Unit 6
Statistics

Statistical Questioning

Mean, Median, Mode & Range

Dot Plots

Frequency Tables

Histograms

Box Plots

Mean Absolute Deviation

Name: _____

Mrs. Bothers / Period _____

Math 6/7
Unit 6 IXL Log

	Required Skills	
	Skill	**Your Score**
Week of 1/6	**HH.1** (Identify Statistical Questions)	
	HH.2 (Calculating Mean, Median, Mode & Range)	
	GG.3 (Interpret Line Plots)	
	HH.5 (Changes in Mean, Median, Mode & Range)	
	HH.New (Identify an Outlier and Describe the Effect)	
Week of 1/13	**HH.4** (MMMR: Find the Missing Number)	
	GG.12 (Create Histograms)	
	HH.3 (Interpret Charts and Graphs to Find MMMR)	
	GG.11 (Interpret Histograms)	
Week of 1/21	**GG.18** (Interpret Stem and Leaf Plots)	
	HH.7 (Calculate Quartiles and IQR)	
	GG.20 (Interpreting Box & Whisker Plots)	

Unit 6: Statistics
Standards, Checklist and Concept Map

Georgia Standards of Excellence (GSE):

MGSE6.SP.1: Recognize a statistical question as one that anticipates variability in the data related to the question and accounts for it in the answers.

MGSE6.SP.2: Understand that a set of data collected to answer a statistical question has a distribution which can be described by its center, spread, and overall shape.

MGSE6.SP.3: Recognize that a measure of center for a numerical data set summarizes all of its values with a single number, while a measure of variation describes how its values vary with a single number.

MGSE6.SP.4: Display numerical data plots on a number line, including dot plots, histograms, and box plots.

MGSE6.SP.5 : Summarize numerical data sets in relation to their context, such as by:
a. Reporting the number of observations.
b. Describing the nature of the attribute under investigation, including how it was measured and its units of measurement
c. Giving quantitative measures of center (median/mean) and variability (interquartile range, mean absolute deviation), as well as describing any overall pattern or any striking deviations from the overall pattern with reference to the context in which the data was gathered.
d. Relating the choice of measures of center and variability to the shape of the data distribution and the context in which the data was gathered.

••

What Will I Need to Learn??

_____ Know that a statistical question will receive many different answers

_____ To produce questions that include a variety of answer choices

_____ To interpret graph of data by describing its center, spread, and shape

_____ Know that mean, median, & mode describe the center of a set of data

_____ Know that range describes how much a set of data varies

_____ How to display data in a dot plot

_____ How to display data in a histogram

_____ How to display data in a box plot

_____ To describe the topic in question, including methods, units of measure

_____ How to calculate mean and median

_____ How to find IQR (interquartile range) and MAD (mean absolute deviation)

_____ To explain outliers in data

••

Math 6/7
Unit 6 Calendar

1/6	1/7	1/8	1/9	1/10
Computer Lab: Unit 6 Pre-Test MSG, Statistical Questions	Mean, Median, Mode, Range	Measures of Center, Spread, and Shape	Dot Plots (Line Plots)	QUIZ
1/13	**1/14**	**1/15**	**1/16**	**1/17**
Frequency Tables & Histograms	Box Plots and IQR	Box Plots	Stem & Leaf Plots	Review
1/20	**1/21**	**1/22**	**1/23**	**1/24**
MLK Holiday	Computer Lab; Unit 6 Test	MAD	Review Activity OR Work on Project	Project Due

Unit 6 Vocabulary

Vocabulary Term	Definition	Picture/ Example
Box-and-Whisker plot (or "box plot")	A graph that uses a number line to show how data is distributed. It shows the minimum, lower quartile, median, upper quartile, and maximum values of the data.	
Dot plot (line plot)	A graph in which each value is shown as a dot above a number line	
Frequency	The number of times a value appears in a set of data	
Frequency Table	A table that lists the number of times (frequency) that a piece of data occurs. This table is often used as a method of recording data.	
Gap	A "hole" in the data, where no data is present	
Histogram	A bar graph used to display numerical data grouped in equal intervals.	
Interquartile range (IQR)	The range between the upper and lower quartiles. This represents the middle 50% of the data.	
Maximum value	The largest value in a set of data	
Mean	The "average" of a set of data, found by adding all values in a set of data and dividing by the number of values you added.	

Vocabulary Term	Definition	Picture/ Example
Measures of center	A number that describes the middle of the data.	
Measures of spread (or variation)	A number that describes how spread out the data is. (how much the data varies)	
Median	The middle number when a set of data is arranged in order.	
Minimum value	The smallest number in a set of data, and the endpoint of one whisker on a box plot.	
Mode	The number that occurs the most in a set of data.	
Outlier	A value that is far away from most other values in a data set.	
Peak	The highest point on a graph, where most data is present (also the mode)	
Range	The difference between the highest and lowest values in a set of data.	
Skew	This describes the "direction" or "side" that has the least data	
Stem-and-leaf plot	A graph that organizes numerical data by splitting each piece of data into a "stem" and a "leaf", which each represent place value.	

