(A) Lesson Context

BIG PICTURE of this UNIT:	• How do we analyze and then make conclusions from a data set? (Math) • How do I present my data and the outcomes of my analysis? (Math) • How do I use data & statistics to make decisions? • How do I decide on the validity/reliability of my data? Of my analysis? Of my conclusions? Of my decision?
CONTEXT of this LESSON:	**Where we've been** From your MS math experience, you may have had an introductory unit on Statistics	**Where we are** Using data & visual representations, present your current understandings of what Statistics is	**Where we are heading** How do I analyze and make conclusions from a data set, in whatever way this data gets presented?

(B) Lesson Objectives:

a. Record ideas and questions that students currently have about Statistics
b. Use the graphing calculator to determine the regression equations of the data sets
c. Introduce key features of the graphs of quadratic relations (the graphs are called parabolas)

## (C) Activity #1 – Current Thinking & Puzzles ➜ Ideas & Questions – POSTER

Part 1: In this section of the lesson your group will be responsible for writing, drawing, and discussing all of the things you collectively know and understand about the topics of STATISTICS. These are the only rules.

- Everyone must write as much as they can.
- Each team member needs a different color marker.
- Everyone writes at the same time.
- The only time you should stop writing is to talk with your group members about Statistics.
- NO TECHNOLOGY: Only your brain for this… if you are a robot… turn off your wireless… no internet.
- You should work on this poster for 10 – 15 minutes.
- Write: Words, Definitions, and thoughts/Questions.
- Draw: Diagrams, Graphs, Examples, and thoughts/Questions.

Part 2: Sticky Notes: Each member will get two sticky notes. Once all the posters are done, each group member needs to put two sticky notes down with LEGITIMATE questions they have about things on other peoples poster. We will have 8 minutes for sticky note questions.

(D)Activity #2 – Ideas & Questions ➜ From a GRAPH

a. You will be given 5 graphs, showing the mark distribution of our current classes of IM2.

b. From your group's thinking about the graphs, you will record:
1. Initial thoughts that arise from the graphs and the context
2. Initial conclusions that arise from the graphs and the context
3. Initial questions that arise from the graphs and the context

c. Rotate one spokesman to another group and share thoughts, conclusions, questions. Use post-it notes to add new ideas
d. Consolidate as a class on ideas

CLASS 1
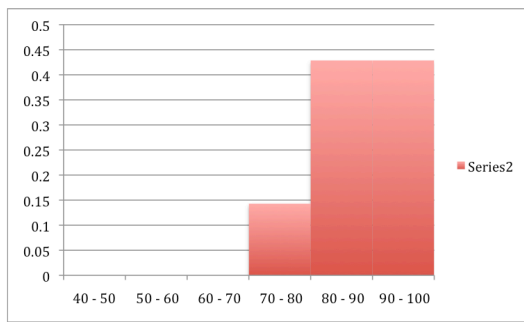
CLASS 2
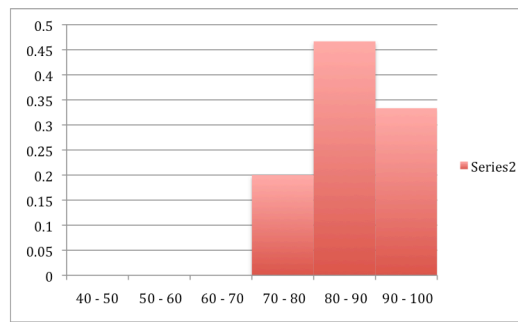
CLASS 3
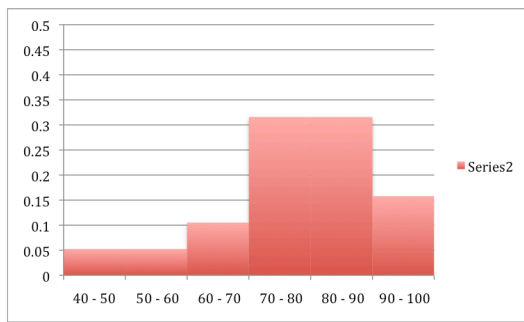
CLASS 4
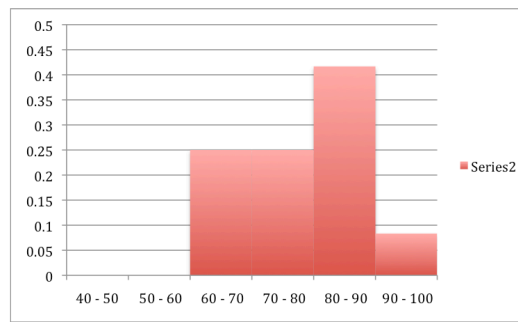
CLASS 5
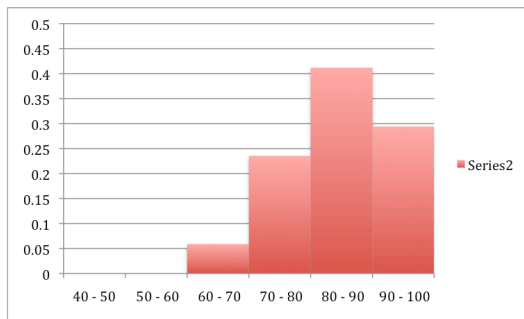
(E) Activity #3 – Ideas & Questions ➜ From a Data Set

a. You will be given 3 data sets of sprinters and be presented with the task of making a decision as to which two sprinters deserve to be selected for the CAC ISST team. The event in question is the 400m sprint.
b. From your group's thinking about the graphs, you will record:
  1. Initial thoughts that arise from the data sets and the context
  2. Initial conclusions that arise from the data sets and the context
  3. Initial questions that arise from the data sets and the context
c. Rotate one spokesman to another group and share thoughts, conclusions, questions. Use post-it notes to add new ideas
d. Consolidate as a class on ideas

Sprinter #1	57.54	55.23	59.32	58.42	58.21	56.37	57.41	56.10
	53.11	55.42	57.31	58.46	57.19	55.16	54.12	56.25
Observations			Conclusions			Questions		

Sprinter #2	54.26	55.71	54.70	53.89	56.25	54.68	53.19	53.82
	57.29	55.54	54.96	53.87	55.21	53.32	54.10	53.61
Observations			Conclusions			Questions		

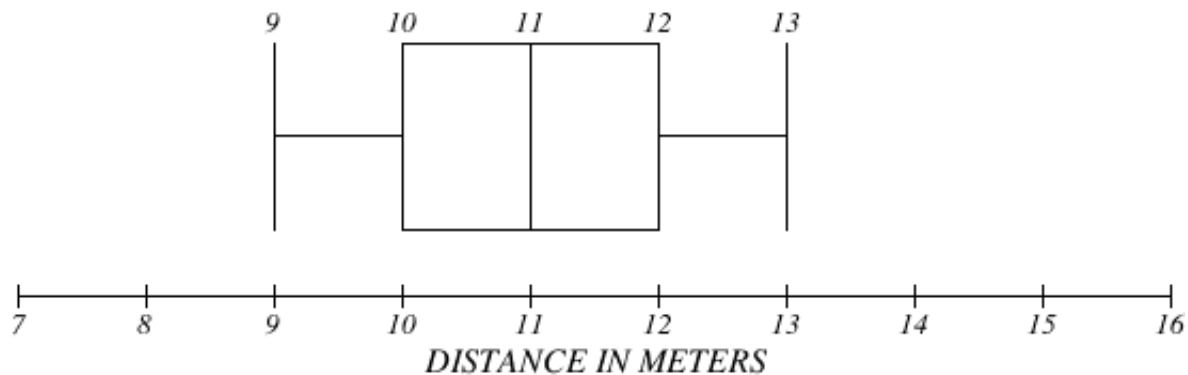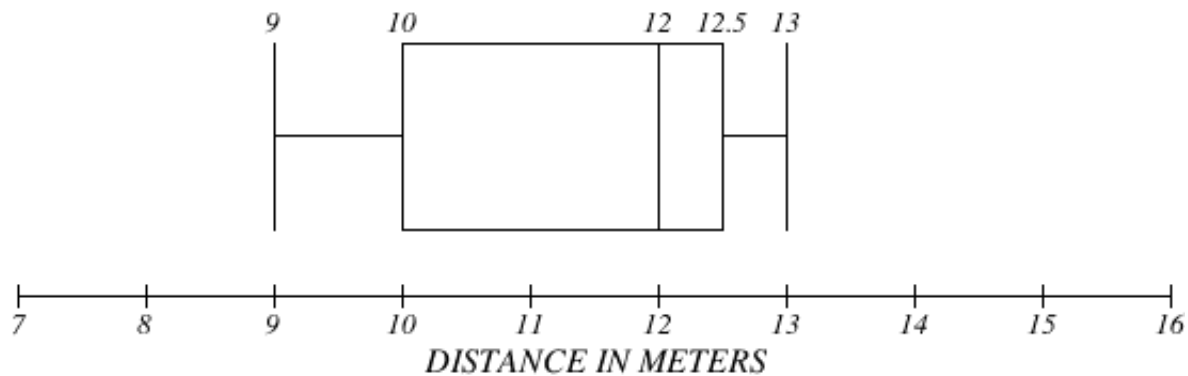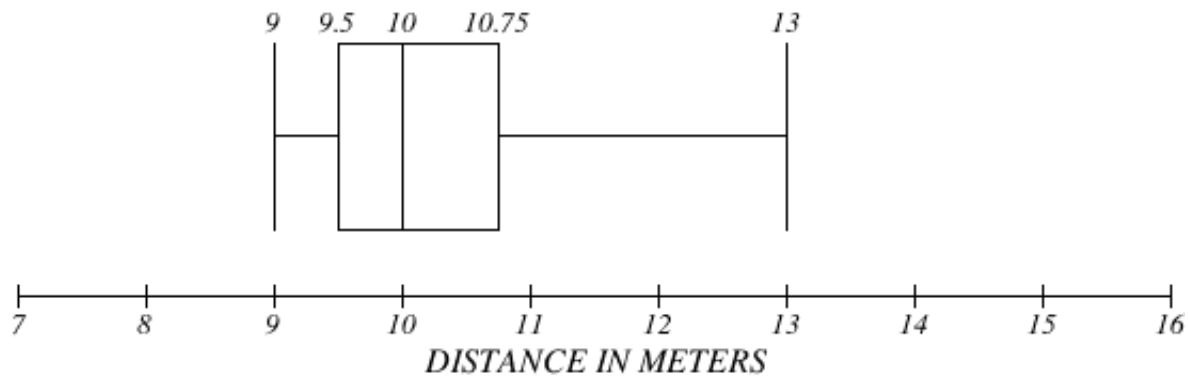| Sprinter #3	55.49	53.15	54.64	55.91	57.80	54.48	53.93	55.12
	57.91	58.13	52.95	52.62	53.10	54.54	58.04	56.81
Observations			Conclusions			Questions		

(F) Acticity #4 – Who's the Best???

