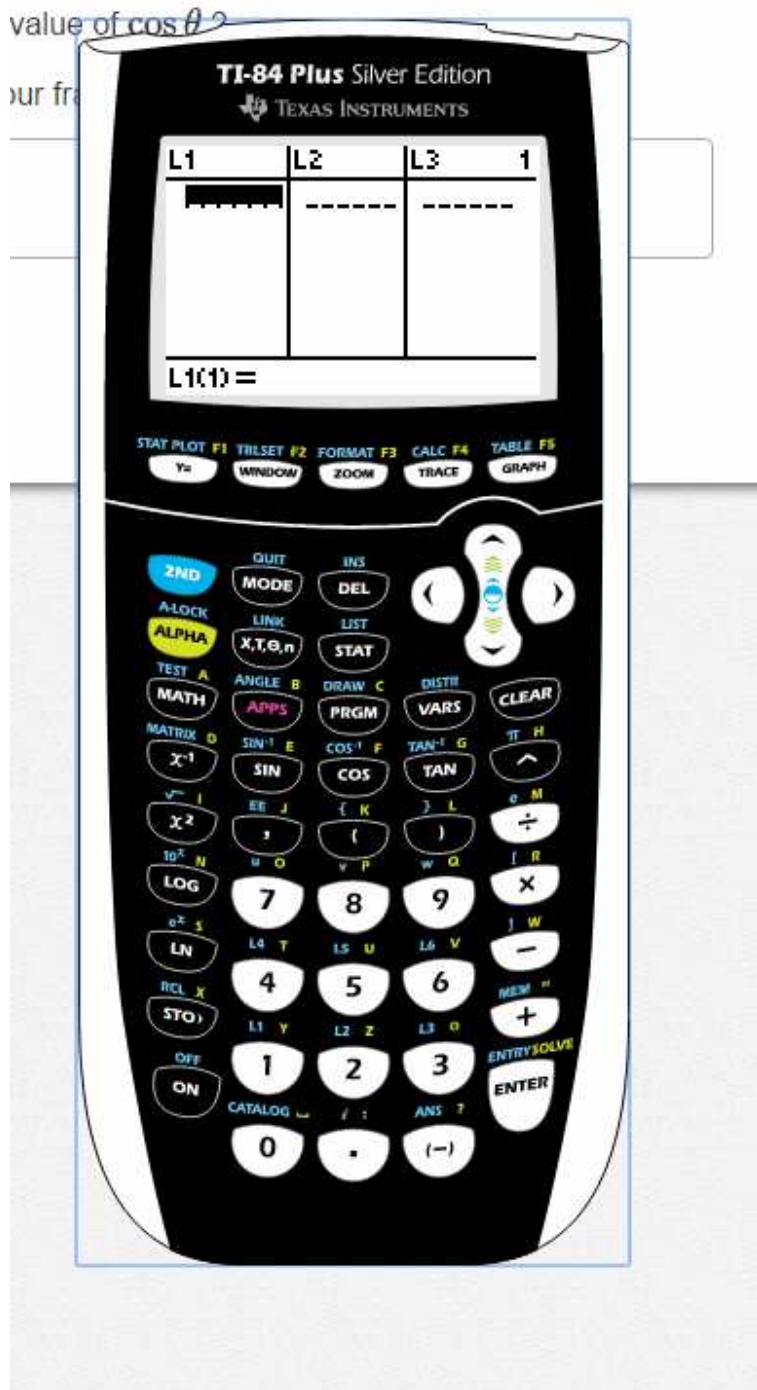
P15.

You should have created the above histogram without using the TI-84 calculator. Can you create one using the calculator?

