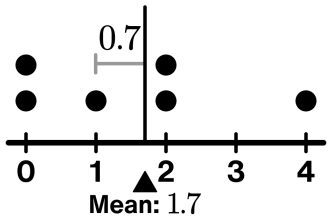
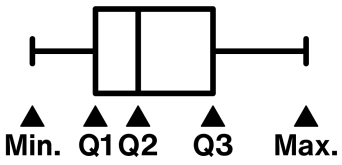
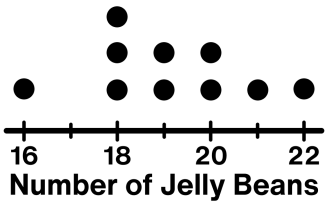
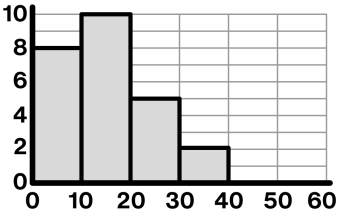
Describing Data
Student Guide

Math 6 Unit 8 Accelerated
Part 1

desmos

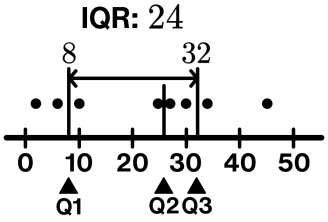
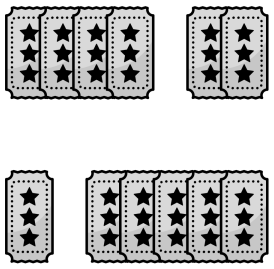
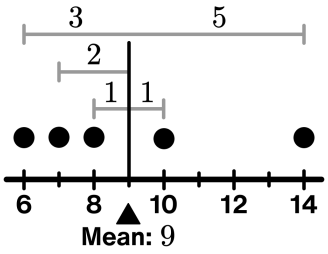
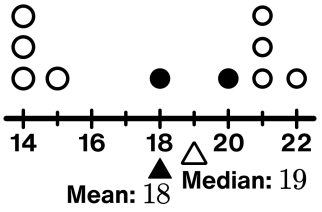
Unit 6.8, Student Goals and Glossary

Glossary

Term	Definition
absolute deviation (from the mean)	Absolute deviation is the distance between a value and the mean of a data set. If a class's mean number of pets is 1.7, then the absolute deviation of a student who has 1 pet is 0.7. 
box plot	A box plot is one way to visualize numerical data. The data is divided into four sections using five numbers: the minimum, Q1, Q2 (or the median), Q3, and the maximum. The box is drawn between Q1 and Q3, and the line inside the box represents the median. 
categorical data	Categorical data has values that are words instead of numbers. *What kind of pet do you have?* is a question that asks for categorical data.
dot plot	A dot plot is one way to visualize data. Each data point is shown as a dot above its value, stacking on top of other dots with the same value. For example, this dot plot shows that 3 students guessed that there were 18 jelly beans in a jar. 
histogram	A histogram is one way to visualize numerical data. The data in a histogram is grouped into bins each shown by a rectangle. The height of each rectangle shows how many values are in that bin. For example, this histogram shows that there are 8 values between 0 and 10. 

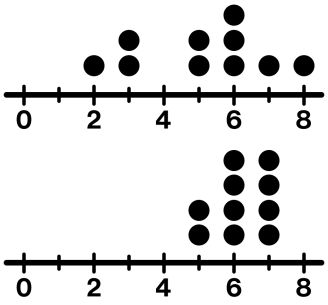
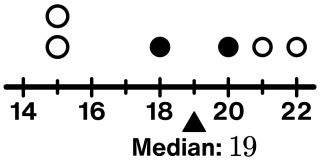
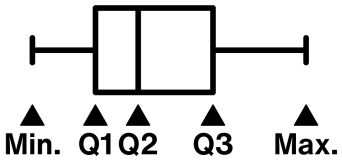
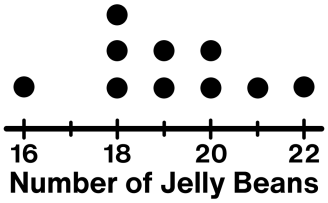
desmos