Here are 6 graphs showing you run times for 100m sprinters, competing for being selected for the CAC team going to the ISST T&F meet in Brussels.
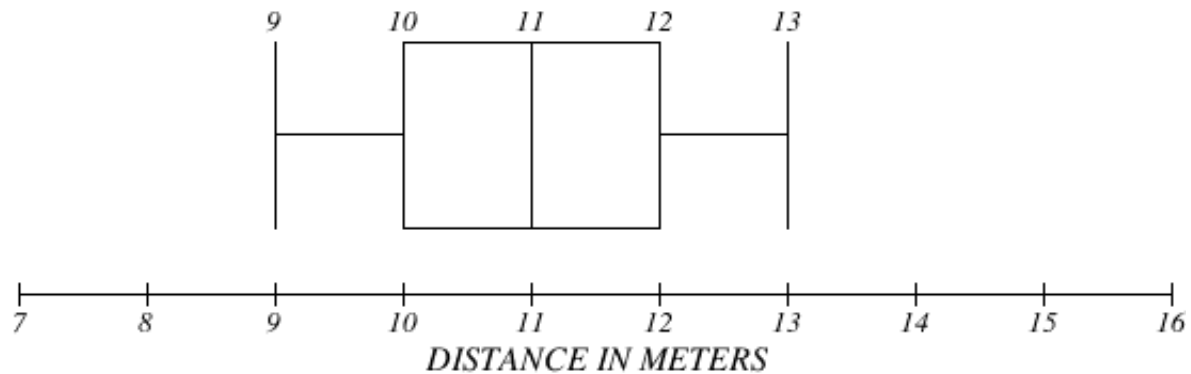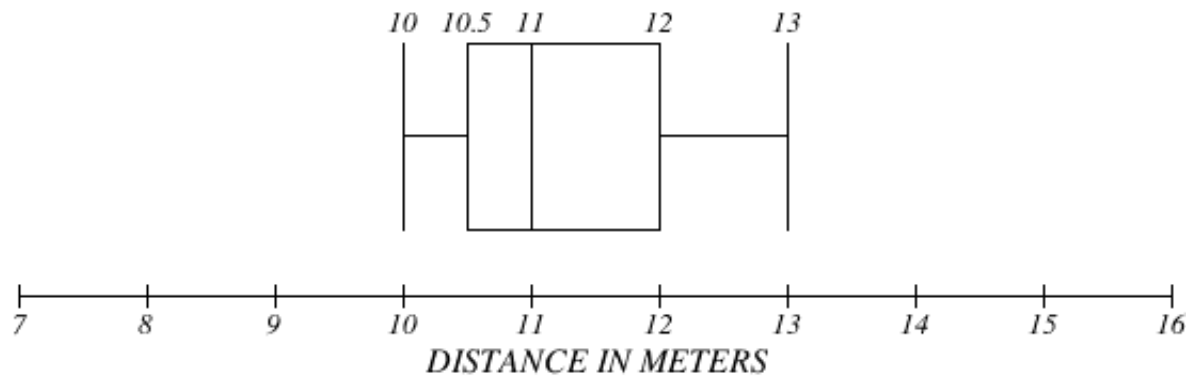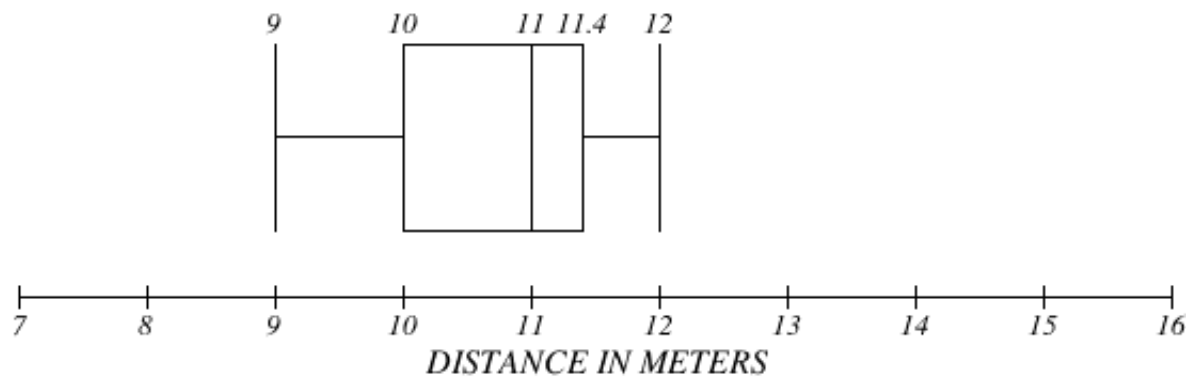
   a. From your group's thinking about the graphs, you will record:
       1. Initial thoughts that arise from the data sets and the context
       2. Initial conclusions that arise from the data sets and the context
       3. Initial questions that arise from the data sets and the context
   b. Rotate one spokesman to another group and share thoughts, conclusions, questions. Use post-it notes to add new ideas
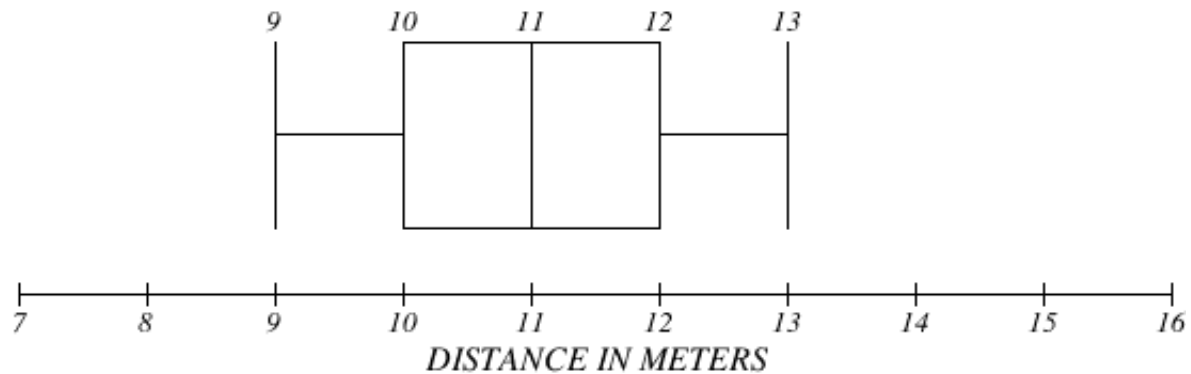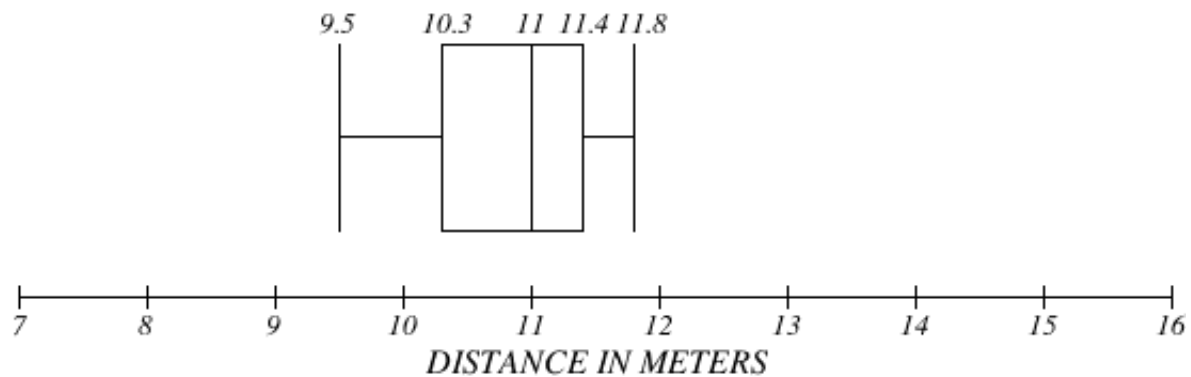   c. Consolidate as a class on ideas

A.

B.

C.

D.

E.

F.

COMPARISONS:
http://www.imathas.com/stattools/boxplot.html

Thrower #1

9   10   11   12   13

7   8   9   10   11   12   13   14   15   16
DISTANCE IN METERS

Thrower #2

9   10   12 12.5 13

7   8   9   10   11   12   13   14   15   16
DISTANCE IN METERS

Thrower #3

9 9.5 10 10.75   13

7   8   9   10   11   12   13   14   15   16
DISTANCE IN METERS

SET #2

Thrower #1

9   10   11   12   13

7   8   9   10   11   12   13   14   15   16

DISTANCE IN METERS

Thrower #2

10 10.5 11   12   13

7   8   9   10   11   12   13   14   15   16

DISTANCE IN METERS

Thrower #3

9   10   11 11.4   12

7   8   9   10   11   12   13   14   15   16

DISTANCE IN METERS

SET #3

Thrower #1

```
          9           10          11          12          13
```

DISTANCE IN METERS
7   8   9   10   11   12   13   14   15   16

Thrower #2

```
          9.5        10.3     11  11.4 11.8
```

DISTANCE IN METERS
7   8   9   10   11   12   13   14   15   16

Thrower #3

```
          8           10          11       11.8 12.2
```

DISTANCE IN METERS
7   8   9   10   11   12   13   14   15   16

Set #1 ➜ http://www.socscistatistics.com/descriptive/bar/Default.aspx

Thrower #1

Thrower #2

Thrower #3

Thrower #4
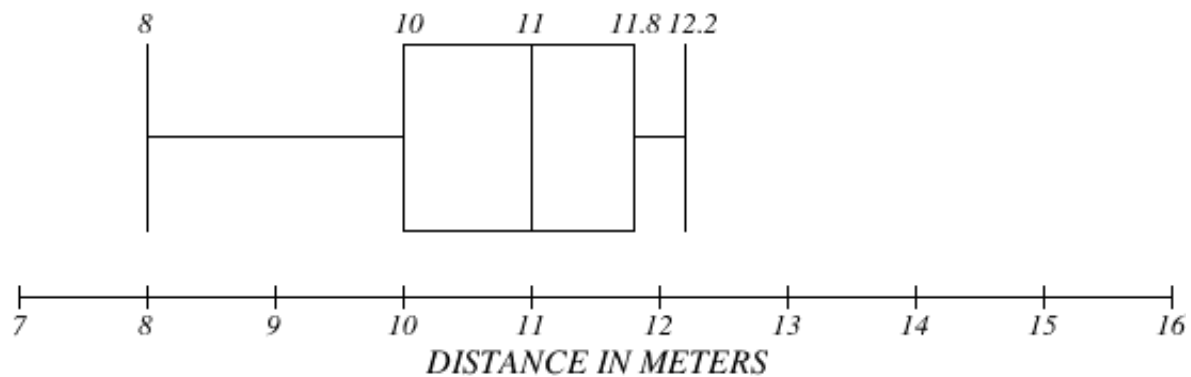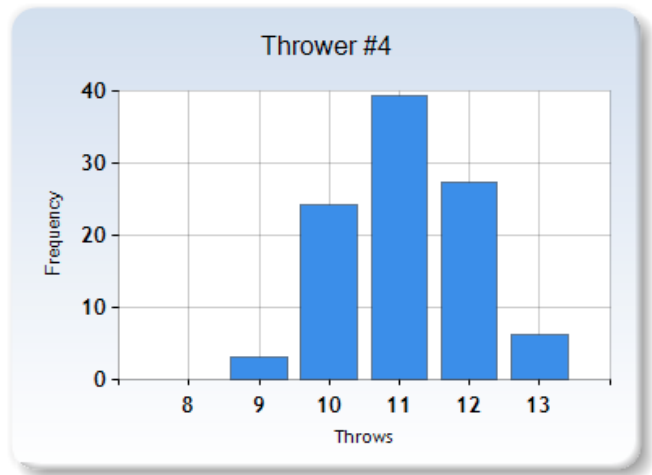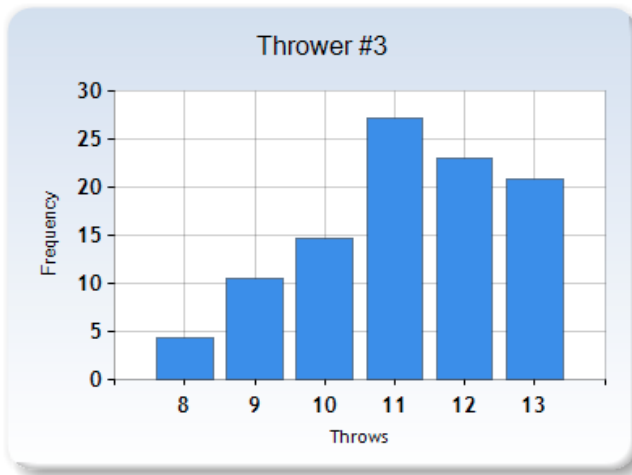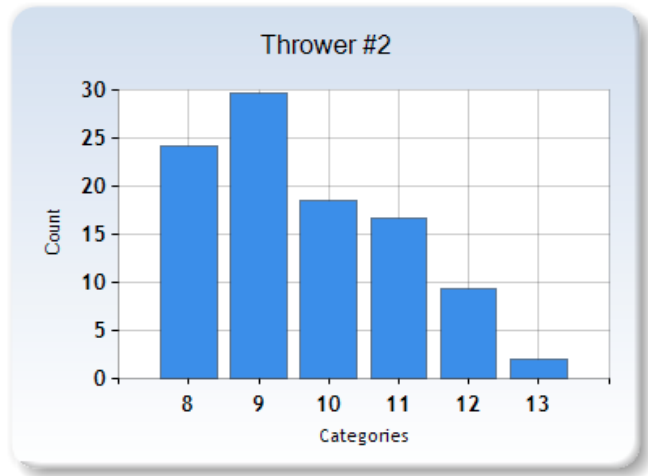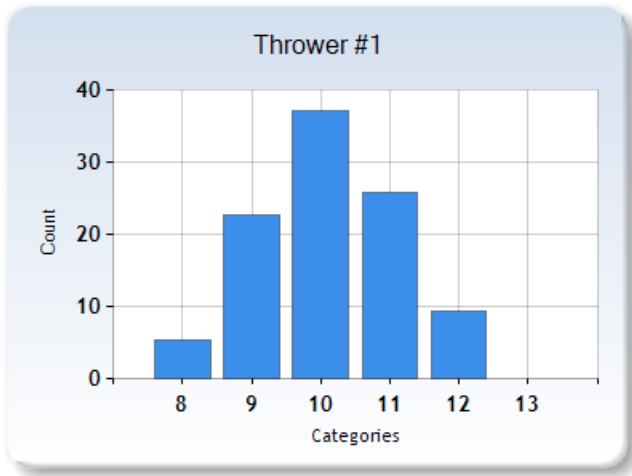
Comparing Sprinters

(A) Lesson Context

BIG PICTURE of this UNIT:	• How do we analyze and then make conclusions from a data set? (Math) • How do I present my data and the outcomes of my analysis? (Math) • How do I use data & statistics to make decisions? • How do I decide on the validity/reliability of my data? Of my analysis? Of my conclusions? Of my decision?

	Where we've been	Where we are	Where we are heading
CONTEXT of this LESSON:	Using data & visual representations, present your current understandings of what Statistics is	How do we visually represent data?	How do I analyze and make conclusions from a data set, in whatever way this data gets presented?