Unit 6 Study Guide

1) What are the measures of center? _____

_____

2) What are the measures of spread? _____

_____

3) How do you calculate the interquartile range (IQR)? _____

_____

4) What do you look for in the shape of data? _____

_____

5) Big Bob scored the following points at eight basketball games: {21, 24, 9, 11, 16, 7, 24} Calculate the following:

   a. Mean: _____ b. Median: _____
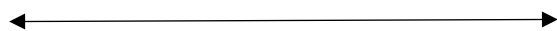
   c. Mode: _____ d. Range: _____

6) What is a statistical question? _____

_____

   Give an example: _____

_____

   Give a non-example: _____

_____

7) Use the following data to create a **box** plot:

Ages of Students Who Downloaded "Divergent"								
10	18	17	12	13	15	15	14	14

Minimum: _____
Q₁: _____
Median: _____
Q3: _____
Maximum: _____

8) Which measure of center is MOST affected by an outlier? ____

9) Identify the IQR from the box plot below: _____

50   53   56   59   62   65   68   71   74   77   80
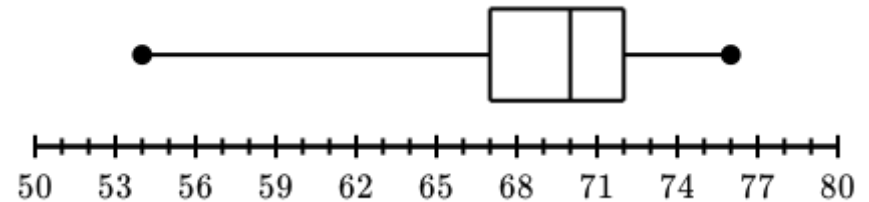
10) Use the box plot above to answer the following questions:

   a) Minimum: _____ b) Lower Quartile (Q₁): _____

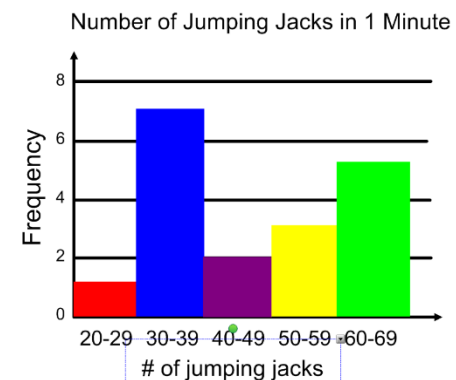   c) Median: _____ d) Uppoer Quartile (Q₃): _____

   e) Maximum: _____ f) Range: _____

   g) What percent of the data is 15 or greater? _____

   h) What percent of the data is between 15 and 25? _____

   i) The data is (Circle One): symmetrical   skewed right   skewed left

Use the Histogram below to answer questions 11-14.

Number of Jumping Jacks in 1 Minute

20-29 30-39 40-49 50-59 60-69
# of jumping jacks

11) How many students can do more than 49 jj per minute?_____

12) How many students participated in the survey?_____

13) Which interval shows the mode/peak of the histogram? _____

14) How many people did **EXACTLY** 45 jumping jacks? _____

15) At a car dealership, the salespeople sold the following numbers of cars during 2018.

28, 15, 35, 19, 22, 59, 23, 28, 19, 11

Determine the following:

a. mean = _____   b. median = _____   c. mode = _____

d. range = _____   e. IQR = _____   f. MAD = _____

g. outlier(s) = _____   h. minimum = _____   i. maximum = _____

16) Determine a set of data that has a mean of 12, a range of 10, and a median of 14.

17) Elisabeth's test scores are 72, 90, 94, 83, and 85. If she needs to maintain an 85 test average, what is the minimum score she needs on her next test?

18) Use the dot plot to find the following values:

Minutes To Eat Breakfast
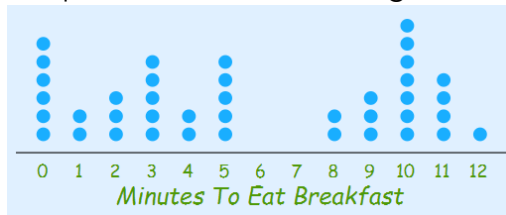
a. mean = _____   b. median = _____
c. mode = _____   d. range = _____
e. Describe at least 3 attributes about the SHAPE of the data.

19) Make a line plot for each set of data. Find the mean, median, mode, range, and any outliers of the data shown in the line plot.