1. Access the calculator.

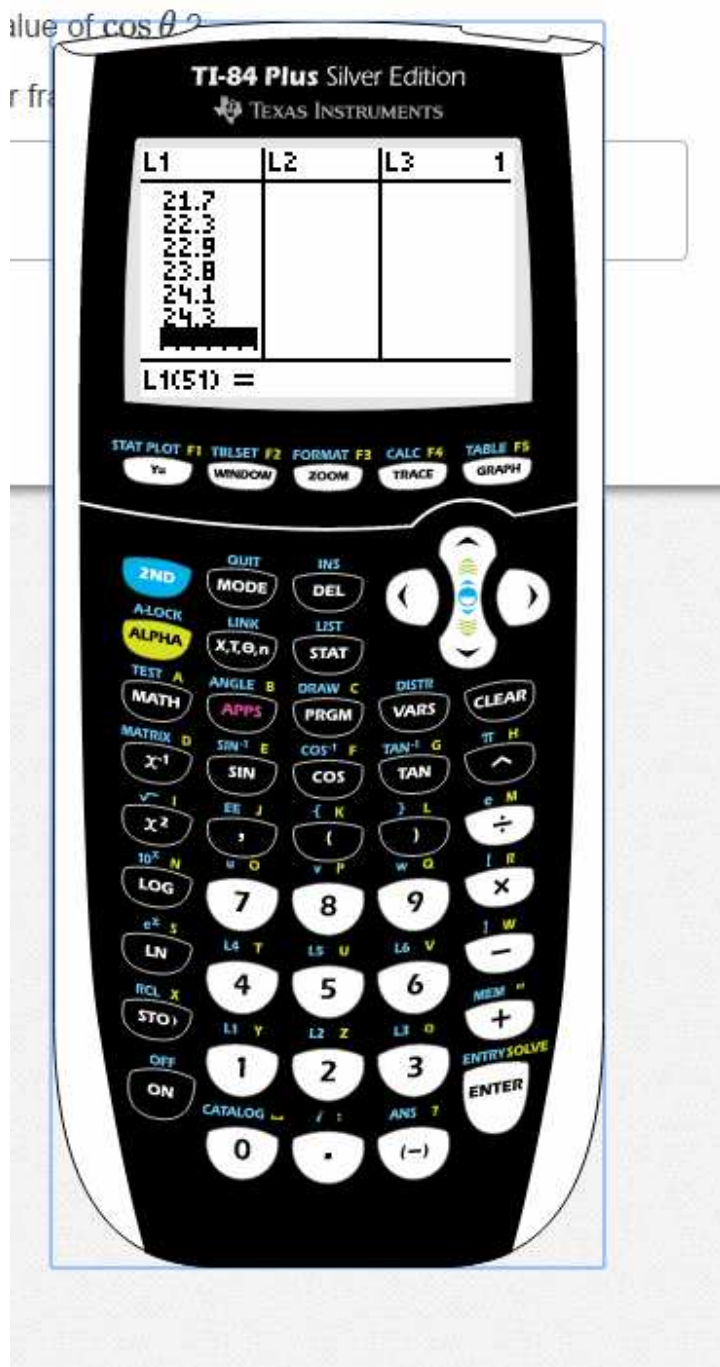


2. Stat – Edit

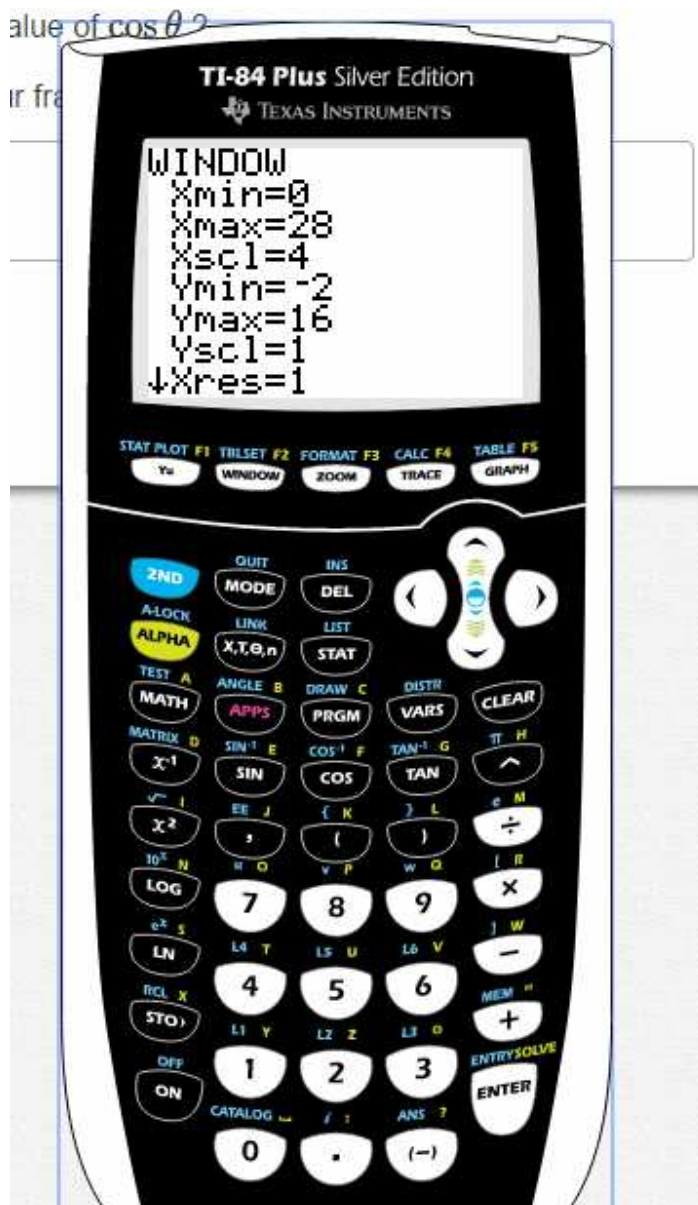


You can now enter the data into L1.

3. Enter the data.



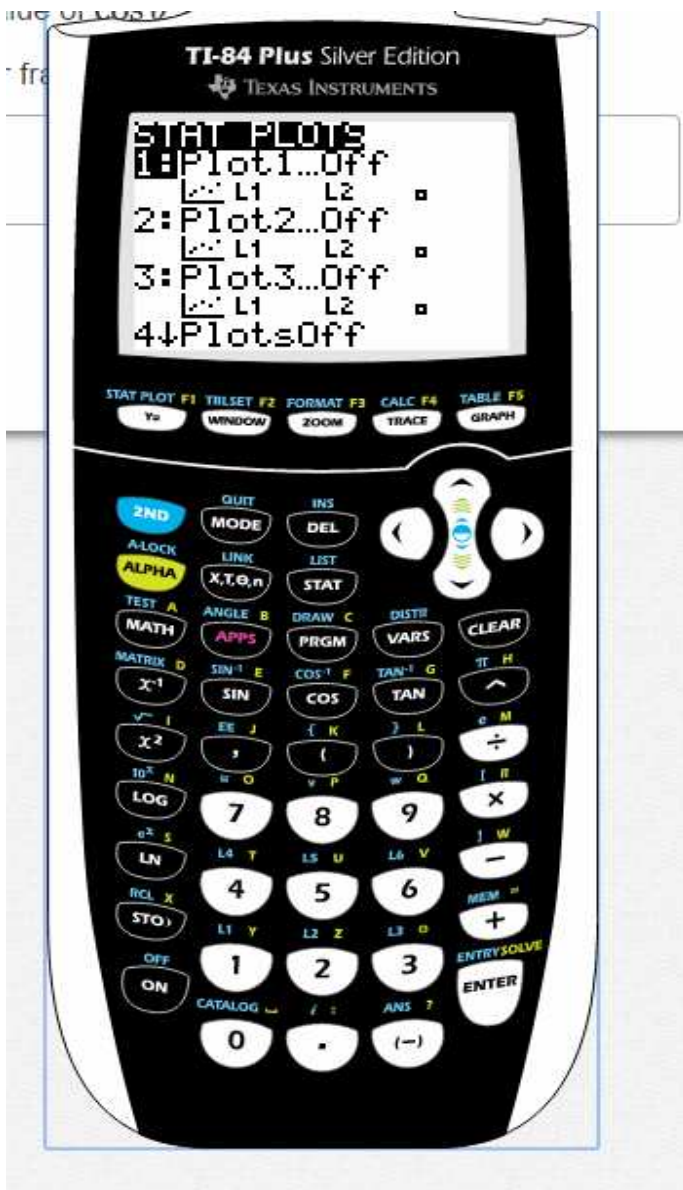
4. Set the window for the graph. Choose Window and set the window using the information from the frequency table.