(B) Lesson Objectives:

a. Describe data as either being qualitative (categorical) or quantitative (numerical)
b. Introduce the idea that quantitative data can be either discrete or continuous
c. Introduce different ways to organize & visually represent data ➔ using dot plots, stem & leaf plots, frequency tables, grouped data, histograms/column graphs

(C) Describing Data (go online to find definitions/clarifications/examples for these terms)

a. Qualitative Data ➔

b. Quantitative Data ➔

1. Discrete quantitative data ➔

2. Continuous quantitative data ➔

(D) Activity #1 – Organizing and Representing Data

Mr Santowski has two athletes competing in a 400m sprint and he needs to collect data in order to help the coaching staff decide which sprinter is the better of the two. Over the course of the training season, the two sprinters have had numerous time trial data collected and the results are listed below

Sprinter #1	57.54	55.23	59.32	58.42	58.21	56.37	57.41	56.10
	53.11	55.42	57.31	58.46	57.19	55.16	54.12	56.25
	54.26	55.71	54.70	53.89	56.25	54.68	53.19	53.82
	57.29	55.54	54.96	53.87	55.21	53.32	54.10	53.61
	55.49	53.15	54.64	55.91	57.80	54.48	53.93	55.12
	57.91	58.13	52.95	52.62	53.10	54.54	58.04	56.81

Sprinter #2	57.67	56.29	58.01	54.28	54.63	57.95	56.21	55.04
	55.45	57.96	53.82	57.31	56.19	54.67	56.91	53.43
	55.17	54.44	56.28	57.34	54.99	57.22	56.23	56.31
	53.52	55.20	54.83	54.96	57.03	54.05	54.27	53.61
	55.26	57.36	57.38	54.93	53.82	53.01	54.82	56.67
	55.93	57.21	57.58	56.91	55.33	58.11	54.96	54.60

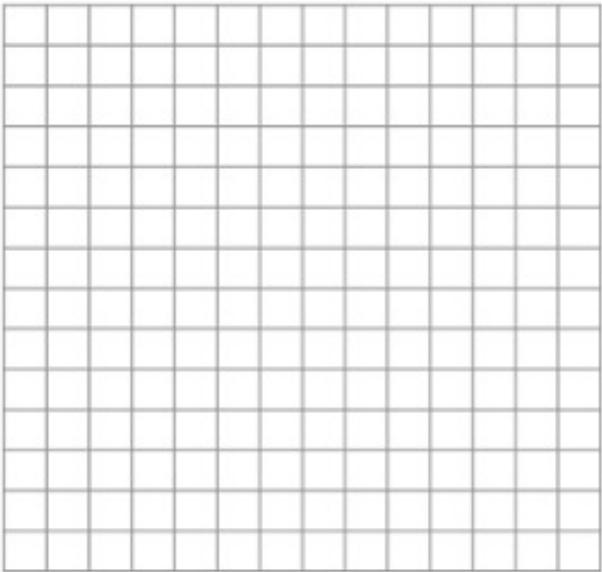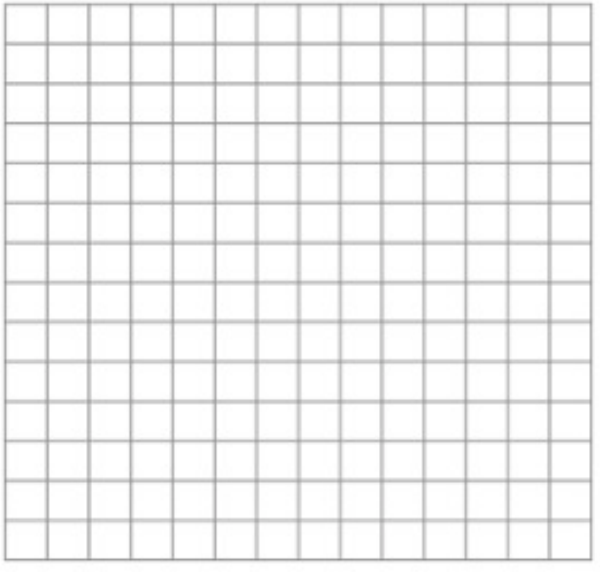### *To help with the data analysis from each experiment, your group will:*

(a) Prepare a Stem and Leaf plot for each data set. (Here are three links showing you how to prepare a stem and leaf plot) ➔ Link #1 from Math is Fun; Link #2 from PurpleMath and Link #3 from our Haese and Harris reference text on page 113.

SPRINTER #1		SPRINTER #2	
Stem	Leaf	Stem	Leaf

(b) Prepare a Frequency table for each data set. (Here are two links showing you how to prepare a frequency table) ➔ Link #1 from MathsTeacher.com and Link #2 from our Haese and Harris reference text on page 113

SPRINTER #1			SPRINTER #2		
Interval	Tally	Frequency	Interval	Tally	Frequency

(c) Prepare a Column graph by hand for each data set. (Here are three links showing you how to prepare a frequency histogram from a frequency table) ➔ Link #1 from Math is Fun; Link #2 from Open High School Courses; and Link #3 from our Haese and Harris reference text on page 114.

SPRINTER #1	SPRINTER #2

(d) Now you need to make a decision ➔ who's the BEST and why?

(e) Now you have an opportunity to ASK QUESTIONS about the data, about the data collection, about the conditions, etc

(f) Prepare a Column graph on your graphing calculator for each data set (instructions linked here.)

(A) Lesson Context

BIG PICTURE of this UNIT:	• How do we analyze and then make conclusions from a data set? (Math) • How do I present my data and the outcomes of my analysis? (Math) • How do I use data & statistics to make decisions? • How do I decide on the validity/reliability of my data? Of my analysis? Of my conclusions? Of my decision?

	Where we've been	Where we are	Where we are heading
CONTEXT of this LESSON:	Using data & visual representations, present your current understandings of what Statistics is	How do we prepare and analyze frequency histograms, frequency polygons and cumulative frequency?	How do I analyze and make conclusions from a data set, in whatever way this data gets presented?

(B) Lesson Objectives:

a. Starting from a set of raw data, prepare a grouped frequency table – using absolute and relative frequencies
b. Use either the grouped frequency tables or the histograms to prepare frequency polygons
c. Use the grouped frequency tables, calculate the cumulative frequencies and prepare cumulative frequency graphs (or ogives)
d. Introduce simple analysis questions, which can be answered from any of these visual representations, most of which involve percentiles.

(C) Misleading Graphs

You are going to be presented with 4 slides showing statistical information via graphs. You are asked to record your ideas about "What's wrong with this Picture"

Slide #1	Slide #2
Slide #3	Slide #4

(D)Opening Exercise/Review ➔ Below is a frequency table for 3 different classrooms' test scores.

Score	Frequency		
	Class A	Class B	Class C
0-9	0	0	0
10-19	1	0	1
20-29	0	0	5
30-39	0	0	7
40-49	0	1	2
50-59	1	1	2
60-69	4	2	1
70-79	9	3	1
80-89	3	6	1
90-100	1	4	0
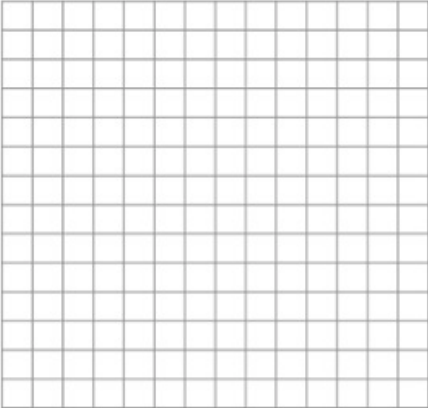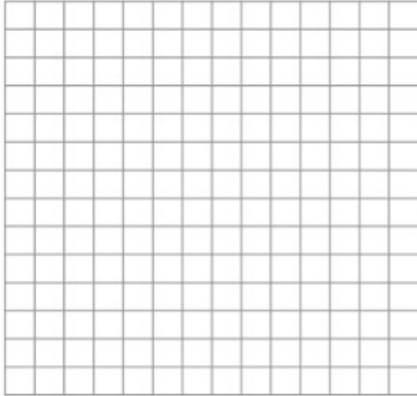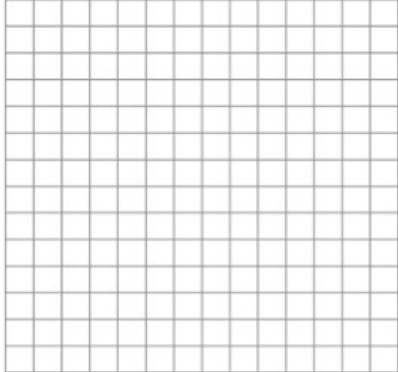
a. How many students are in each class? How do you know?

b. Please create three sets of data that will match the given frequency tables.

Class A

Class B

Class C

c. Create Frequency Bar Graphs for each class.

Class A	Class B	Class C
1 Initial Observation for each class.		
2 Conclusions for each class.		
3 Questions for each class about the data, or data collection.		
Circle one for each class that best describes the data distribution.		
Class A	Class B	Class C
Positively Skewed?	Positively Skewed?	Positively Skewed?
Negatively Skewed?	Negatively Skewed?	Negatively Skewed?
Symmetrical Data?	Symmetrical Data?	Symmetrical Data?
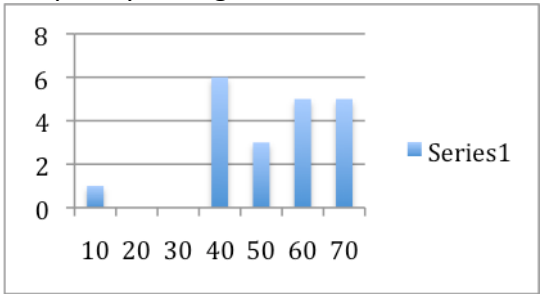
Do any of the data points stand out to you?
   (A) Outliers
     a. Go online and look up the term "Outlier" Please give a definition of this term in your own words, and then given examples of this term visually, as well as in a set of data.

Outlier	
Definition:	
Visual Example	Example with a data set

   b. Given the data representation below, fill out the rest of the table.

Frequency Chart	Frequency Histogram
X Y	
Actual Data: Based off of the Histogram	Stem and Leaf Chart
Outliers:	Average:

Story Behind the Data?

Story Behind the Outlier?

Questions you may have about this problems?

(E) Frequency Distribution Tables ➔ Example #1

Prepare and use frequency distribution tables (using both frequencies & relative frequencies) to create (i) histographs, (ii) frequency polygons and (iii) cumulative frequency graphs

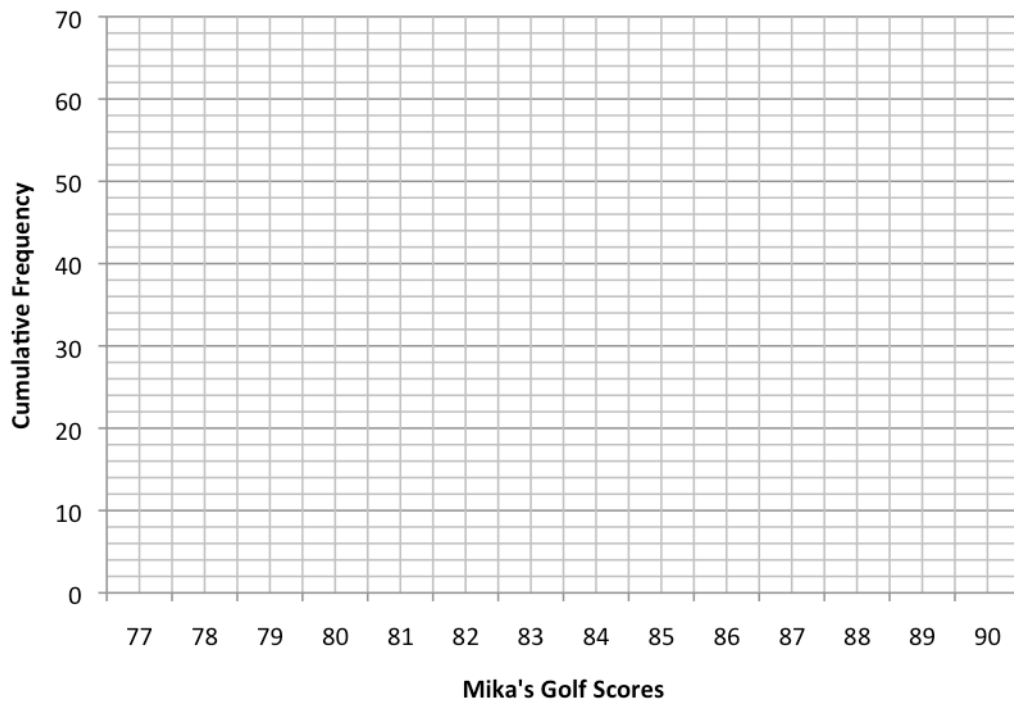Example #1 – FDT of ages of 200 first year college students at Juan Fine University

age of students	number	cum. Freq.
16	0	
17	3	
18	72	
19	62	
20	28	
21	11	
22	9	
23	5	
24	4	
25	6	
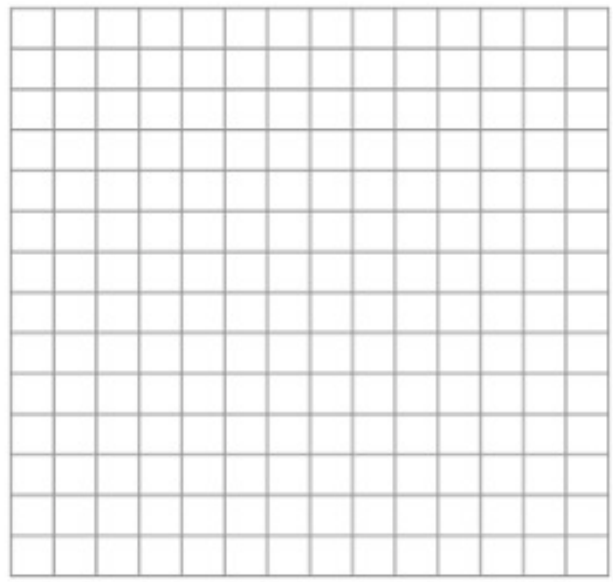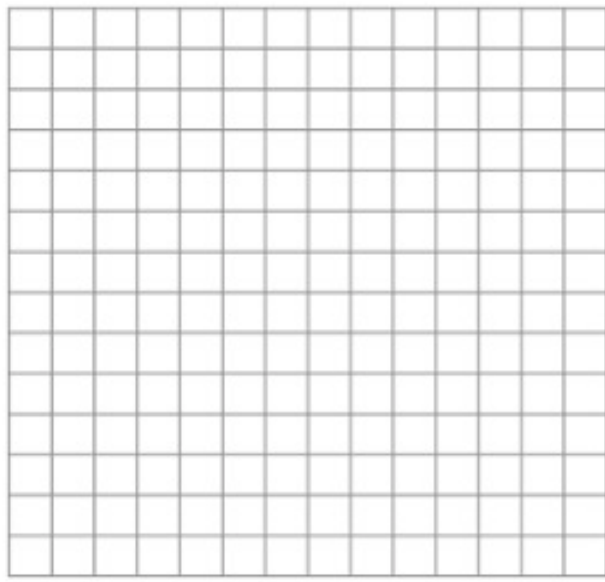
Example #2 – Mika's golf scores this past summer

a. Prepare a Frequency Histogram, frequency Polygon & CFG

Mika's Golf Scores	number	cum. Freq.
77	0	
78	1	
79	3	
80	0	
81	5	
82	7	
83	8	
84	9	
85	10	
86	8	
87	7	
88	3	
89	2	
90	1	