Unit 6.8, Student Goals and Glossary

interquartile range (IQR)	Interquartile range (or IQR) is a measure of spread. It is the distance from Q1 to Q3 and the width of the box in a box plot. For example, the IQR of this data set is $32 - 8 = 24$.	
mean	The mean or average is a measure of center. The mean is the number of items in each group if the items are distributed equally or the balance point of a dot plot. To calculate the mean, you can add up all the data values, and divide by the number of data points. In this situation, the mean is 3 tickets.	 $4 + 2 + 1 + 5 = 12$ $12 \div 4 = 3$
mean absolute deviation (MAD)	The mean absolute deviation (or MAD) is one way to measure how spread out a data set is. It is the average of all of the absolute deviations of the points in a data set. To calculate the MAD, determine the distance between each data point and the mean, then calculate the mean of those distances. In this example, the MAD is 2.4 because $3 + 2 + 1 + 1 + 5 = 12$ and $12 \div 5 = 2.4$.	
measure of center	A measure of center is a single number that summarizes all of its values. It is usually a typical value for a data set. Mean and median are measures of center.	

desmos

Unit 6.8, Student Goals and Glossary

measure of spread	A measure of spread tells us how bunched up or spread out the values in a data set are. Range, interquartile range, and mean absolute deviation are measures of spread. For example, the dot plot on the top has a larger spread than the dot plot on the bottom. 
median	Median is a measure of center. It is the middle value of a data set when the values are in numerical order. If there are two values in the middle of the data set, then the median is the mean of those two values. 
numerical data	Numerical data has values that are numbers and can be measured. *How many pets do you have?* is a question that asks for numerical data.
quartile	Quartiles divide a data set into four sections. Quartile 1 is the median of the lower half of the data. Q2 is also the median. Q3 is the median of the upper half of the data. Q4 is also the maximum. 
range	Range is a measure of spread. It is the difference between the maximum and minimum values in a data set. For example, the range of this data set is 6 jelly beans because $22 - 16 = 6$. 
statistic	A statistic is a single number that measures something about a data set. Examples of statistics: mean, median, MAD and IQR.
statistical question	A statistical question requires more than one piece of data to answer it. Here are some examples of statistical questions: • *What is the most popular band at your school?* • *When do students in your class typically eat dinner?*

desmos

Unit 6.8, Family Resource

Unit 8 Summary

Prior Learning	Math 6, Unit 8	Future Learning
Grades 3–5 • Fractions and decimals on a number line • Visualizing data using line plots • Calculating distances on a number line Math 6, Unit 3 • Calculating percentages	• Visualizing data • Measuring data: mean and MAD • Measuring data: median and IQR	Math 7, Unit 8 • Probability and sampling data Math 8, Unit 6 • Associations in bivariate data High School • Standard deviation and outliers

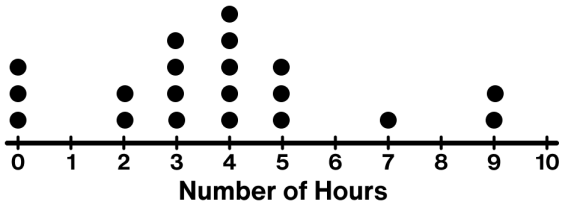
Visualizing Data

Asking questions and collecting data can help us make claims about a group.

Visualizing the data we collected can help us interpret the responses.

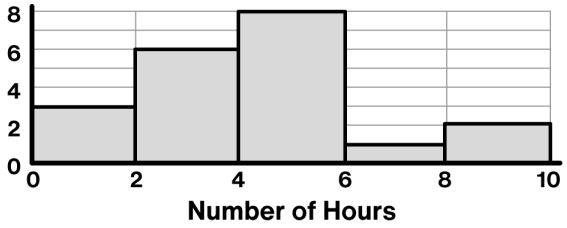
This *dot plot* and *histogram* show the number of hours a day that 20 adults spend on their phone.

Dot Plot



In a dot plot, each data point is represented by a dot above the number line.