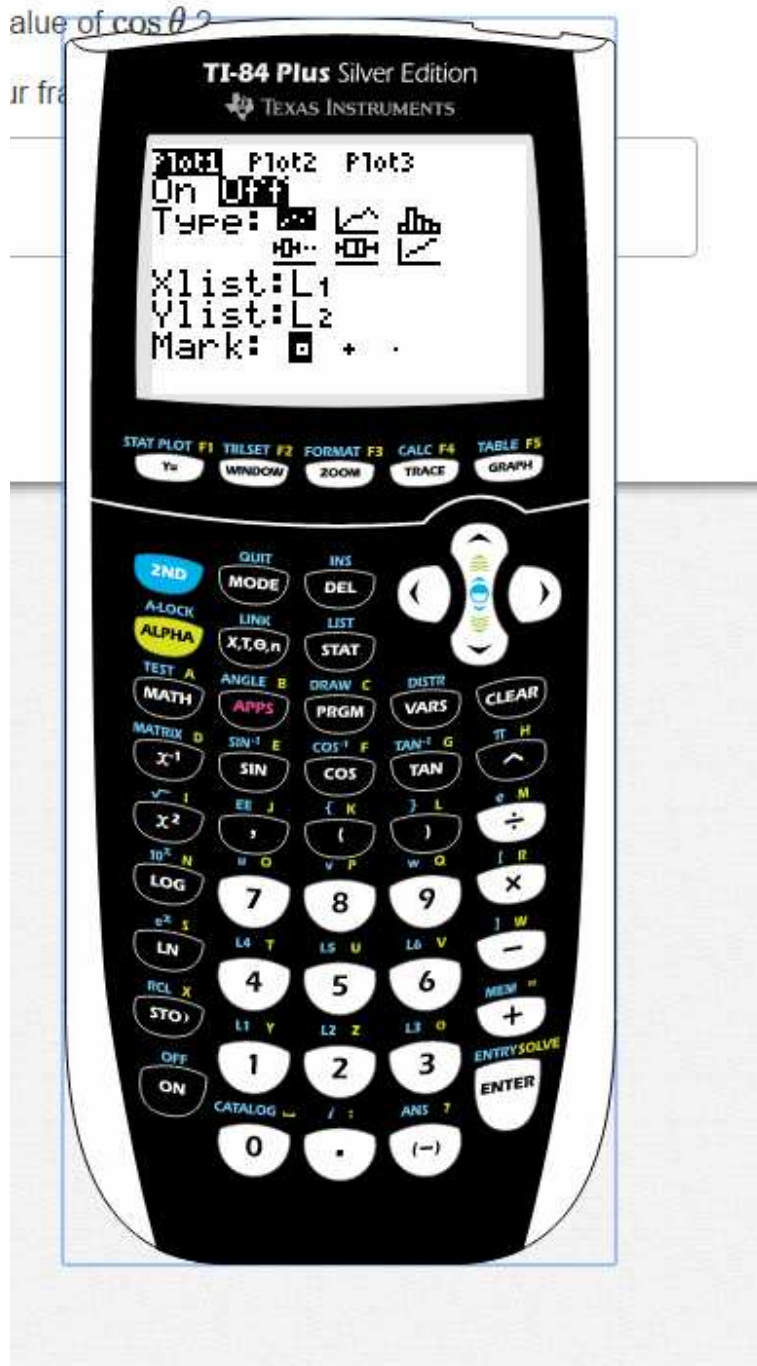
The first class starts at 0 (Xmin) and the last class ends at 28 (Xmax). The width is 4 (Xscl).

The Y values are the frequencies. The maximum frequency is 14 and the minimum frequency is 2. I chose $Y_{\min} = -2$ and $Y_{\max} = 16$ so that I can better see the graph. $Y_{\text{scl}} = 1$.