Example 3 - The length of 40 insects of a certain species were measured correct to the nearest millimeter. The frequency distribution is given below:

Lengths (mm)	Frequency (f_i)	Cumulative frequency
$25 \leq L < 30$	2	
$30 - 35$	4	
$35 - 40$	7	
$40 - 45$	10	
$45 - 50$	8	
$50 - 55$	6	
$55 - 60$	3	

a. Construct a cumulative frequency table for the given data.
b. Draw a cumulative frequency curve for the data.
c. Estimate from the curve
   (i) the number of insects that were less than 43.5 mm long,
   (ii) the percentage of insects that were of length 37.5 mm or more,
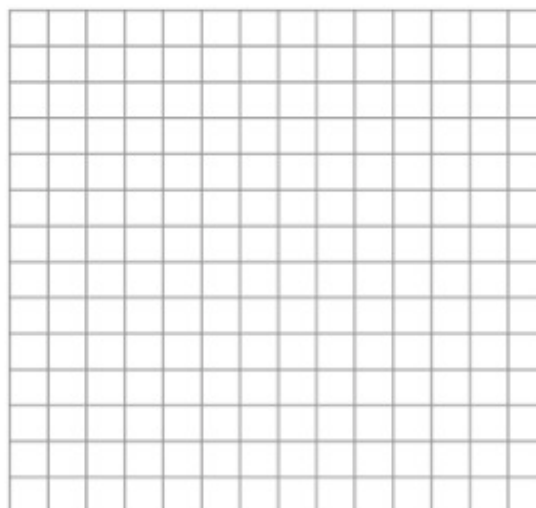   (iii) the value of k, if 75% of the insects were less than k mm long.
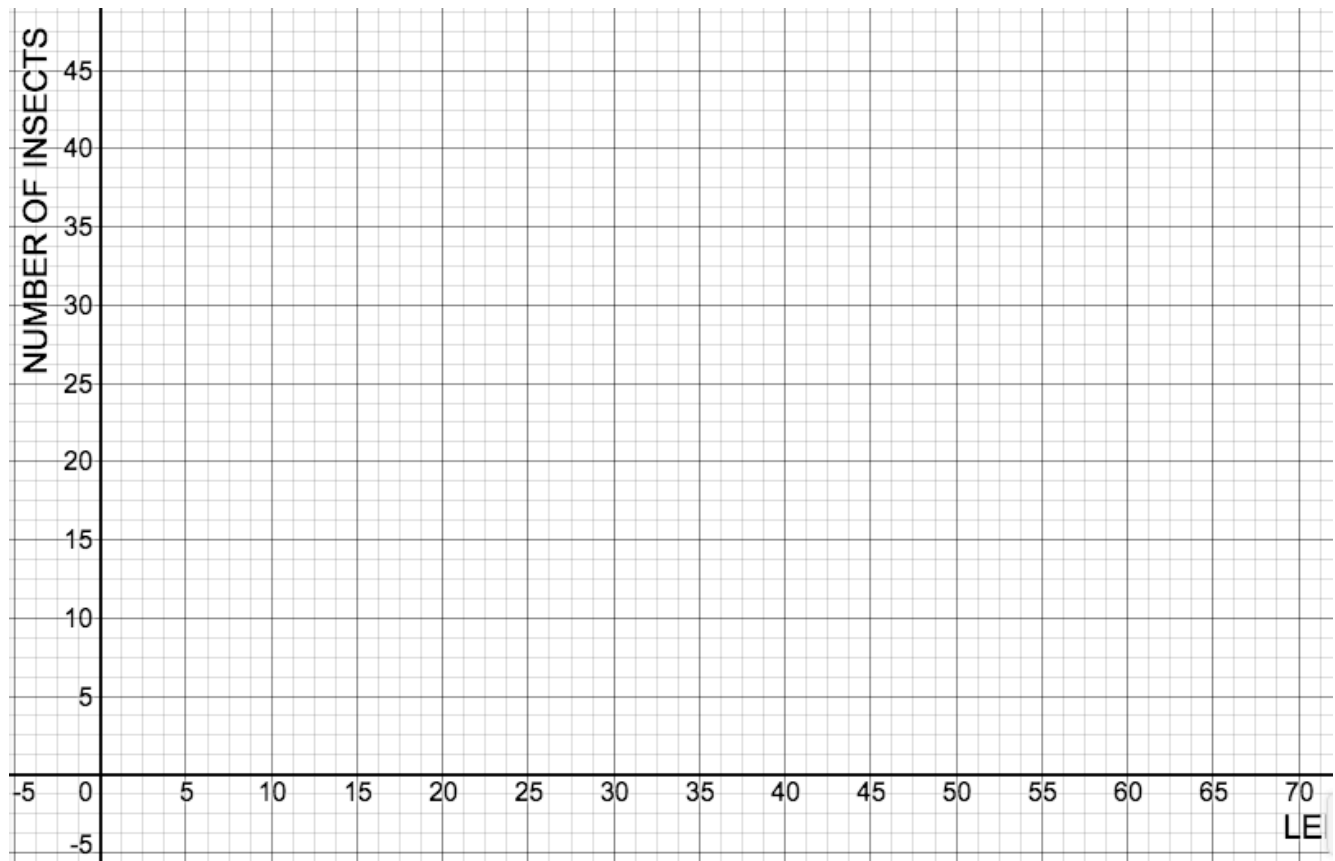
Example #4 – siblings of students in Mr. S's Grade 10 classes

a. 55% of the students in Mr S's classes have at least ………….. siblings (according to the ogive)

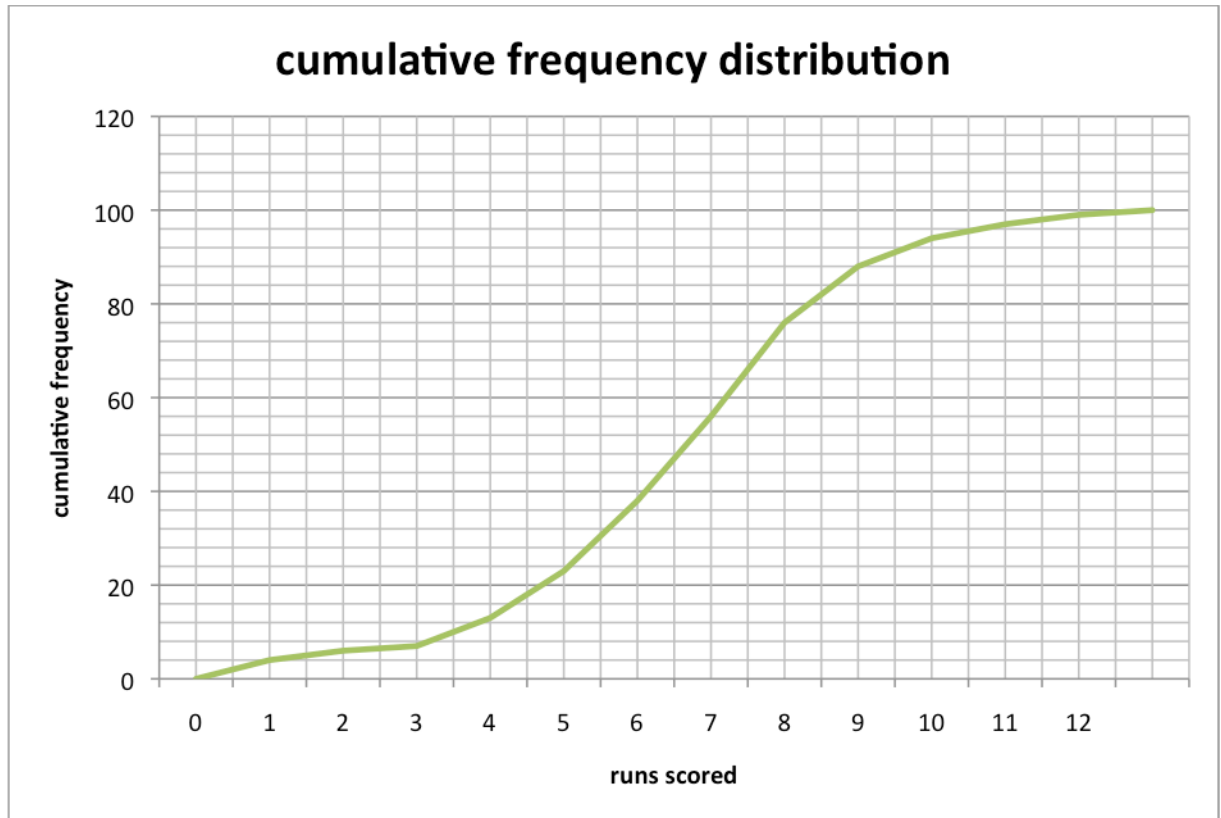b. …….. % of the students had at least 3 siblings

(A)
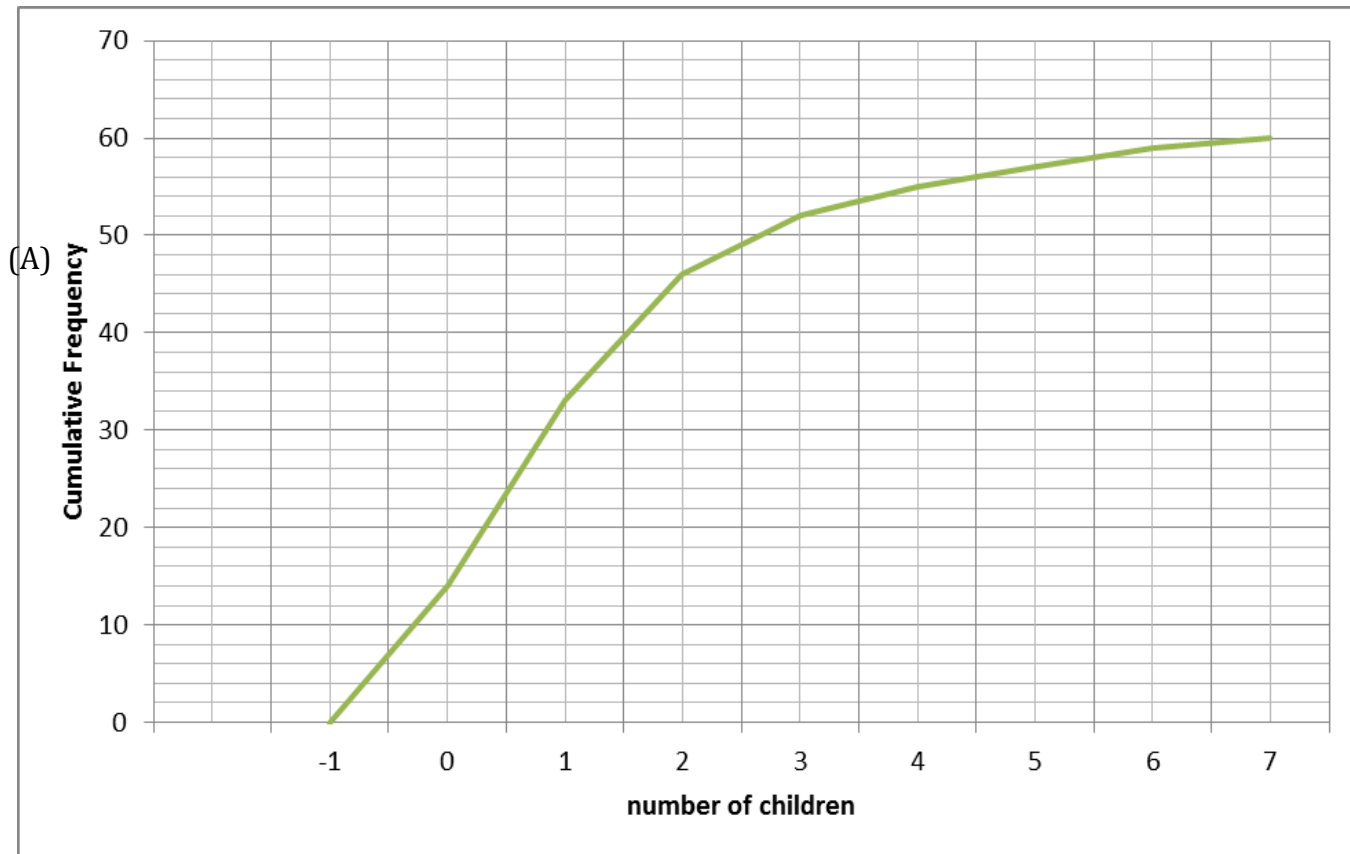
Number of children	number	cum. Freq.	
-1	0	0	
0			
1	19		
2		46	
3			
4		55	
5	2		
6			
7		60	
8	0	60	