52	48	52	51
52	65	58	48
60	45	50	52
56	48	53	58
62	49	51	49

Student Height in Inches
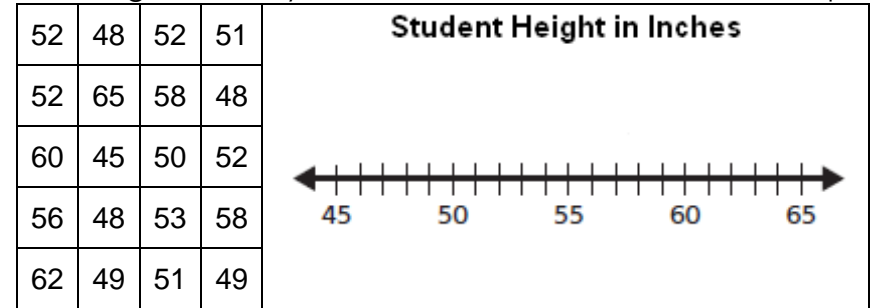
Mean: _____   Median: _____   Mode: _____

Range: _____   Outliers: _____

20) The table shows the daily soda sales for a restaurant. Choose intervals, make a frequency table, and construct a histogram to represent the data.

Number of Sodas Sold Daily					
56	86	74	63	51	94
86	72	53	77	74	88
81	90	72	76	84	92
78	89	85	75	91	87

Statistical Questions

A **STATISTICAL QUESTION** is a question that can have a **variety of answers**.

Examples	Non-Examples
How many books did my friends read this summer?	How many pages are in the Hunger Games?
How tall are my classmates?	How old am I?
During which months were my family members born?	During which month is Valentine's Day?
How many lunches did the cafeteria sell each day last month?	How many milkshakes did Chick-fil-A sell yesterday?

You Try! Check "Yes" or "No" to indicate whether or not each question is statistical.

Question:	Yes	No
1- What grades did the students in my class score on the test?		
2- How many marbles in the jar?		
3- What does this apple cost?		
4- How fast can each of our dogs run 100 yards?		
5- How old are each of the 6th grade teachers at East Cobb Middle School?		
6- How many days are in March?		
7- How many pets does each of my friends have?		
8- What did each of my classmates score on their last math test?		
9- Do you like peaches?		
10- What was the temperature at noon today in Marietta?		
11- What was the winning score in the last Super Bowl?		
12- How many songs has Shawn Mendes written?		
13- How many classmates binged Netflix over the break?		

Analyzing Data

CENTER
SPREAD
SHAPE

....
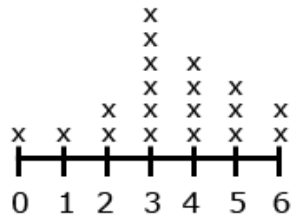
CENTER

A measure of _____ is a single number that

describes how data looks in the **middle**!

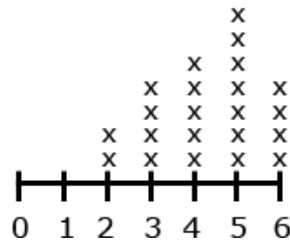Examples include **mean**, **median** and **mode**.

Graph A: Number of Siblings	Graph B: Number of Pets

Let's take a look at the two graphs above, and see what you

notice about their **CENTER**.

The **CENTER** of Graph A is about 3.

The **CENTER** of Graph B is about _____.

What does this tell you? _____

_____

_____

SPREAD

A measure of _____ (or variation) is a single

number that describes how far data is **spread out**.

Examples include **range** and **interquartile range (IQR)**
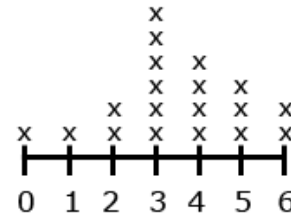
Graph A: Number of Siblings	Graph B: Number of Pets
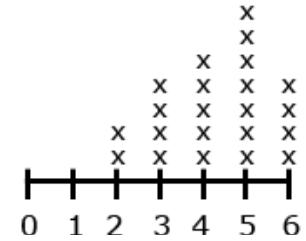
Now, look at the **SPREAD** in both graphs.

Which graph has data that is CLOSER together? _____

Which graph has data that is FARTHER apart? _____

What does this tell you? _____
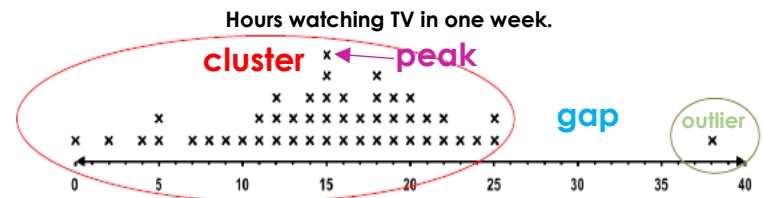
_____

SHAPE

The of a set of data gives a quick snapshot of its characteristics.