5. Create the histogram. 2nd Function – Stat Plot



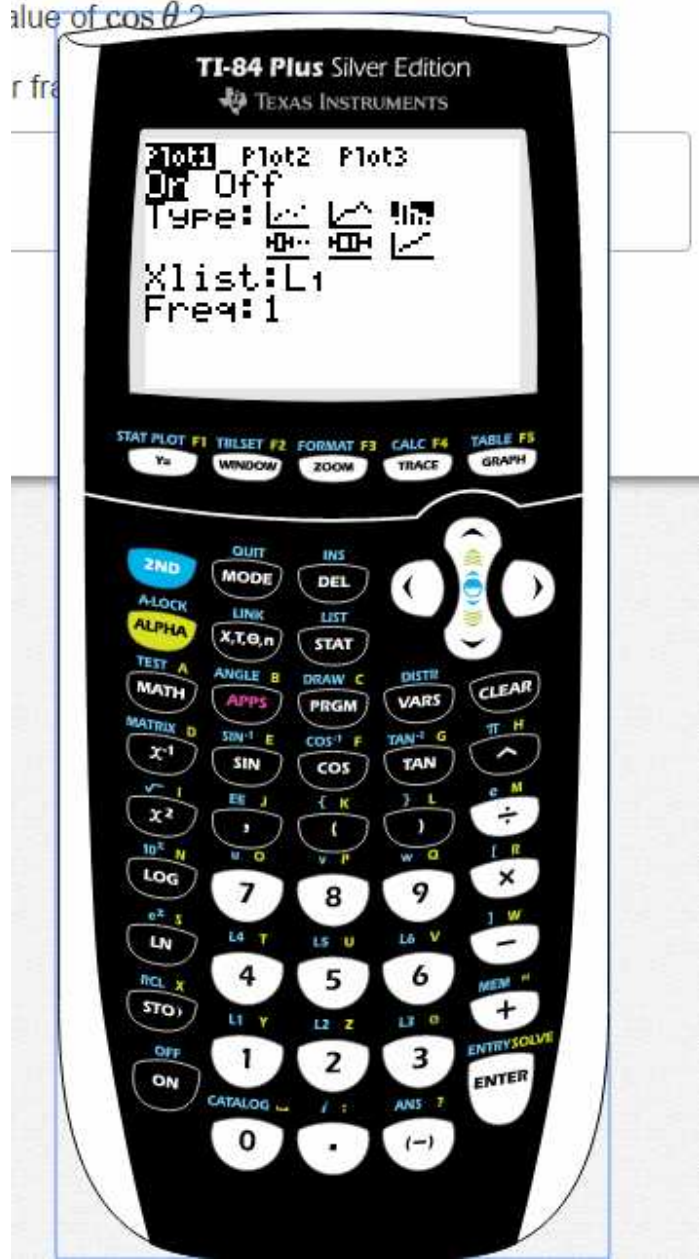
6. Choose Plot 1 (just hit enter).



Turn the plot on, choose histogram (third type), Xlist is in L1, Freq is 1 (no need to change).