Example #3 – Runs scored by Mr. Nicols baseball teams at ISM over the years
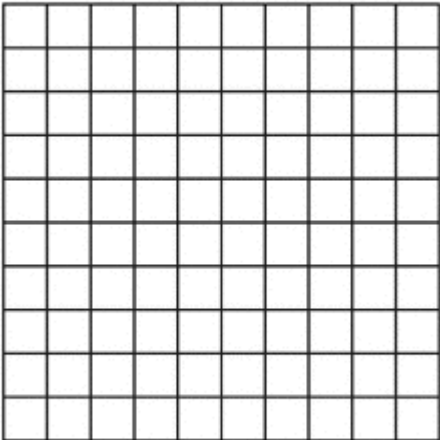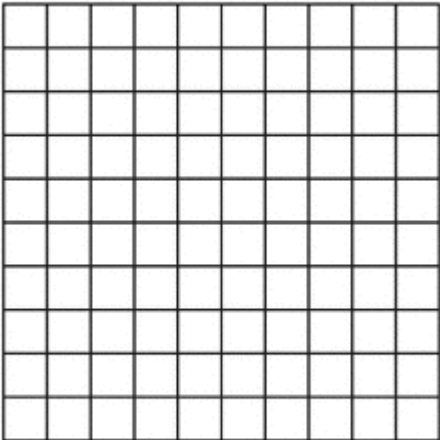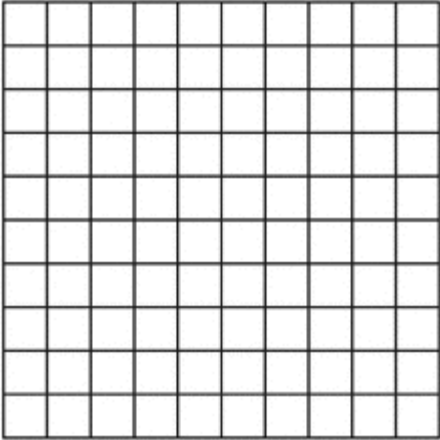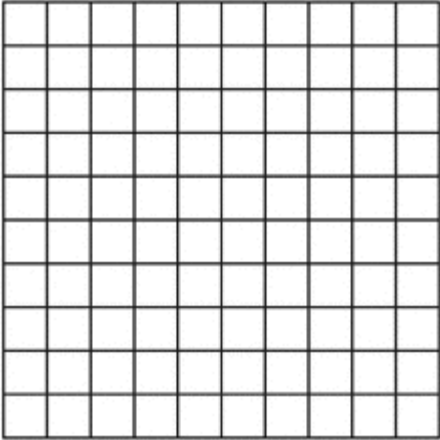
    a. 43% of the time, Mr. Nicol's baseball teams scored ……… runs (according to the ogive)

    b. In …….. % of the games, the team scored at least 5 runs

cumulative frequency distribution

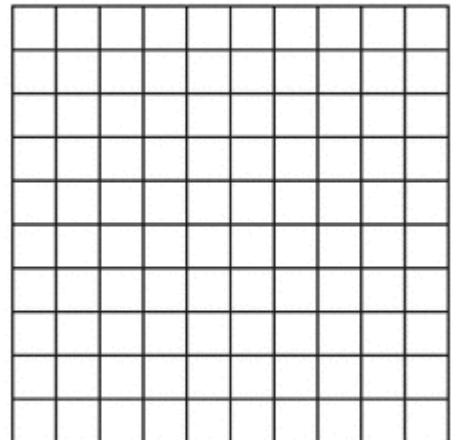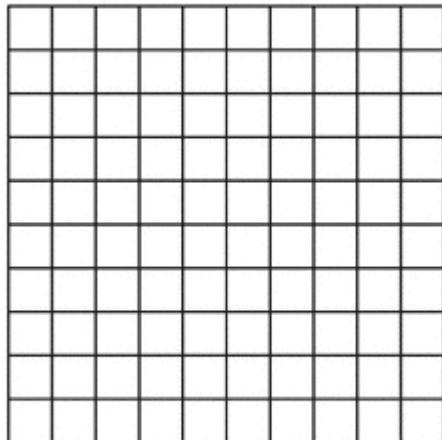| Runs in Baseball Game	number	cum. Freq.	Histogram
0			
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			

Prepare Histograms & Cumulative Frequency Graphs from the following data sets (each of which show different types of data distributions ➔ Let's say these data sets represent IM2 Final Exam scores from the past 4 years at CAC (Work in Groups of 4 to complete this activity)

IM2 SEM 2 Exam Scores	Frequency Histogram	CFG

$0 \leq M < 10$	0
10 – 20	0
20 – 30	0
30 – 40	0
40 – 50	7
50 – 60	16
60 – 70	28
70 – 80	27
80 – 90	15
90 - 100	7

$0 \leq M < 10$	0
10 – 20	2
20 – 30	3
30 – 40	4
40 – 50	3
50 – 60	6
60 – 70	12
70 – 80	18
80 – 90	33
90 - 100	19

0 ≤ M < 10	0
10 – 20	11
20 – 30	26
30 – 40	21
40 – 50	15
50 – 60	13
60 – 70	12
70 – 80	2
80 – 90	0
90 - 100	0
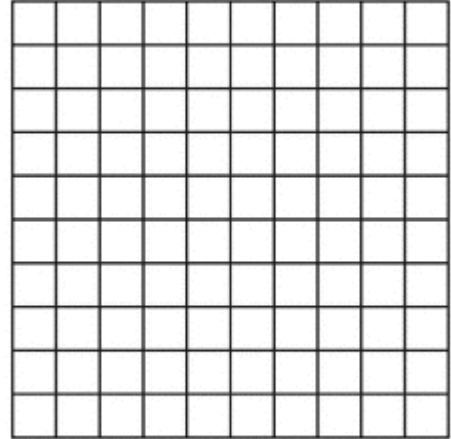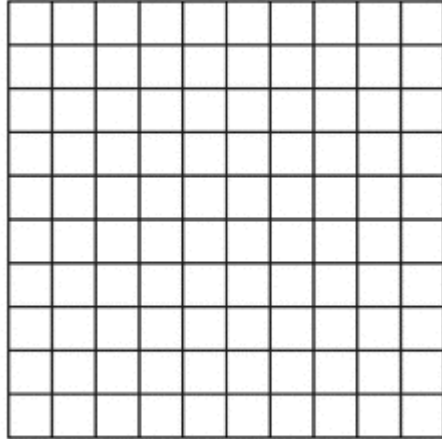
0 ≤ M < 10	2
10 – 20	15
20 – 30	20
30 – 40	12
40 – 50	3
50 – 60	2
60 – 70	10
70 – 80	20
80 – 90	14
90 - 100	2

(A) Lesson Context

BIG PICTURE of this UNIT:	• How do we analyze and then make conclusions from a data set? (Math) • How do I present my data and the outcomes of my analysis? (Math) • How do I use data & statistics to make decisions? • How do I decide on the validity/reliability of my data? Of my analysis? Of my conclusions? Of my decision?		
CONTEXT of this LESSON:	Where we've been Prepare and analyze frequency histograms, frequency polygons and cumulative frequency graphs	Where we are One set of numbers we can calculate in order to analyze a data set is the measures of central tendency … how do you find the "center" of a data set?	Where we are heading How do I analyze and make conclusions from a data set, in whatever way this data gets presented?

(B) Lesson Objectives:

a. Starting from a set of raw data, calculate three measures of central tendency (mean, median & mode)
b. Determine the effect of different data distributions on measures of central tendency
c. Perform different types of calculations involving means

Opening Question ➔ What is the purpose of finding an "average?"

(C) Measures of Central Tendency – Define each of the following terms

a. Mean

b. Median

c. Mode

d. Range

(D) Measures of Central Tendency from a Data List – Example #1

Here are the weights of 7 students in class (measured in pounds): 150, 160, 173, 155, 160, 175, 170. Calculate the mean, median, mode and range of student weights.

(E) Measures of Central Tendency from a Data List – Example #2

Samuel is trying to determine the average height of high school male students. Because he is on the basketball team, he uses the heights of the 14 players on the team, which are given below in inches.

 69, 70, 72, 72, 74, 74, 74, 75, 76, 76, 76, 77, 77, 82

 (a) Calculate the mean, median, mode and range for this data set. Round any non-integer answers to the nearest tenth.

(b) Is the data set above a fair sample to use to determine the average height of high school male students? Explain your answer.

(F) Measures of Central Tendency from a Data List – Example #3

The tuition costs for ten private schools in Florida are $7568, $8650, $9225, $5880, $6720, $8840, $7820, $ 8260, $ 8432, and $8990. Find the mean & median & mode & range of tuition costs.

(G)Measures of Central Tendency from a Data List – Example #4

Example – The stem and leaf plot represents the scores on the Chapter 5 test in Mrs. Jones' geometry class. Find the mean, median, mode and range of scores.

Geometry Test Scores	
Stem	**Leaf**
5	6 8 9
6	1 6 9
7	4 5 77 99
8	2 4 6 7 7 8 8 9
9	1 3 3 4 4 5 5 5 7
10	0 0

(H)Measures of Central Tendency from a Data List – Example #5

The high temperatures for a 7-day week during December in Chicago were 29°, 31°, 28°, 32°, 29°, 27°, and 55°.
(a) Find the mean high temperature for the week.

(b) In this example, is this mean temperature a good representation of the data?

(c) Find the median temperature for Chicago during this week.

(d) Which measure of central tendency is a better indicator of the "central tendency" of temperatures in Chicago this week?

(I) Measures of Central Tendency from a Data List – Example #6

Exercise #1: Students in Mr. Ramirez's statistics class were trying to determine if people speed along a certain section of roadway. They collected speeds of 20 vehicles, as displayed in the table below.

(a) Find the mean, median, and mode for this data set.

Speed (mph)	Number of Cars
29	1
33	2
34	4
35	5
36	3
38	2
39	2
54	1

(b) The speed limit along this part of the highway is 35 mph. Based on your results from part (a), is it a fair to make the conclusion that the average driver does speed on this roadway?

Exercise #2: In Mr. Smith's Advanced Calculus Course, eight students recently took a test. Their grades were as follows: 45, 78, 82, 85, 87, 89, 93, 95

(a) Calculate the mean and median of this data set.   (b) What score is an outlier in this data set?

(c) Which value, the mean or the median, is a better measure of how well the average student did on Mr. Smith's quiz?

(J) <u>Measures of Central Tendency from a Data List – Example #7</u>

Exercise #1: Tony's Luncheonette is open six days a week. His lunch business income for the five days this week is $120, $110, $200, $300, and $140.

(a) How much money must he total for the six days if he wants to have a mean income of $200 per day?

(b) What must Tony bring in on the 6th day in order to have an average of $200 for the six days?

Exercise #2: In order to pass Algebra 1 for the year, Darrien must have a mean of 65% for the four marking periods and the final, where all five grades have equal weight. Darrien has already earned a 67%, a 61%, and a 53% for the first three marking periods.

(a) What must Darrien's final two grades average (4th marking period and final exam) so that she will pass the class?

(b) What is the minimum grade Darrien can earn in the fourth marking period and still be able to pass the overall class?

Exercise #3: At Charleston High School, a student needs a 93% average to earn an A. Chris has four scores in Algebra 1 of 94%, 93%, 96%, and 91%. What is the lowest score he could earn on the fifth test and still earn an A?

Exercise #4: High temperatures for a five day span in May happened to be consecutive even integers whose mean was 66 degrees. Find the lowest temperature of this five day span.
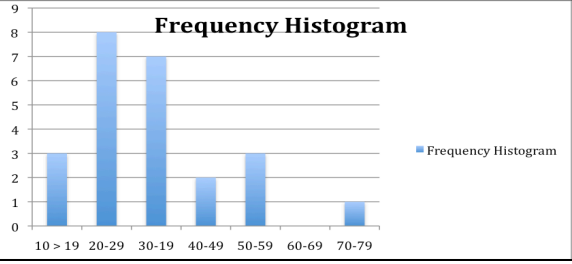
(A) Lesson Context

BIG PICTURE of this UNIT:	• How do we analyze and then make conclusions from a data set? • How do I present my data and the outcomes of my analysis? • How do I use data & statistics to make decisions? • How do I decide on the validity/reliability of my data? Of my analysis? Of my conclusions? Of my decision?