Look for: _____ (groups of data close together),

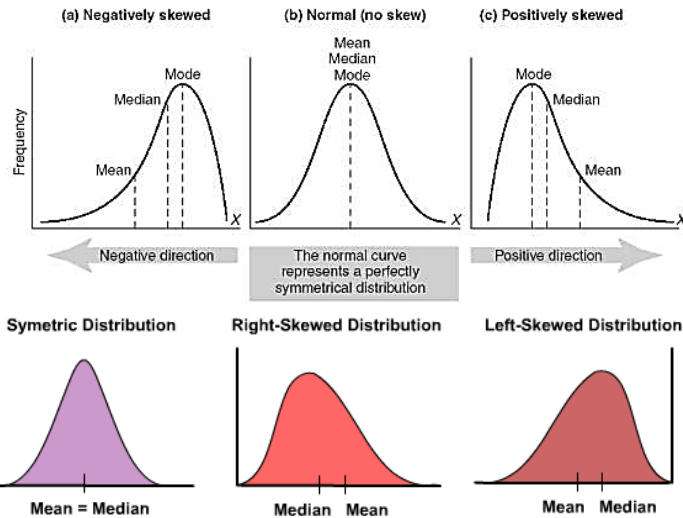_____ (a "hole" where no data is present),
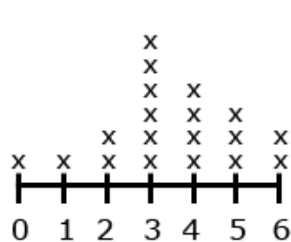
_____ (data that is far from the rest),

and _____ (the highest point on the graph – the

mode)

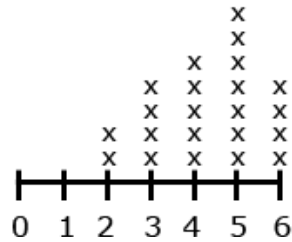**Hours watching TV in one week.**

cluster — peak

gap

outlier

Data can also be **skewed** based on the relationship between the Mean, Median and Mode. The way we describe the skew is based on the direction of the "tail".

(a) Negatively skewed   (b) Normal (no skew)   (c) Positively skewed

Negative direction   The normal curve represents a perfectly symmetrical distribution   Positive direction

Symetric Distribution   Right-Skewed Distribution   Left-Skewed Distribution

Mean = Median   Median Mean   Mean Median

Graph A:	Graph B:
Number of Siblings	Number of Pets

Now, look at the **SHAPE** of both graphs.

Which graph is skewed left? _____

Do either of the graphs have an outlier? _____

What is the peak of Graph B? _____

Does Graph A have any gaps? _____

Measures of Center: Mean, Median & Mode

Mean (a.k.a. "average")

To find the mean, first _____

Then, _____

Hint: *The mean is "mean" because it's the hardest to solve!*

Example: Find the mean of 6, 4, 10, 11 and 4.

$$6 + 4 + 10 + 11 + 4 = 35 \quad \rightarrow \quad 35 \div 5 = 7$$
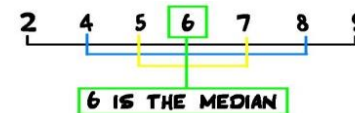
The mean of this data is **7**.

You Try: Find the mean of 8, 33, 20, 11, 6 and 12.

Median (a.k.a. "the middle")

The median is the _____ number when all data values are in _____. If there are two middle numbers, find their _____.

Example: Find the median of 2, 8, 9, 5, 6, 7, and 4.

2 4 5 6 7 8 9

6 IS THE MEDIAN

You Try: Find the median of 8, 33, 20, 11, 6, and 12.

Mode (a.k.a. "the most")

The mode is the number that occurs _____ in a set of data. You can have ____ _____ if all of the numbers in your data have the same frequency. You will have _____ than one mode if more than one number occurs most in a data set.

Example: Find the mode of 6, **4**, 10, 11, and **4**. Mode = **4**

You Try:

a. Find the mode of 8, 33, 20, 11, 6, and 12. _____

b. Find the mode of 1, 3, 4, 1, 5, 6, and 3. _____

c. Find the mode of 15, 62, 76, and 62. _____

You Try!

Find the mean, median and mode for the following data:

$$3, 5, 13, 6, 1, 2, 3, 2, 1$$

Mean: _____

Median: _____

Mode: _____

$$100, 111, 122, 133, 144, 155, 166$$

Mean: _____

Median: _____

Mode: _____

$$8.4, 14, 10.6, 2.4$$

Mean: _____

Median: _____

Mode: _____

Measures of S P R E A D (variation): Range & Interquartile Range (IQR)

Measures of Spread tell you how spread out your data is, or how much it varies.