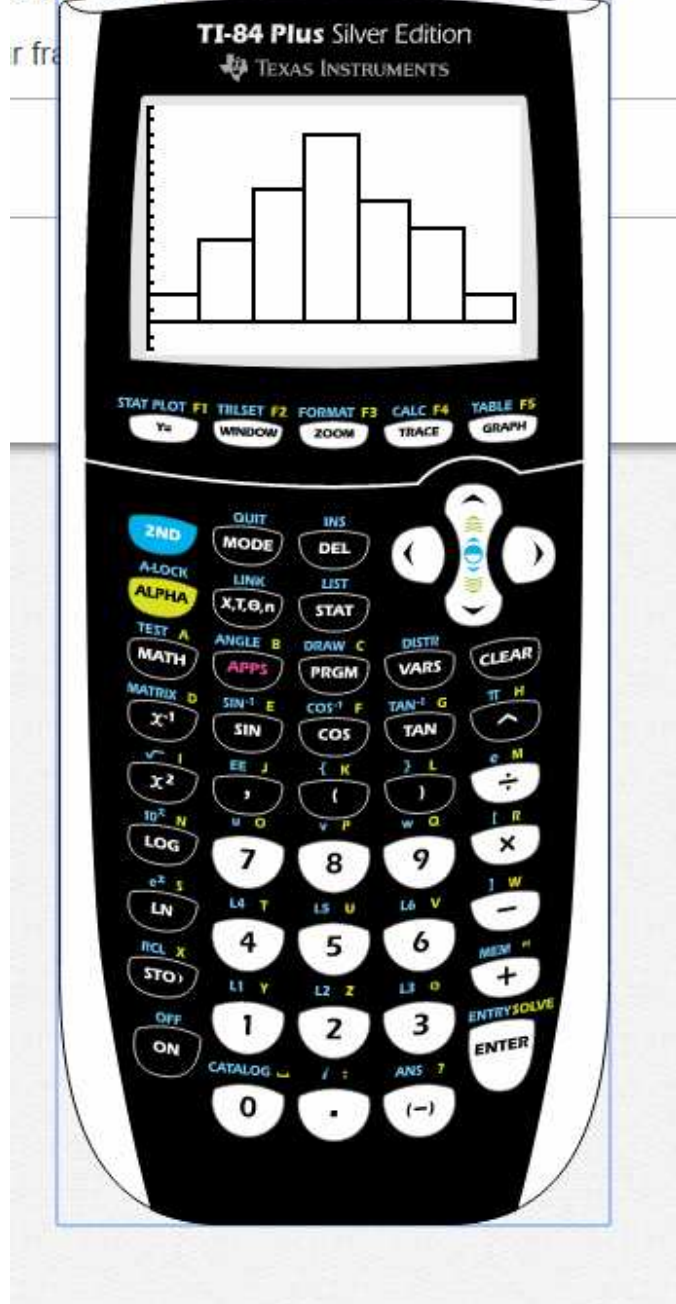
value of $\cos \theta$

r fr



7. Hit Graph.

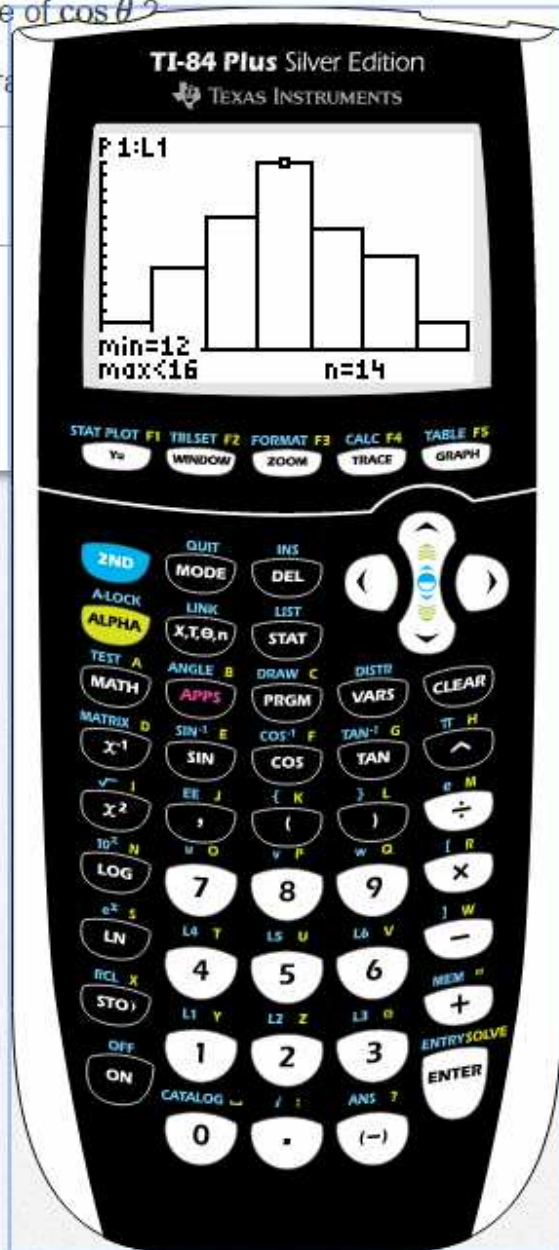
value of $\cos \theta$



Hit Trace and you can move through the classes using the gray arrow keys. Compare this graph to the one that you generated above.

value of $\cos \theta$

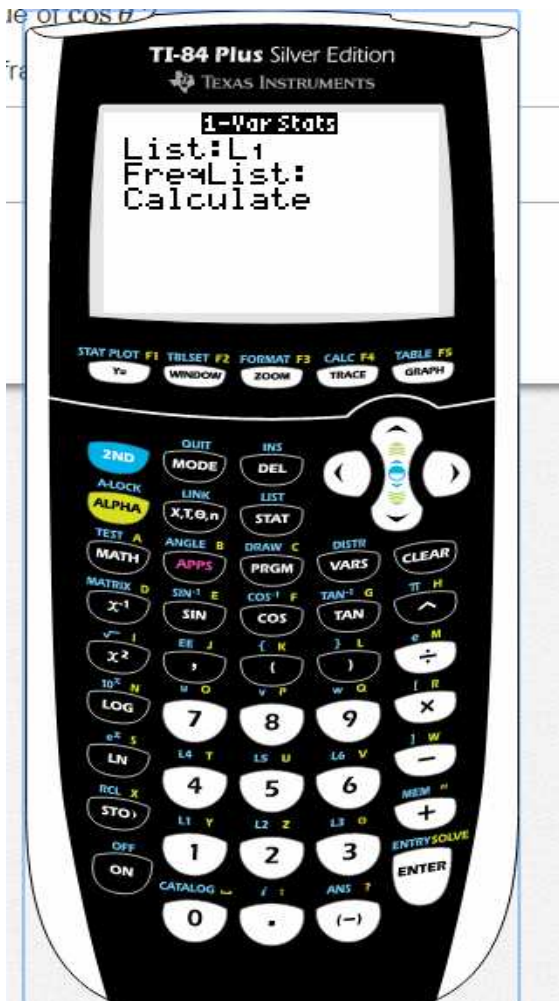
ur fra



You can see that I am looking at the fourth class. The value of n is the frequency and the X values for the class are on the left.

Note: You can perform One Variable Statistics on the data in L1 to obtain the summary statistics.

Stat – Calc – One Var Stat – Calculate



8. Clear the calculator once you are finished with this question.

2nd Function – Mem – Reset – All RAM – Reset – Clear

P16.

Consider the set {3, 11, 12, 19, 22, 23, 24, 25, 27, 29, 35, 36, 37, 45, 49}. Create a histogram. You may want to practice using the TI-84. You only have 15 observations. The square root of 15 is 3.87. However, this rule does not work well for small sets of data. We will choose 5 classes of width 10.

Class One: $0 \leq X < 10$




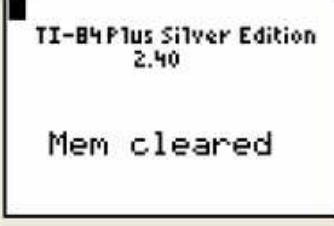
Class Five: $40 \leq X < 50$

Describe the distribution. Generate the summary statistics using One-Var Stats. What is the SD of the sample data? How can you interpret the SD if the data represent the ages of a random sample of people at a certain movie theater on a Sat afternoon?

End of Project

Using the TI-84 Plus

Source: http://www.ccti.edu/asc/Documents/AP_TI84.pdf

Basic Operations	
<p>Clearing the memory Sometimes, you may find it necessary to clear the memory on your calculator and reset all the defaults. <u>Warning:</u> doing this will wipe out any stored programs which you may have on your calculator.</p> <p>To clear the memory and reset defaults: press [2nd], <Mem>, {Reset}, {All}, {All Memory}, {Reset}</p> <p>You will receive a message informing you that the memory has been cleared.</p>	 <p>MEMORY 1:About 2:Mem Mgmt/Del... 3:Clear Entries 4:ClrAllLists 5:Archive 6:UnArchive 7:Reset...</p>  <p>RAM ARCHIVE [ALL] [1] All Memory...</p>  <p>RESET MEMORY 1:No 2:Reset Resetting ALL will delete all data, programs & Apps from RAM & Archive.</p>  <p>TI-84 Plus Silver Edition 2.40</p> <p>Mem cleared</p>

Powers and roots

The calculator has three keys which are used to calculate powers and roots.