	Where we've been	Where we are	Where we are heading
CONTEXT of this LESSON:	Prepare and analyze frequency histograms, frequency polygons and cumulative frequency graphs	One set of numbers we can calculate in order to analyze a data set is the measures of central tendency … how do you find the "center" of a data set?	How do I analyze and make conclusions from a data set, in whatever way this data gets presented?

(B) Lesson Objectives:

a. Starting from a frequency table for discrete data, calculate three measures of central tendency (mean, median & mode)
b. Starting from a grouped frequency table for continuous data, calculate three measures of central tendency (mean, median & mode)
c. Starting from a cumulative frequency graph, calculate three measures of central tendency (mean, median & mode)

(C) Opening Activity ➜ Fitting Data to Central Tendencies

Your job will be to create a data set that matches the mean, median, mode, and frequency table given.	
Mean: 34	
Median: 31	
Mode: 25	

Data:

Story Behind the Data:

(D)Measures of Central Tendency of Discrete Data from a Frequency Table

Example #1: Here are the scores for the last season for Mr. S's Football team.

Football Score	Frequency		List data in order:
0	2		
1	2		
2	10		
3	12		
4	5		
5	1		Mean:
6	0		Median:
7	2		Mode:
8	0		

(a) What percentage of their games did they score more than 3 Points?

(b) If they had an average of 3.1 points scored against them per game, estimate what percentage of games they won.

Example #2: A survey was given to a random sample of freshman at a local college asking their ages. This is the data that came back.

Age	Frequency		Find the following. Discuss what are good methods for doing this.
17	6		
18	34		Mean:
19	59		
20	81		
21	16		
22	11		Median:
23	5		
24	8		Mode:
25	2		

(a) What percentage of students are over 19?

(b) If the entire freshman class consists of 15,698 students, estimate how many people on campus are under the age of 21.

(E) Measures of Central Tendency of Continuous Data from a Grouped Frequency Table

Example #1: Mr. Mello's class went on a field trip to the Amazon Jungle. While exploring their class took some height samples for local plant life. Here is the data they recorded.

Length in cm	mid interval length	Frequency	Cumulative Frequency	Find the following. Discuss what are good methods for doing this.
[0 – 10]		6		
(10 – 20]		34		**Mean:**
(20 – 30]		81		
(30 – 40]		59		
(40 – 50]		16		**Median:**
(50 – 60]		11		
(60 – 70]		5		**Modal Class/Interval:**
(70 – 80]		8		
(80 – 90]		2		

Which measure of central tendency best reflects the height of an "average" plant in the Amazon? Explain

Example #2: Below are the times recorded for Grade 9 students' mile run in Mr. S's PE Classes.

Time in minutes	mid-interval time	Frequency	Cumulative Frequency	Estimate the following. Discuss what are good methods for doing this.
(5:00 – 5:20]	5.167	1		
(5:20 – 5:40]	5.5	0		**Mean:**
(5:40 – 6:00]	5.833	3		
(6:00 – 6:20]	6.167	2		
(6:20 – 6:40]	6.5	5		**Median:**
(6:40 – 7:00]	6.833	4		
(7:00 – 7:20]	7.167	8		**Modal Class/Interval:**
(7:20 – 7:40]	7.5	15		
(7:40 – 8:00]	7.833	12		

(a) How many students does Mr. S. have in his P.E. Classes?

(b) Which measure of central tendency best reflects the mile run time of an "average" student in Grade 9 . Explain.

(F) Measures of Central Tendency from a Cumulative Frequency Graph

Example #1: Below is a Cumulative Frequency Graph for the AP Math Scores from the past two year at CAC.

(a) Estimate the mean given the following CFG. Discuss what are good methods for doing this.

"Cumulative Frequency AP Math Scores"
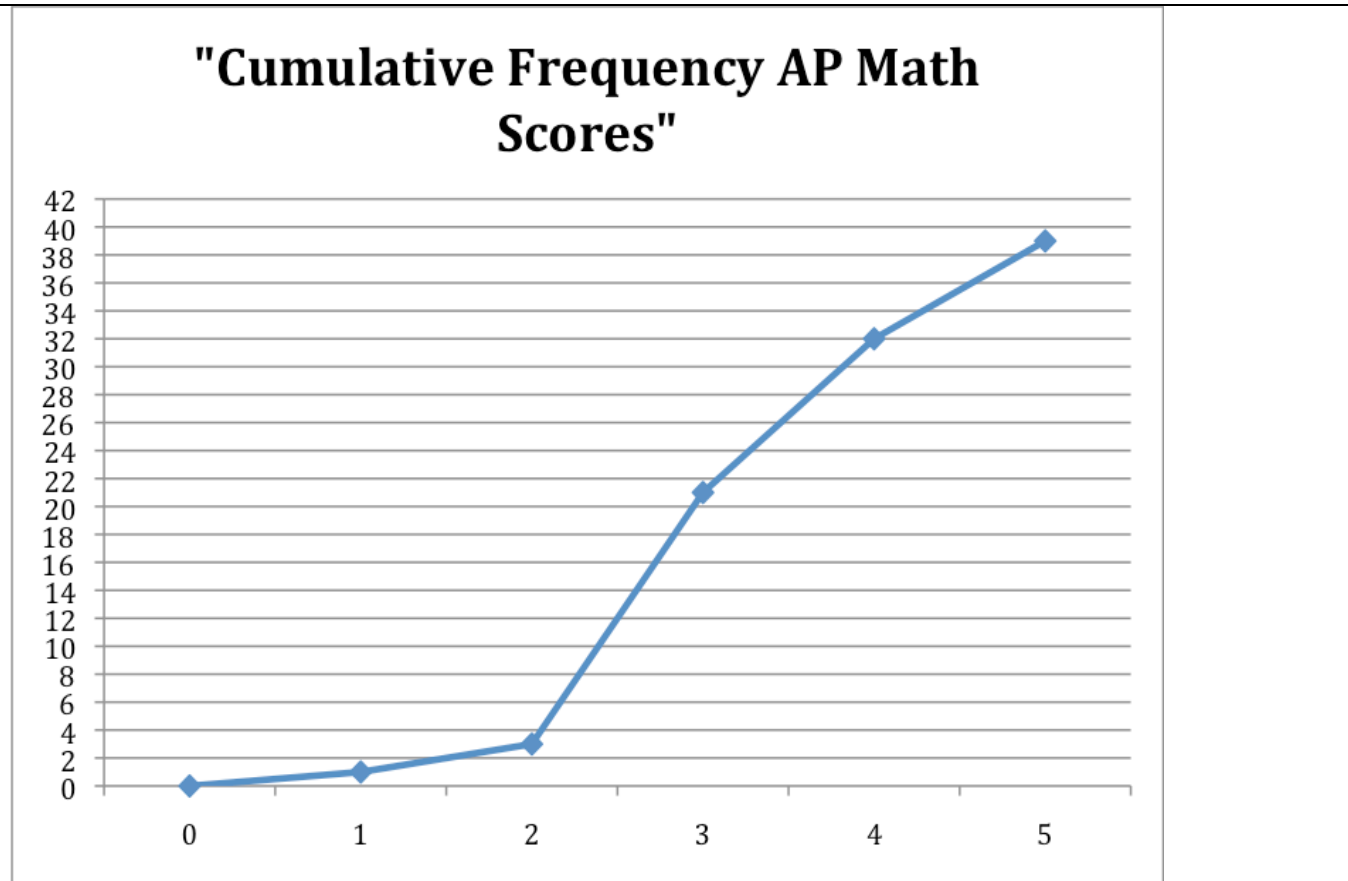
From the CFG, prepare a frequency table, then use your frequency table to list all the AP Scores for the last two years:

Score	Frequency
0	
1	
2	
3	
4	
5	

(b) From your list, determine the mean, median & mode

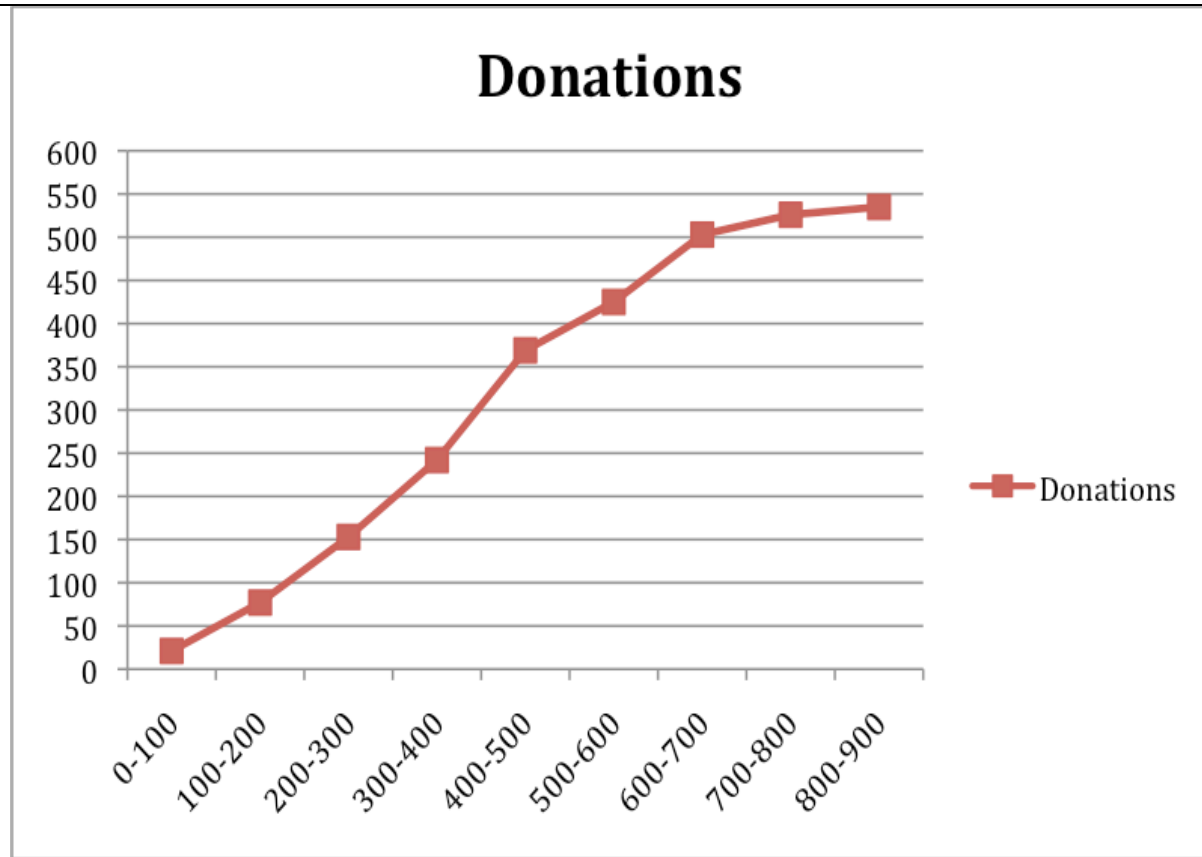Example #2: The High School has just finished a fundraiser for a local charity during the latest Spirit Week. This fundraiser brought in quite a lot of money. Below is a cumulative Frequency table showing the amount of donations that came in.

(a) Estimate the mean given the following CFG. Discuss what are good methods for doing this.

Donations

(b) From the CFG, prepare a frequency table.

(b) Use the CFG to estimate the donation amount given by the lower 25% of the donors	(c) Use the CFG to estimate the donation amount given by the upper 25% of the donors

(A) Lesson Context

BIG PICTURE of this UNIT:	• How do we analyze and then make conclusions from a data set? • How do I present my data and the outcomes of my analysis? • How do I use data & statistics to make decisions? • How do I decide on the validity/reliability of my data? Of my analysis? Of my conclusions? Of my decision?
CONTEXT of this LESSON:	Where we've been Determining the central tendency of a data set, given an assortment of data presentations	Where we are Measuring the dispersion of a data set, using range & quartiles and visualizing the dispersion using a box and whisker plot	Where we are heading How do I analyze and make conclusions from a data set, in whatever way this data gets presented?

(B) Lesson Objectives:

    a. Calculate the range of a data set
    b. Calculate the quartiles of a data set
    c. Present the quartiles using a visual ➔ a box & whisker plot

(C) Opening Activity ➔ Which Class did "better"?

Central Tendencies: Results Summary from Class A	Central Tendencies: Results Summary from Class B
—Mean grade ➔ 63% —Median grade ➔ 60% —Modal Grade ➔ 60%, 63%	—Mean grade ➔ 63% —Median grade ➔ 60% —Modal Grade ➔ 60%
Can we determine which class did "better" if all we know are the central tendencies?	

(D) **Opening Activity ➔ Which Class did "better" – More Information added?**

Central Tendencies: Results Summary from Class A	Central Tendencies: Results Summary from Class B
—Mean grade ➔ 63% —Median grade ➔ 60% —Modal Grade ➔ 60%, 63% —Range ➔ 70%	—Mean grade ➔ 63% —Median grade ➔ 60% —Modal Grade ➔ 60% —Range ➔ 40%
Can we determine which class did "better" after we have now added the range to the central tendencies?	