Range: the difference between the _____ and _____ values in a data set (simply subtract the highest and lowest numbers!)
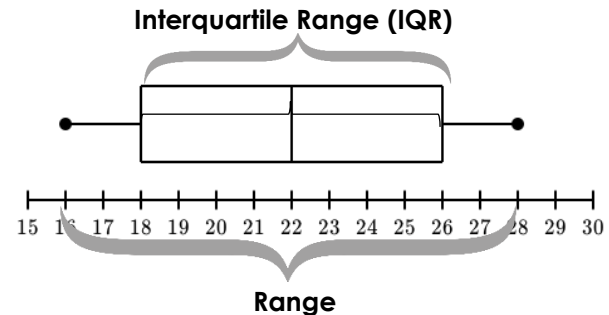
Maximum Value   Minimum Value

Example: 20, 13, 22, 17, (28) (10) 25   **Range** → 28 – 10 = 18

You Try: Find the range for the following data.

1) 34, 15, 9, 33, 27, 12, 27, 25, 30 _____

Interquartile Range (IQR): The range between the upper and lower quartiles on a box plot. This represents the middle 50% of the data. Simply subtract the upper and lower quartiles $Q_1 – Q_3$.

Interquartile Range (IQR)

Range

How to find the IQR:

1) Find the Median.
2) Find the medians of the lower (Q_1) and the higher (Q_3) group.
3) Subtract those two medians to get the IQR.

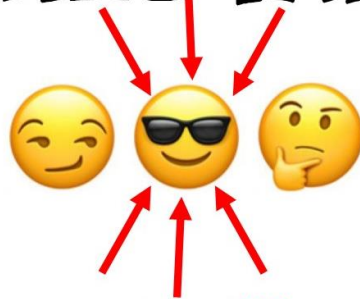**You Try:** What is the IQR in the box plot above?

Measures of Center and Spread Summary

Mean

The **MEAN** is the **average** of a set of numbers. You find it by finding the sum of a set of numbers and dividing by the number of numbers in the set.

MEDIAN

The **MEDIAN** is the **middle** of a set of numbers. You find it by putting your data in order and finding the one in the middle.

If there are two, you average the two middle numbers.

MODE

The **MODE** is the number that occurs **most** in a set of numbers. You only have a mode if one or more numbers are repeated in a data set.
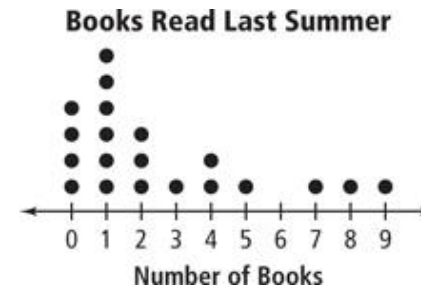
RANGE

The **RANGE** is the **difference** between the biggest (maximum value) and the smallest (minimum value) of numbers in a data set.

Dot Plots (Line Plots)

A dot plot (also called a line plot) is used to easily organize large sets of data. It is a graph in which each value is shown as a _____ (or and x) above a number line. Each dot (or x) represents _____ response.

Books Read Last Summer

1) How many people read 4 books last summer? _____

2) How many people read 1 book last summer? _____

3) How many people were surveyed? _____

4) Only 2 people read _____ books last summer.

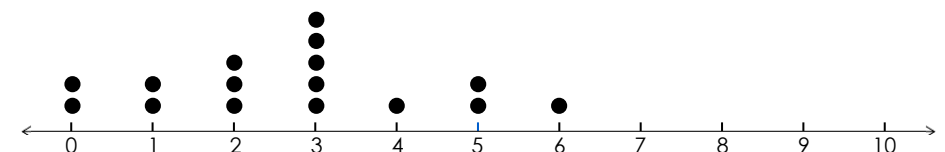## How to make a dot plot?

1) Draw a number line.
2) Mark off the minimum and maximum values and ALL numbers in between.
3) Make a dot (or x) for EACH data value above its number on the number line. Take care to make it neat and easy to read.
4) Title your dot plot.

Example:

Number of Jolly Ranchers eaten per day:

1, 2, 4, 3, 5, 2, 3, 6, 1, 0, 0, 5, 2, 3, 3, 3

Use the sets of data below to create dot plots and then use the data to find the mean, median, mode and range.

1) Number of siblings:

   3, 2, 0, 4, 1, 1, 1, 2, 1, 3, 5, 3, 4, 0, 2, 1, 0, 8

Mean: _____   Median: _____

Mode: _____   Range: _____

Outliers: _____   Min: _____ Max: _____
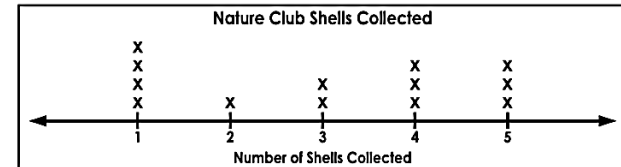
2) Number of downloaded apps on teens' cell phones:
   8, 12, 10, 15, 11, 20, 12, 12, 9, 10, 11, 13, 12, 9, 10, 13, 11, 13, 9, 12, 14, 9, 12, 15, 10, 11

Mean: _____   Median: _____

Mode: _____   Range: _____

Outliers: _____   Min: _____ Max: _____

Interpreting Dot Plots (Line Plots)

Use the data in the dot plot to answer questions 1-4.