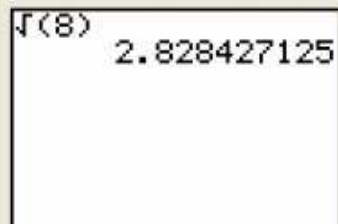
- $[x^2]$ is used to square a number
- $\langle\sqrt{}\rangle$ is used to determine the square root of a number (Note: the $\langle\sqrt{}\rangle$ automatically opens a set of parentheses.)
- $[^]$ is used to determine any power or root of a number

Examples:

- to determine 8 squared, press the following keys: 8, $[x^2]$, [Enter]
- to determine the square root of 8, press the following keys: $[2^{nd}]$, $\langle\sqrt{}\rangle$, 8, $[)]$, [Enter]
- to determine the cube root of 8, press the following keys: 8, $[^]$, $[()]$, 1, $[+]$, 3, $[)]$, [Enter]



8² 64



$\sqrt{}(8)$ 2.828427125



8^(1/3) 2

The following is very important because the adjustment of the window is needed for most graphs.

Adjusting the window for graphs

You may change the size, scale, and resolution of the graphing window in order to change what the calculator displays for a graph or boxplot.

The features you may change are:

- Xmin: the minimum value for x displayed
- Xmax: the maximum value for x displayed
- Xscl: the number of units between index marks on the x axis
- Ymin: the minimum value for y displayed
- Ymax: the maximum value for y displayed
- Yscl: the number of units between index marks on the y axis



```
WINDOW
Xmin=10
Xmax=10
Xscl=1
Ymin=-10
Ymax=10
Yscl=1
Xres=1
```


To change the window settings, press [Window]. You may now use the [V] (down arrow) key to move down the list and make adjustments. When you are finished, press [2nd], <Quit>.

Entering data into a list

Enter the data using the Stat editor. To enter data this way, press [Stat], {Edit}, the appropriate list, the value(s) to be stored, [V] (down arrow) or [Enter] after each entry. Press [2nd], <Quit> when complete.

- To enter the values 125, 150, 138, 180, and 205 into list 2: press [Stat], {Edit}, {L2}, 125 [V] 150 [V] 138 [V] 180 [V] 205 [V] [2nd], <Quit>

L1	L2	L3	2
64	125		
69	150		
71	138		
71	180		
73	205		


L2(6) =			


Box plots

The box plot feature plots one-variable data. The values shown are the minimum (left whisker), Q1 (left side of box), median (line inside box), Q3 (right side of box), and maximum (right whisker).

More than one box plot may be viewed simultaneously. When this capability is used, the box plots are plotted with the first box plot at the top of the screen, second in the middle, and third at the bottom.

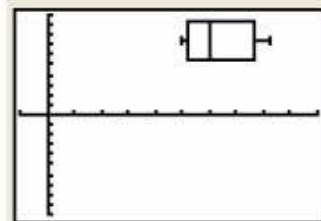
In order to view the box plots, you may need to adjust the Xmin and Xmax for the viewing window.

To create a box plot, press: [2nd], <Stat Plot>, plot to use, {On}, {}, location of the data for Xlist, [Graph].

Example: To create a box plot using the data in L₂ as plot 1, press: [2nd], <Stat Plot>, {Plot 1}, {On}, {}, [2nd], <L₂>, [Graph] (Note: for this example, the window is set for Xmin: -25, Xmax: 250, Xscl: 25. For information about adjusting the window, see page 6.)


```
STAT PLOTS
1: Plot1...On
  [2nd] L2 1
2: Plot2...Off
  [2nd] L1 L2
3: Plot3...Off
  [2nd] L1 L2
4↓ PlotsOff
```


```
Plot1 Plot2 Plot3
Off Off
Type: [2nd] [2nd] [2nd]
  [2nd] [2nd] [2nd]
Xlist:L2
Freq:0
```



Modified box plots

The modified box plot is similar to the box plot. The difference is that the modified box plot individually plots outliers (data which is 1.5 times the interquartile range beyond the quartiles).

To create a modified box plot, press: [2nd], <Stat Plot>, plot to use, {On}, {}, location of the data for Xlist, the mark to use for outliers, [Graph].

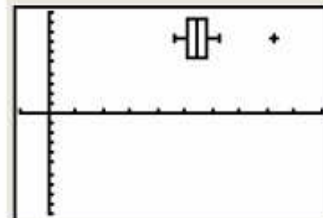
Example: To create a modified box plot using the data in L₃ as plot 2 with "+" marking the outlier(s), press: [2nd], <Stat Plot>, {Plot 2}, {On}, {}, [2nd], <L₃>, {+}, [Graph] (Note: for this example, the window is set for Xmin: -25, Xmax: 250, Xscl: 25. For information about adjusting the window, see page 6.)



```
STAT PLOTS
1:Plot1...Off
  Xlist: L2  1
2:Plot2...On
  Xlist: L3  1  +
3:Plot3...Off
  Xlist: L1  L2  □
4↓PlotsOff
```



```
Plot1 On Plot3
Off Off
Type:   
Xlist: L3
Freq: 1
Mark:   
```



Scatter plots

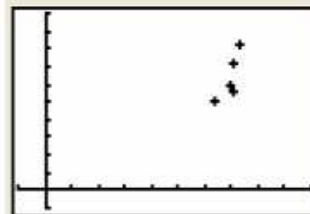
Scatter plots plot the data from two lists as coordinate pairs.

To create a scatter plot, press: [2nd], <Stat Plot>, plot to use, {On}, { \llcorner }, location of the data for Xlist, location of the data for Ylist, the mark to use for the coordinate pairs, [Graph].