(E) **Measures of Spread/Dispersion/Variance – Range**

So what else do we need? ➔ we need some way to come up with some other "summary numbers" that allow us to get a picture of the "spread" or "dispersion" or "distribution" or "variance" in the data

One such number is the RANGE (as we have discussed before)

The range is simply the difference between the largest and the smallest observed values in a data set. Thus, range, including any outliers, is the actual spread of data. A great deal of information is ignored when computing the range, since only the largest and smallest data values are considered.

(F) **Measures of Spread/Dispersion/Variance – Quartiles**

Quartiles simply split data into quarters:

The first quartile or the lower quartile is ¼ of the way into an ordered data set ➔ cuts off lowest 25% of data = *25th percentile*

The third quartile or the upper quartile is ¾ of the way into an ordered data set ➔ **third quartile** *(designated Q_3)* = **upper quartile** = cuts off highest 25% of data, or lowest 75% = *75th percentile*

The second quartile or the median is 2/4 (or ½) of the way into an ordered data set ➔ **second quartile** *(designated Q_2)* = **median** = cuts off data set in half (50% of data) = *50th percentile*

(G) Example #1 – Analyzing Quiz Scores ➔ 5 Number Summary

Here are QUIZ results from BOYS in our IM2 classes. The scores have already been sorted/ordered.

3, 3, 3, 3, 4, 4, 4, 5, 6, 6, 7, 7, 8, 8, 8, 8, 8, 9, 9, 9, 9, 10, 10

(a) Determine the minimum score ➔
(b) Determine the maximum score ➔
(c) Calculate the range of the scores ➔
(d) Determine the lower quartile score ➔
(e) Determine the median score ➔
(f) Determine the upper quartile score ➔

Here are QUIZ results from GIRLS in our IM2 classes. The scores have already been sorted/ordered.

4, 4, 5, 5, 5, 6, 6, 6, 7, 8, 8, 8, 9, 9, 9, 9, 9, 9, 10, 10, 10, 10, 10

(a) Determine the minimum score ➔
(b) Determine the maximum score ➔
(c) Calculate the range of the scores ➔
(d) Determine the lower quartile score ➔
(e) Determine the median score ➔
(f) Determine the upper quartile score ➔

What observations & conclusions can we make from looking at the results of the data calculations?

(H) Example #1 – Analyzing Quiz Scores ➔ Visualizations with Box & Whisker Plots

A box plot is a way of illustrating key information about a set of data

They are also very useful for comparing the distribution of two sets of data

To draw a box plot, you need FIVE pieces of information: (i) The minimum value, (ii) the lower quartile (iii) the median, (iv) the upper quartile (v) the maximum value

Interquartile Range

Minimum Value   Lower Quartile   Upper Quartile   Maximum Value

Median
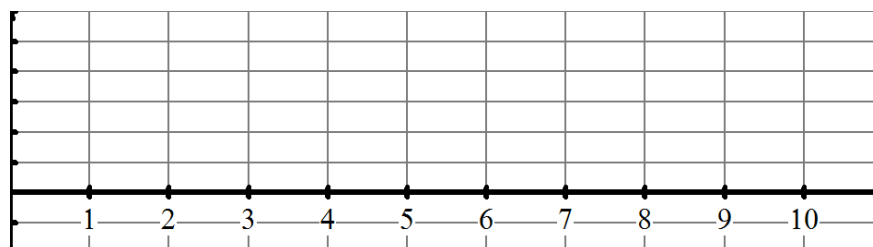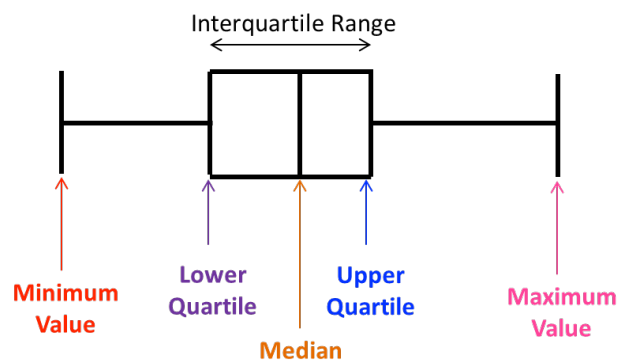
Example #2 – Comparing Golf Clubs

Below are the ages of 15 members of a golf club.

20, 20, 21, 24, 27, 29, 33, 33, 36, 39, 50, 57, 60, 65, 65

Below are the ages of members of a second golf club.

18, 23, 25, 27, 27, 37, 38, 45, 47, 49, 49, 50, 50, 58, 61, 66, 69

(a) Determine the minimum age ➔
(b) Determine the maximum age ➔
(c) Calculate the range of the age ➔
(d) Determine the lower quartile age ➔
(e) Determine the median age ➔
(f) Determine the upper quartile age ➔

(g) Determine the minimum age ➔
(h) Determine the maximum age ➔
(i) Calculate the range of the age ➔
(j) Determine the lower quartile age ➔
(k) Determine the median age ➔
(l) Determine the upper quartile age ➔
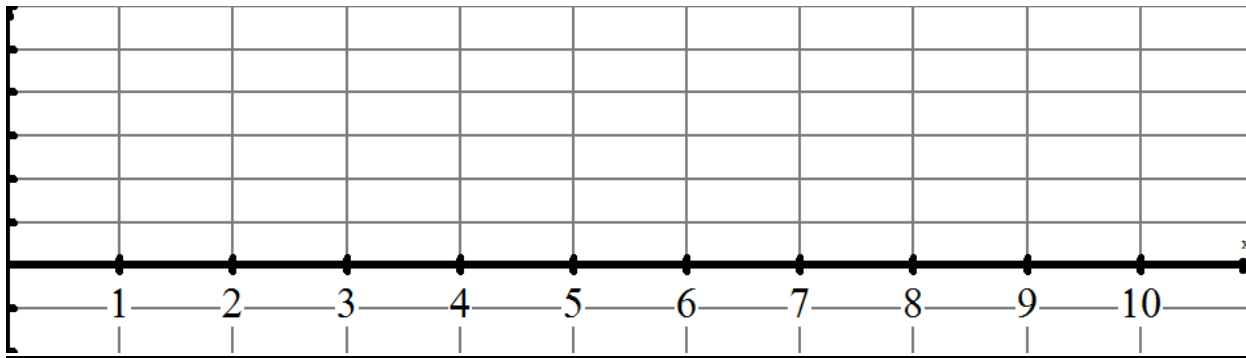
•Which golf club has more younger members?

•Which golf club has more older members?

•What can you say about the age profile of each golf club in general?

(I) **Example #3**

1. Here are the marks scored by 15 students in their quadratic functions test.
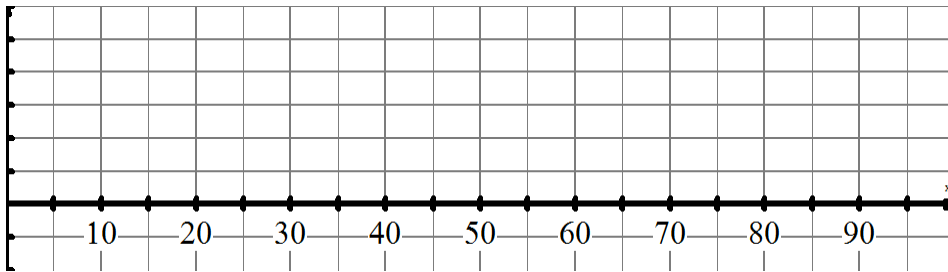
    75 82 69 82 95 80 74 77 60 51 68 84 90 100 45

    a) Find the median, quartiles and interquartile range for these data

   b) In a probability test the following summary data was collected

Minimum	Lower Quartile	Median	Upper Quartile	Maximum
60	72	78	84	98

    c) On the same axis draw box plots for both tests.

    d) Compare and contrast the results from the two tests. (Which test did they do better? Which test had more varied results?).
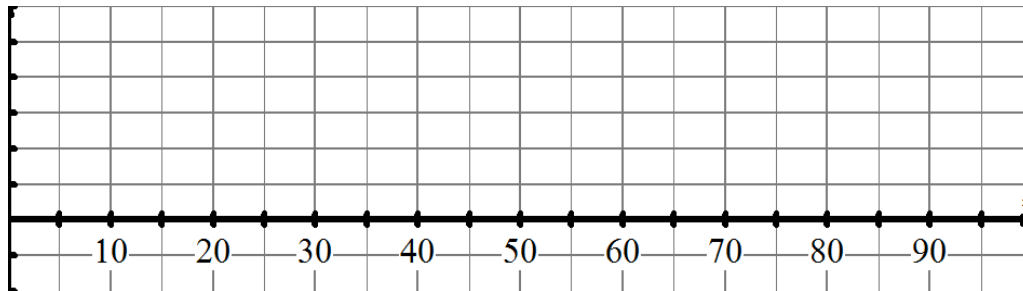
2. This back-to-back stem and leaf diagram shows heart rates of two groups of students before and after they walk up the stairs

				Before			After					
8	8	4	3	2	6	8						
9	9	7	5	4	3	7	0	1	5	6	6	7	9
		3	2	0	8	0	0	1	4	8		
				0	9	2						
					10	0						
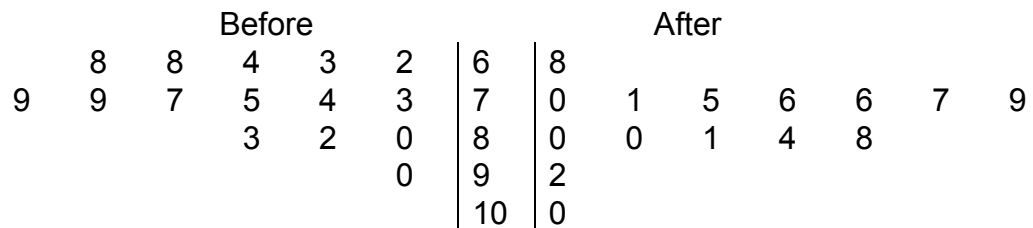
    legend | 6 | 7   = 67 beats per minute

a) Find the median and quartiles for these data.

b) Draw a box plot for each on the same scale

c) Compare the heart rates before and after the walk up the stairs.

3. The data below shows the price of gas (to the nearest cent) at different locations for Shell and BP.

   SHELL: 130, 129, 132, 133, 136, 130, 129, 130, 130, 131, 133, 134, 140

   BP: 131, 133, 133, 133, 132, 134, 138, 130, 133, 134, 135, 136, 137, 132, 135

a) Calculate the quartiles, median and interquartile range for each subject

b) Use your data to draw box plots for each subject

c) Comment on at least two aspects of your box plots to compare the IQs of the graduates.

4. This summary data has been collected on times taken, in minutes, to run 5km in two different fun runs in Victoria Park.

Run 1	
Median	35
Minimum	23
Range	40
Interquartile range	12
Upper Quartile	42

Run 2	
Median	36
Maximum	55
Range	30
Interquartile range	15
Lower Quartile	29

a)   Draw box plots for both runs using the same scale

b)   Write down two comparisons between the races

c)   Emma says that Race 1 attracted more serious runners. Use your box plots to comment on what Emma has said.

5.   The data below shows the IQ of 11 Math and 11 Geography university graduates.

MATH: 98, 103, 105, 99, 110, 94, 98, 100, 120, 117, 113
GEOGRAPHY: 93, 99, 110, 111, 95, 97, 90, 99, 92, 102, 103

a)   Calculate the quartiles, median and interquartile range for each subject

b)   Use your data to draw box plots for each subject

c)   Comment on at least two aspects of your box plots to compare the IQs of the graduates.

(A) Lesson Context

BIG PICTURE of this UNIT:	• How do we analyze and then make conclusions from a data set? • How do I present my data and the outcomes of my analysis? • How do I use data & statistics to make decisions? • How do I decide on the validity/reliability of my data? Of my analysis? Of my conclusions? Of my decision?		
CONTEXT of this LESSON:	Where we've been Determining the central tendency of a data set, given an assortment of data presentations	Where we are Measuring the dispersion of a data set, using range & quartiles and visualizing the dispersion using a box and whisker plot	Where we are heading How do I analyze and make conclusions from a data set, in whatever way this data gets presented?