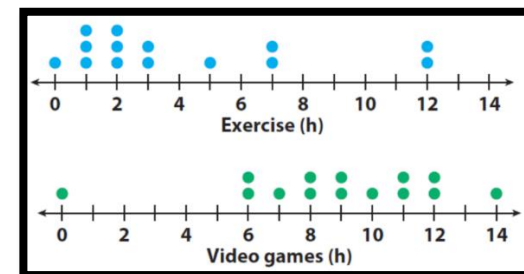Nature Club Shells Collected — Number of Shells Collected

1) What is the mean number of shells collected? _____

2) What is the median number of shells collected? _____

3) What is the mode? _____

4) What is the range? _____

Exercise (h) / Video games (h)

Fourteen students were surveyed about the time they spend exercising and playing video games each week. Compare the data by answering the questions 5-8.

5) What is the **range** for the hours of exercise? _____

   For playing video games? _____

6) What is the **mode** for exercise? _____

   Playing video games? _____

7) What is the **median** hours spent exercising? _____

   Playing video games? _____

8) What is the **mean** number of hours spent exercising? _____

   Playing video games? _____

Frequency Tables

A _____ _____ displays data that has been collected.

Season Soccer Scores

Score	Tally	Frequency
1	/	1
2	/	1
3	///	3
4	/	1
5	////	4

Intervals & Frequency Tables

Number of Cups of Coffee

Intervals	Tally	Frequency
0 – 3	//	2
4 – 7	///	3
8 – 11	///// ///	8
12 – 15	///	3
16 – 19	//	2

Intervals must be:

1) _____

2) _____

3) _____

You Try: If your data ranges from 2 to 38, are the intervals below good (👍) or bad (👎)

1) 1-10, 11-20, 21-30, 31-40   👍   👎

2) 1-10, 10-20, 20-30, 30-40   👍   👎

3) 1-10, 11-15, 16-35, 36-40   👍   👎

4) 1-8, 9-16, 17-24, 25-32, 33-40   👍   👎
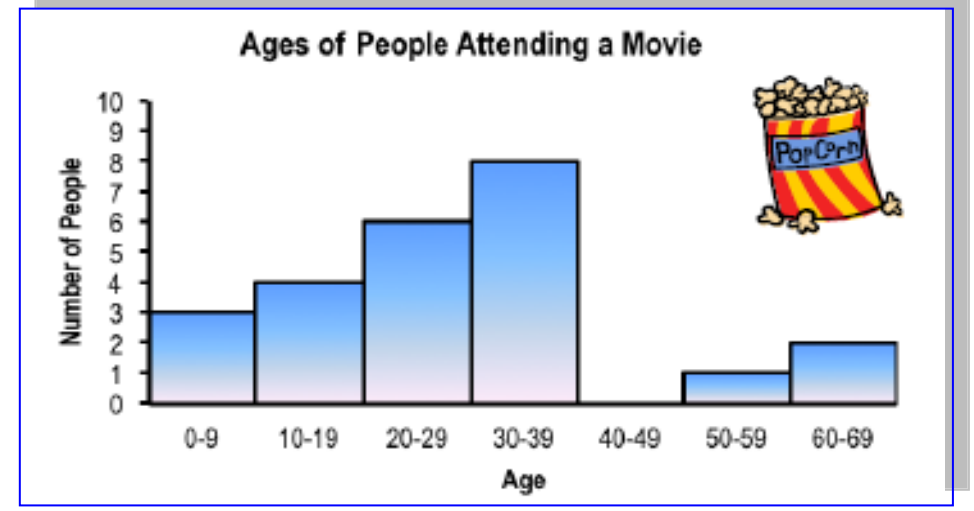
5) 1-10, 11-20, 21-30   👍   👎
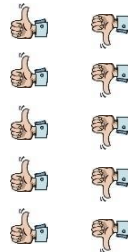
Histograms

A _____ is a bar graph used to display numerical data grouped in equal intervals.

Example:

The students of Monster High took a survey of the ages of everyone attending the "Ghouls Rule" Movie. The results are displayed in the histogram below.