Example: To create a scatter plot using the data in L₁ for X and L₂ for Y as Plot 1 with "+" marking the coordinate pairs, press: [2nd] <Stat Plot>, {Plot 1}, {On}, { \llcorner }, [2nd], <L₁>, [2nd], <L₂>, {+}, [Graph]. (Note: for this example, the window is set for Xmin: -10, Xmax: 100, Xscl: 10, Ymin: -25, Ymax: 250, Yscl: 25. For information about adjusting the window, see page 6.)

```
STAT PLOTS
1:Plot1...On
  L1 L2 +
2:Plot2...Off
  L1 L2 +
3:Plot3...Off
  L1 L2 □
4↓PlotsOff
```

```
Plot1 Plot2 Plot3
On Off
Type:  $\llcorner$   $\llcorner$   $\llcorner$ 
       $\llcorner$   $\llcorner$   $\llcorner$ 
Xlist:L1
Ylist:L2
Mark: □ ■ .
```



xyLine

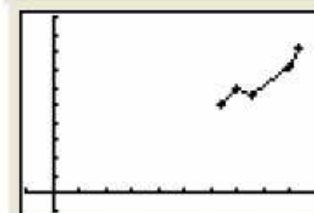
xyLine is a scatter plot in which the data points are plotted and connected in the order they appear on the lists. If the data is not already sequenced, you may need to use the SortA feature of the calculator (see page 25.)

To create an xyline, press: [2nd], <Stat Plot>, plot to use, {On}, { \llcorner }, location of the data for Xlist, location of the data for Ylist, the mark to use for the coordinate pairs.

Example: To create an xyline using the data in L₁ for X and L₂ for Y as Plot 1 with "+" marking the coordinate pairs, press: [2nd] <Stat Plot>, {Plot 1}, {On}, { \llcorner }, [2nd], <L₁>, [2nd], <L₂>, {+}, [Graph]. (Note: for this example, the window is set for Xmin: -10, Xmax: 100, Xscl: 10, Ymin: -25, Ymax: 250, Yscl: 25. For information about adjusting the window, see page 6.)

```
STAT PLOTS
1:Plot1...On
  L1 L2 +
2:Plot2...Off
  L1 L2 +
3:Plot3...Off
  L1 L2 □
4↓PlotsOff
```


```
Plot1 Plot2 Plot3
On Off
Type:  $\llcorner$   $\llcorner$   $\llcorner$ 
       $\llcorner$   $\llcorner$   $\llcorner$ 
Xlist:L1
Ylist:L2
Mark: □ ■ .
```




You have to set the window before you can create a histogram.

Histograms

The histogram plots one-variable data. The Xscl value (located on the window screen) determines the width of each bar of the histogram.

To create a histogram, press: [2nd], <Stat Plot>, plot to use, {On}, {}, location of the data for Xlist, [Graph].

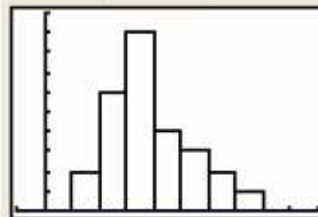
Example: To create a histogram using the data in L₄ as Plot 1, press: [2nd], <Stat Plot>, {Plot 1}, {On}, {}, [2nd], <L₄>, [Graph].



```
STAT PLOTS
1:Plot1...On
  ▬ L1 1
2:Plot2...Off
  ▬ L1 L2 +
3:Plot3...Off
  ▬ L1 L2 □
4↓PlotsOff
```



```
Plot1 Plot2 Plot3
On Off
Type: ▬ ▬ ▬
      ▬ ▬ ▬
Xlist:L4
Freq:1
```



Normal Probability Plots

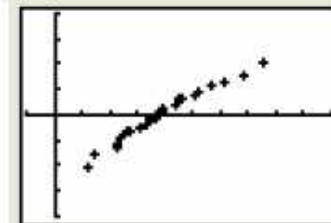
The normal probability plot plots each observation X in a list versus the corresponding quartile of the standard normal distribution. If the plotted points appear to lie close to a straight line, then the plot indicates that the data is normal.

To create a normal probability plot, press: [2nd], <Stat Plot>, plot to use, {On}, { \surd }, location of data for Data List, {X}, the mark to use to plot data, [Graph].

Example: To create a normal probability plot using the data in L_4 as Plot 1, press: [2nd], <Stat Plot>, {Plot 1}, {On}, { \surd }, [2nd], <L4>, {X}, {+}, [Graph]. (Note: for this example, the window is set for Xmin: -10, Xmax: 100, Xscl: 10, Ymin: -4, Ymax: 4, Yscl: 1. For information about adjusting the window, see page 6.)

```
STAT PLOTS
1:Plot1...On
  L4 xAxis +
2:Plot2...Off
  L1 L2 +
3:Plot3...Off
  L1 L2 □
4↓PlotsOff
```

```
Plot2 Plot3
Off Off
Type: L1 L2 L3
      D1 D2 D3
Data List:L4
Data Axis:Y
Mark: □ ■ .
```



1 Variable Statistics

The 1-variable statistics function analyzes data with one variable. This function may be executed two ways.

1. From a single list. This procedure is used if all of the data values are entered into a single list, with values being repeated if the value occurs more than once. Press: [Stat], {Calc}, {1-Var Stats}, location of data, [Enter].
2. From two lists. This procedure is used if one list contains the values and a second list contains the frequency of occurrence for each value. Press: [Stat], {Calc}, {1-Var Stats}, location of values, [,], location of frequencies, [Enter].

Examples:

- To determine the 1 variable statistics for the data in L_3 , press: [Stat], {Calc}, {1-Var Stats}, [2nd], <L3>, [Enter].

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

```
1-Var Stats L3
```

```
1-Var Stats
x̄=22.30769231
Σx=290
Σx²=7026
Sx=6.811566332
σx=6.544341248
↓n=13
```

- To determine the 1 variable statistics for the data in L_3 , with L_5 being the frequency of each occurrence, press: [Stat], {Calc}, {1-Var Stats}, [2nd], < L_3 >, [,], [2nd], < L_5 >, [Enter].

```

EDIT  [2nd] CALC TESTS
1:1-Var Stats
2:2-Var Stats
3:Med-Med
4:LinReg(ax+b)
5:QuadReg
6:CubicReg
7↓QuartReg

```

```

1-Var Stats L3,L
5

```

```

1-Var Stats
x=21.93548387
Σx=680
Σx²=15800
Sx=5.427924612
σx=5.339659793
↓n=31

```

Other

Sorting lists

Lists may be sorted in either ascending or descending order.