(B) Lesson Objectives:

a. Construct frequency histograms, polygons & CFG and use them to determine quartiles & draw BW plots
b. Calculate the quartiles of a data set
c. Present the quartiles using a visual ➔ a box & whisker plot

(C) Requirements for each data set

From the following grouped frequency table, you are required to:

(a) Prepare a frequency histogram & then a Frequency Polygon
(b) Use the table to ESTIMATE the mean
(c) Prepare Cumulative Frequency Graph
(d) Use the CFG to ESTIMATE the median, Q1 & Q3
(e) Prepare a B&W plot

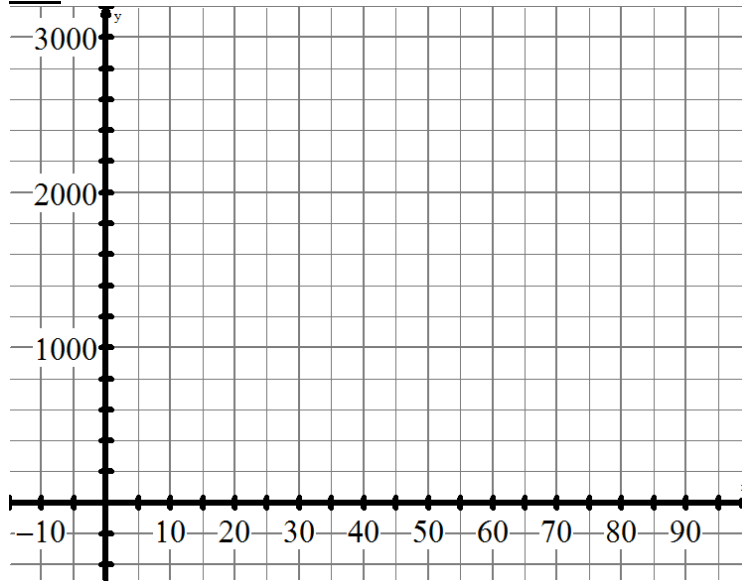**(D)Example #1** Ages of players in a community football league

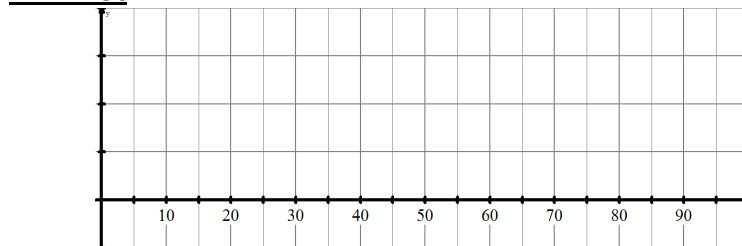| ages	# of players
0 - 5	70
5 - 10	400
10 - 15	850
15 - 20	600
20 - 25	400
25 - 30	200
30 - 35	150
35 - 40	100
40 - 45	100
45 - 50	80
50+	50

Histogram

Mean and Median Estimates:

CFG
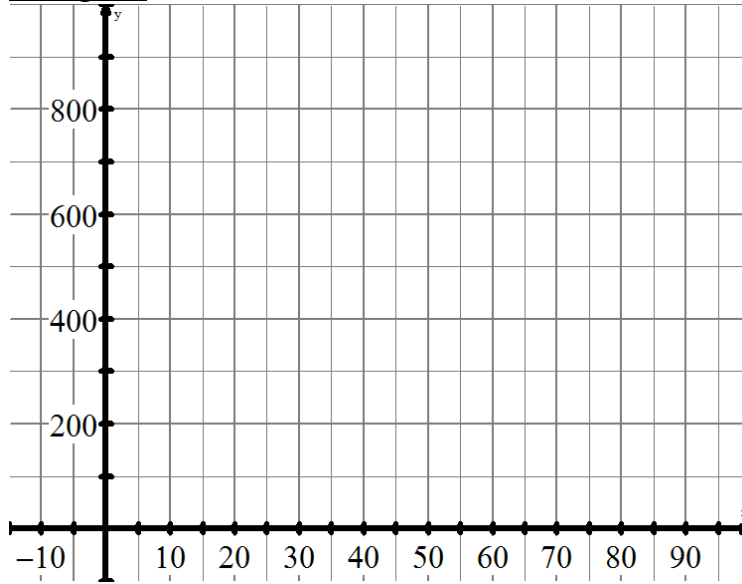
Quartiles

B-W Plot

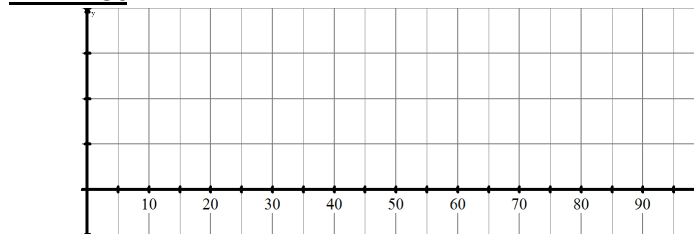**Example #2** years of experience of doctors in city A hospitals

years of experience	Frequency
0 - 5	1400
5 - 10	1750
10 - 15	1150
15 - 20	800
20 - 25	400
25 - 30	200
30 - 35	120
35 - 40	80
40 - 45	50
45 - 50	30
50 - 70	20

Histogram

Mean and Median Estimates:

CFG
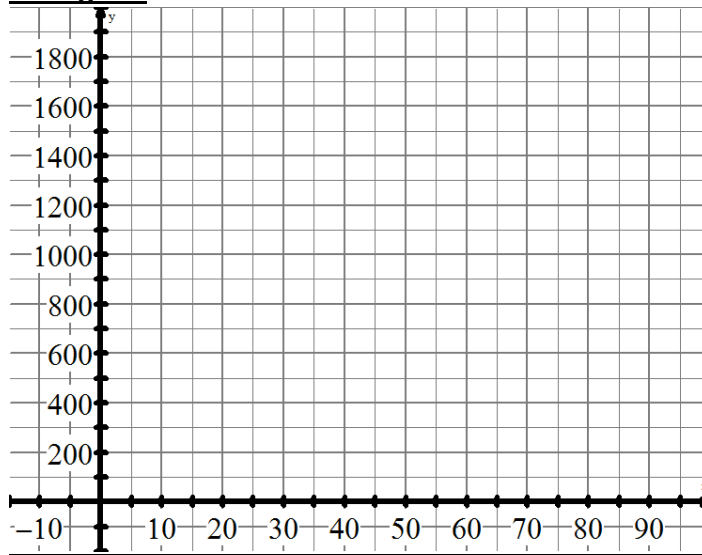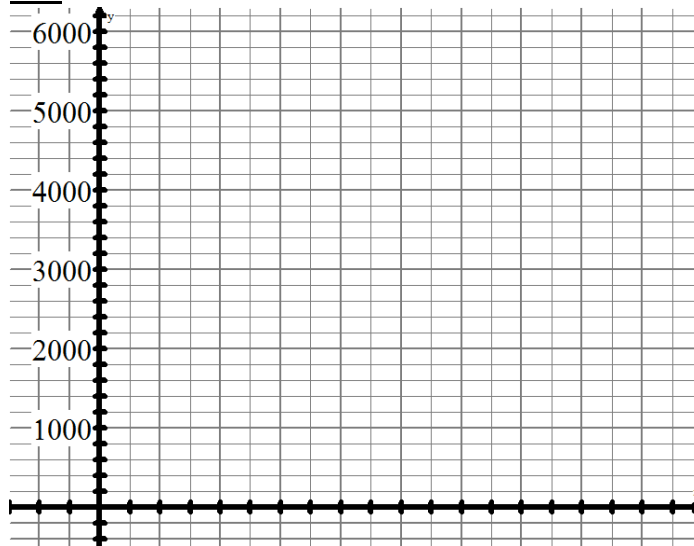
Quartiles

B-W Plot

(E) Example #3 years of experience of doctors in city B hospitals
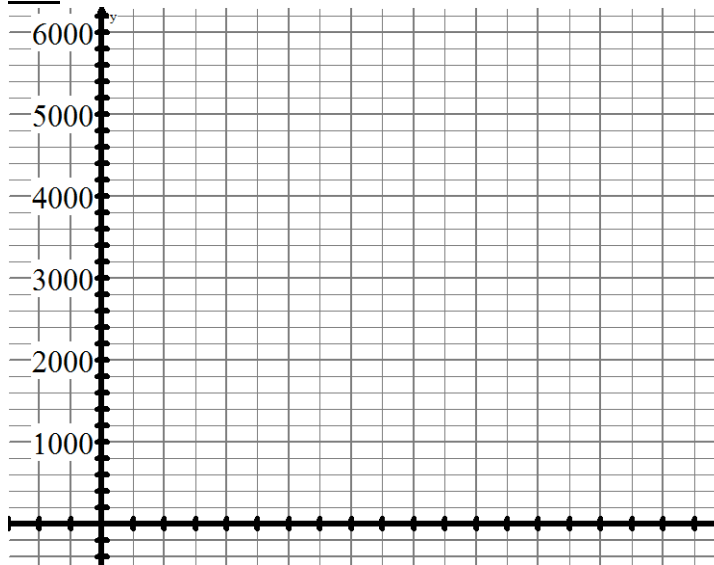
years of experience of
doctors in city B hospitals

Years	Frequency
0 - 5	250
5 - 10	600
10 - 15	1100
15 - 20	1800
20 - 25	1200
25 - 30	800
30 - 35	180
35 - 40	50
40 - 45	20
45 - 50	0
50 - 70	0

Histogram

Mean and Median Estimates:

CFG
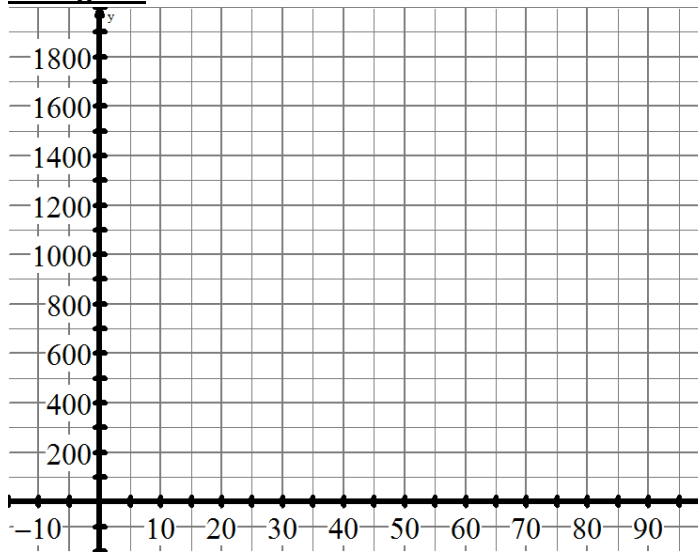
Quartiles

B-W Plot
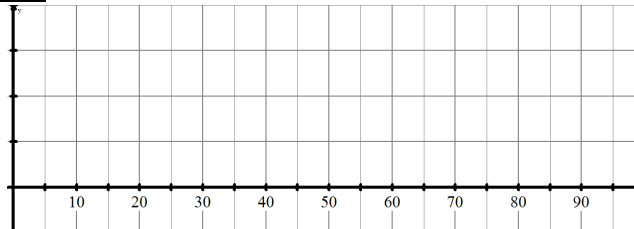
Example #4 Number of Research Papers published

Number of Research Papers published	Frequency
0 - 3	10
3 - 6	30
6 - 9	40
9 - 12	75
12 - 15	150
15 - 18	225
18 - 21	500
21 - 24	600
24 - 27	270
27 - 30	100

Histogram

Mean and Median Estimates:

CFG
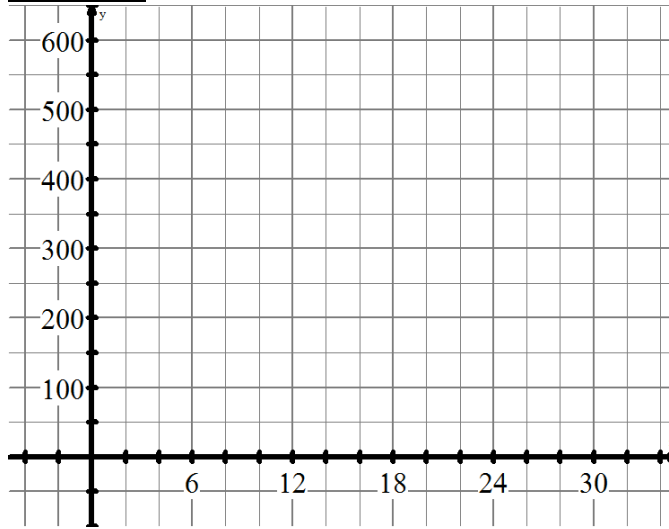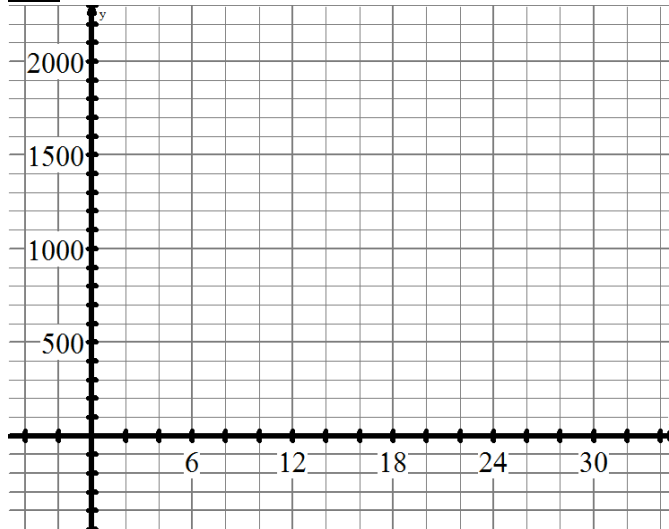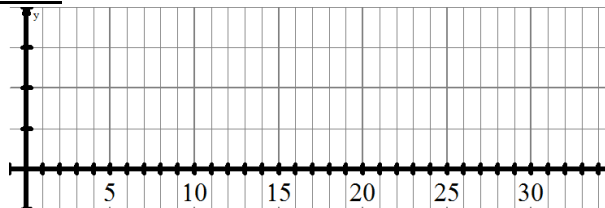
Quartiles

B-W Plot

(F) Example #5 - Ages of Teaching Faculty in U of T

Ages of Teaching
Faculty in U of T

Ages	Frequency
25 - 30	25
30 - 35	50
35 - 40	90
40 - 45	120
45 - 50	210
55 - 60	220
60 - 65	135
65 - 70	90
70 - 75	40
75 - 80	20

Histogram

Mean and Median Estimates:

CFG

Quartiles

B-W Plot