1) How many people from ages 10-19 attended the movie? _____

2) How many people aged 50 or over attended the movie? _____

3) How many kids younger than 20 attended the movie? _____

4) How many total people attended the movie? _____

5) What does the gap at the interval 40-49 mean? _____

_____

6) Can you tell whether a 25-year-old attended the movie? _____

Why or why not? _____

7) Why must the bars on a histogram always be touching (unless there is a gap in data)? _____

Making a Histogram

Determining Intervals

Look at your data. What is the best way to break that data up?

Examples:

Data Range	Scale	Intervals
3 to 46	0 – 50	0-10, 11-20, 21-30, 31-40, 41-50
1 to 248	0 – 300	0-50, 51-100, 101-150, 151-200, 201-250
4.1 to 5.4	4 – 5.5	4-4.2, 4.3-4.5, 4.6-4.8, 4.9-5.1, 5.2-5.4
52 to 964		

Organize the data in a _____ _____ using the intervals.

Example:

Pages Read per Student Last Weekend				
78	15	40	19	188
50	122	96	37	102

The data ranges from _____ to _____. The scale will go from _____ to _____. We can use the interval of _____.

Make a frequency table:

Pages Read per Student Last Weekend				
NUMBER:	1-50	51-100	101-150	151-200
TALLY:	𝓣𝓗𝓛	�II	�II	�I
FREQUENCY:	5	2	2	1

Use the information in the frequency table on the previous page to create a histogram for the data.

Pages Read per Student Last Weekend				
NUMBER:	1-50	51-100	101-150	151-200
TALLY:	𝓣𝓗𝓛	�II	�II	�I
FREQUENCY:	5	2	2	1

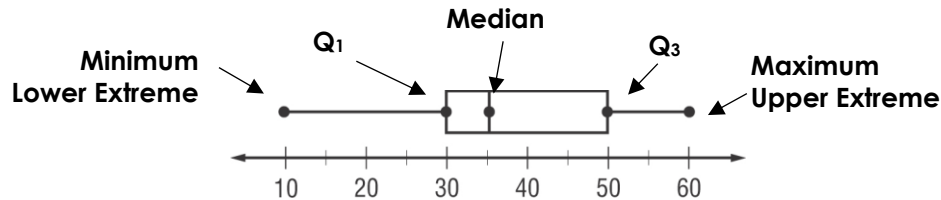Title: _____

y

x

_____

Remember: Bars must _____. Label both _____.

pg.13a

pg. 13b

Box Plot (Box & Whiskers Plot)

A box plot (a.k.a. "box and whisker plot") uses a _____ line to show how data is distributed. It shows the _____, the maximum and _____ values, which are also called the upper extreme and lower _____, and the upper and _____ quartiles.
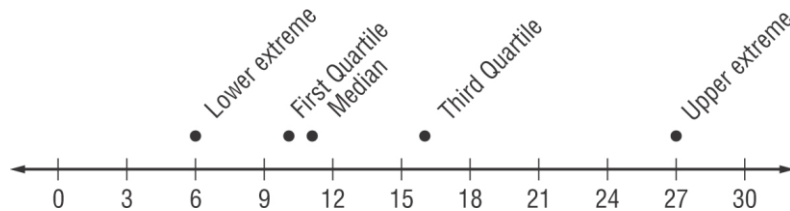

Example:

The list below shows the number of model airplanes owned by the members of the aviation club. Draw a box plot of the data.

6, 8, 10, 10, 10, 11, 12, 14, 16, 18, 27

1) Order the numbers from least to greatest. Then draw a number line that covers the range of the data.

2) Find the median, the extremes, and the first and third quartiles. Mark these points above the number line.

3) Draw the box so that it includes the quartile values. Draw a vertical line through the median value. Extend the whiskers from each quartile to the extreme data points.

4) Title your box plot!

Numbers of Model Airplanes Owned
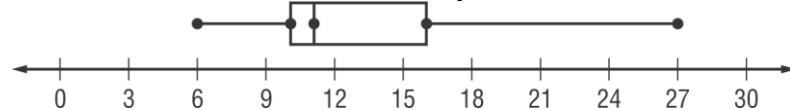
Tip:

One key understanding about box plots is that each section represents 25% of the data. If one section is large, that tells you that the numbers in that section are more spread out. If the section is small, that tells you the data is closer together.
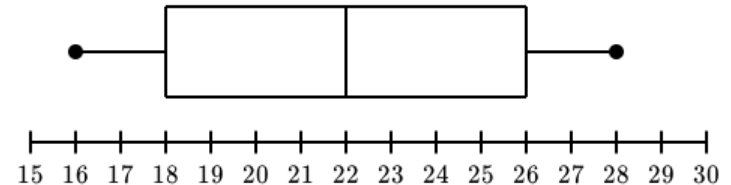
In the box plot below, you can see that the whisker for the upper quartile is much smaller than the whisker for the lower quartile. However, they both have the same number of data values. What does this tell you?
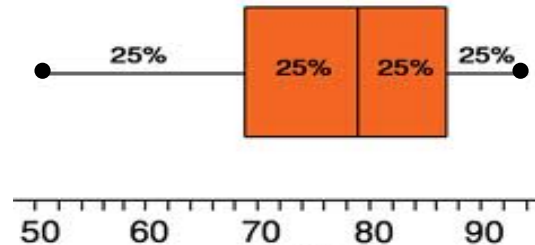
You Try:

1) Use the box and whiskers plot to answer the following questions:

Student Cell Phones per 6th Grade Class

a) What is the lowest number of students with cell phones in the data? _____

b) What is the highest number of students with cell phones in the data? _____

c) What is the median number of students with cell phones? _____

d) What is the range of students with cell phones? _____

2) Use the box and whiskers plot to answer the following questions:

Total Points Scored by Basketball Players in 2018


```
+--+--+--+--+--+--+--+--+--+--+
0   2   4   6   8  10  12  14  16  18  20
```
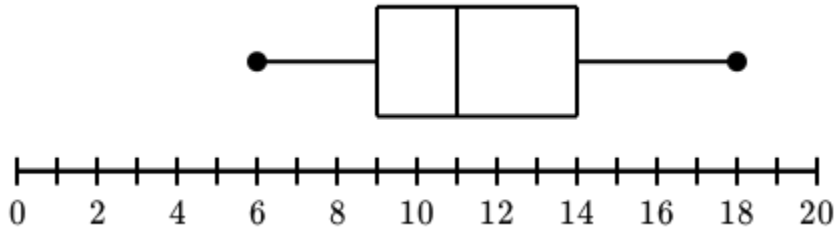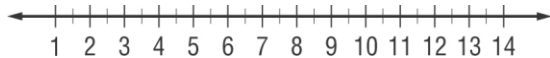
a) Median = _____   b) Lower Quartile (Q1) = _____

c) Maximum = _____   d) Minimum = _____

e) Range = _____   f) Upper Quartile (Q3) = _____

Use the data given for each problem to find the requested information and make a box plot.

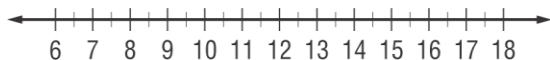1) The number of pencils students have at school:
   {4, 7, 5, 3, 12, 6, 5}

Minimum: _____
Q_1: _____
Median: _____
Q3: _____
Maximum: _____

```
1  2  3  4  5  6  7  8  9  10 11 12 13 14
```

2) Number of books read by the 6th grade teachers:
   {13, 8, 17, 10, 6, 11, 18}

Minimum: _____
Q_1: _____
Median: _____
Q3: _____
Maximum: _____

```
6  7  8  9  10 11 12 13 14 15 16 17 18
```

Pg.15a

3) The heights of students on the soccer team, in inches, are: 56, 69, 60, 64,63, 68, 68 and 66. Make a box plot for this data.

Minimum: _____
Q_1: _____
Median: _____
Q3: _____
Maximum: _____

4) The Young Fashionistas Club tallied up the total pairs of shoes that each member owns. Make a box plot of this data:
   5, 6, 7, 7, 7, 8, 9, 9, 11, 11, 12, 12, 12, 12, 12,
   13, 13, 14, 14, 14, 14, 18, 19, 20, 20

Minimum: _____
Q_1: _____
Median: _____
Q3: _____
Maximum: _____

Mean Absolute Deviation (MAD)

Mean absolute deviation, or **"MAD"**, is the average distance of all data points from the mean. It is a way of looking at variability (spread) in a data set.

STEPS:

1 – _____

2 – _____

_____

3 – _____

_____

Example: Find the MAD of guests' ages at twins Solomon's and Sofia's party: 4, 6, 5, 6, 6, and 9.

Step 1: Find the mean: $\dfrac{4 + 6 + 5 + 6 + 6 + 9}{6} = \dfrac{36}{6} = 6$

Step 2: Find the distance of each data point from the mean.

Age	4	6	5	6	6	9
Distance from mean	2	0	1	0	0	3

Step 3: Find the mean of the distances:
$$\dfrac{2 + 0 + 1 + 0 + 0 + 3}{6} = \dfrac{6}{6} = 1$$

The **MAD**, or average distance from the mean, is **1**.

Let's show "MAD" who's boss!!

1) Number of pets owned by students in Spanish Club: 5, 6, 3, 3, 4, and 3
.

   Step 1: Find the mean.

   Step 2: Find the distance of each piece of data from the mean.

   Step 3: Find the mean of the distances.

   The MAD is_____.

1) Find the MAD of the number of pets owned by students in the Young Vets Club: 8, 2, 3, 1, 8, 1, 0, and 9

2) Notice that both problems above had the same mean of _____. But, the MAD for problem 1 is _____ and the MAD for problem 2 is ____. What does this tell you about the data?