- To sort a list in ascending order, press: [Stat], {Edit}, {SortA()}, list name, [], [Enter].
- To sort a list in descending order, press: [Stat], {edit}, {SortD()}, list name, [], [Enter].

Example: to sort L_1 in ascending order, press: [Stat], {Edit}, {SortA()}, [2nd], < L_1 >, [], [Enter].

```

[2nd] [2nd] CALC TESTS
1:Edit
2:SortA(
3:SortD(
4:ClrList
5:SetUpEditor

```

```

SortA(L1) Done

```

Here are some additional and related notes from a different source.

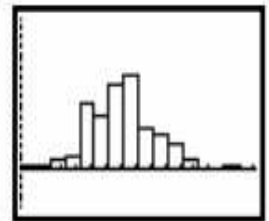
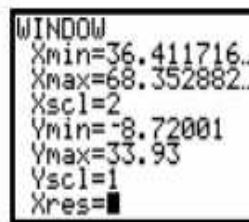
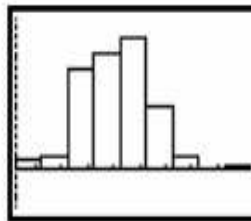
Source:

http://www.medmood.com/aosr/pluginfile.php/30340/mod_resource/content/1/calculator%20use.pdf

A. Graphs

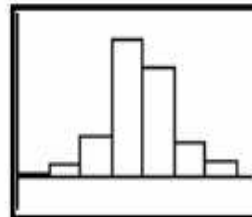
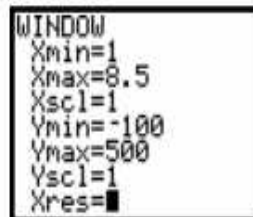
* 1. Histograms - **2nd** **Statplot** Select Plot

- data must be in a selected list. You should clear out anything in Y1 and use **ZOOM** **9**.
- you may adjust the window (Xscl) to change the bin width.



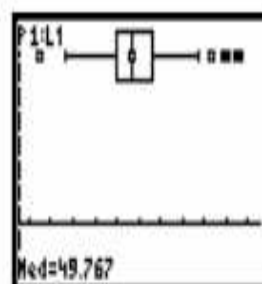
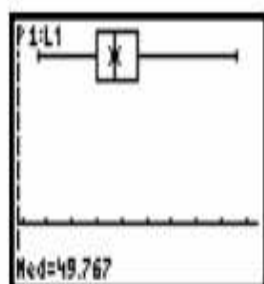
- If you have summary statistics, your methods are somewhat different. Example: Students were asked how many people live in their house. The summary of that data is in L1 (the number of people in the house) and L2 (how many students have that number).

L1	L2	L3	3
1,000	8,000		
2,000	34,000		
3,000	125,000		
4,000	427,000		
5,000	328,000		
6,000	108,000		
7,000	49,000		
L3(1)=			



* 2. **Boxplots** - **2nd** **Statplot** Select Plot

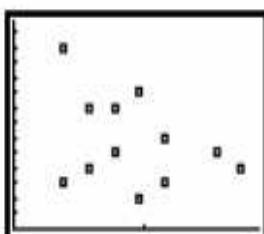
- data must be in a selected list. You should clear out anything in Y1 and use **ZOOM** **9** .
- modified boxplots (showing outliers) can be generated using the 3rd screen below:



- You may scroll through the 5 number summary by pressing **TRACE** . You may show up to 3 boxplots at one time.

* 3. **Scatterplots** - **2nd** **Statplot** Select Plot

- bi-variate data should be in two lists. You should clear out anything in Y1 and use **ZOOM** **9** . Define the mark at the bottom. If you have many points, the dot is preferable. You can scroll through the points by pressing **TRACE** .



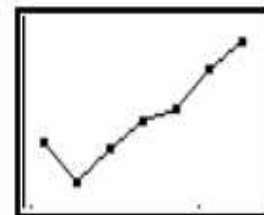
4. Trendlines - $\boxed{2nd}$ $\boxed{Statplot}$ Select Plot

- When data is represented to show change over time, it is best to use a trendline. Bi-variate data should be in two lists. You should clear out anything in Y1 and use \boxed{ZOOM} $\boxed{9}$. Define the mark at the bottom. If you have many points, the dot is preferable. You can scroll through the points by pressing \boxed{TRACE} .

L1	L2	L3	3
1994	1225		
1995	1185		
1996	1220		
1997	1245		
1998	1258		
1999	1296		
2000	1325		
L3(1)=			

```
Plot1 Plot2 Plot3
Off Off Off
Type: L1 L2 L3
Xlist: L1
Ylist: L2
Mark: [dot] [square]
```

```
WINDOW
Xmin=1993.4
Xmax=2000.6
Xscl=5.1249192...
Ymin=1161.2
Ymax=1348.8
Yscl=1
Xres=
```



C. Calculation of Statistics - STAT CALC

* **1. 1-Var Stats** - used to find important statistical information with 1 variable

- assumes that the list that the data you want information about is in L1. If not, you must specify it (ex. 1-Var Stats L3)

- The following information is given:

\bar{x} - the numerical mean (average)

$\sum x$ - the sum of your data

$\sum x^2$ - the sum of the squares of your data

s_x - the sample standard deviation of your data: $= \frac{\sum (x_i - \bar{x})^2}{n-1}$

σ_x - the population standard deviation of your data: $= \frac{\sum (x_i - \bar{x})^2}{n}$

n - the number of data you have in your list

minX - the minimum value in your list

Q1 - the first quartile - 25% of the data is below this value

Med - the median of your data - 50% of the data is both below and above this value

Q3 - the third quartile - 25% of your date is above this value

maxX - the maximum value in your list