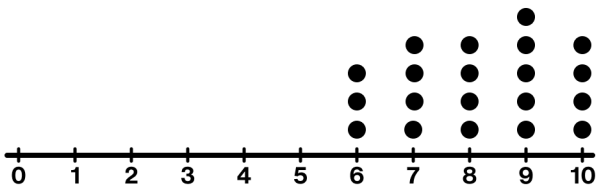
Histogram

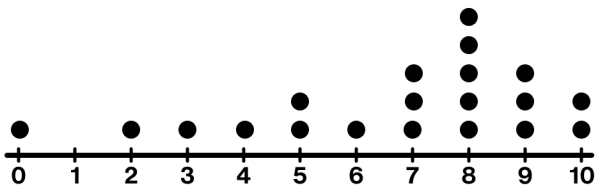


In a histogram, the data is grouped into bins. The height of each rectangle shows how many data points are in that bin.[1]

Visualizing the data can also help us describe its *shape, center,* and *spread*.

For example, the *centers* of these data sets are around 8 and the *spreads* are different.





[1]In this unit, data on the edges, such as 2, are sorted into the bin immediately to the right of it.

desmos

Unit 6.8, Family Resource

This data set has a smaller spread.

This data set has a larger spread.

Mean and MAD

One way to measure the center of a data set is the *mean*, or the average.

The mean can be thought of as the equal share.

For example, the mean is the number of stickers five friends would get if they shared them equally.

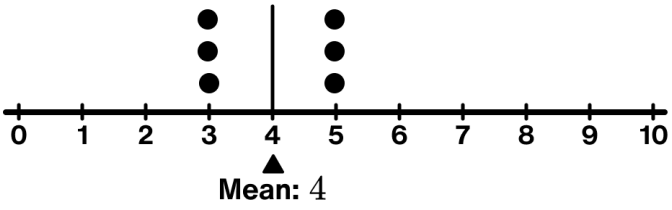
To calculate the mean, add the data and divide the total by the number of data points.

The mean of $7, 8, 10, 7,$ and 8 is:

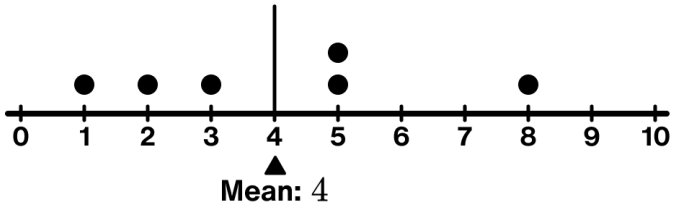
$$\frac{7+8+10+7+8}{5} = \frac{40}{5} \text{ or } 8$$

One way to measure the spread of a data set is the *mean absolute deviation (MAD)*.

The MAD is how far away the data is from the mean on average. The higher the MAD, the more spread out the data.



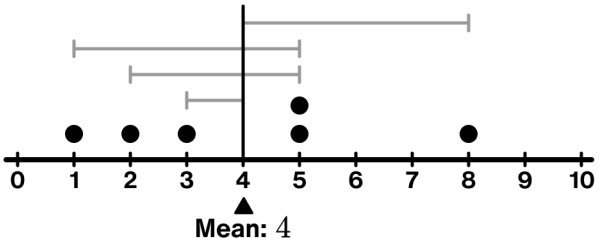
The MAD of this data set is 1.



The MAD of this data set is 2.

To calculate the MAD, first measure the distances between each data point and the mean (these are called *absolute deviations*). Then, calculate the mean of the absolute deviations.

This table shows the distances from each point to the mean.



Data Point	1	2	3	5	5	8
Absolute Deviation (distance from 4)	3	2	1	1	1	4

The MAD of this data is $\frac{3+2+1+1+1+4}{6} = \frac{12}{6}$ or 2.

desmos

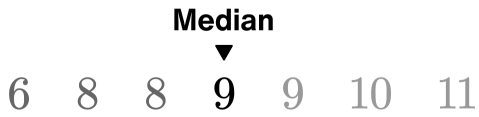
Unit 6.8, Family Resource

Median and IQR

The center of a data set can also be measured by the *median*.

The median is the middle value of a data set when the values are listed in order.

A student kept track of the number of hours they slept each night for a week. The median of this data is 9 hours because it is in the middle of the list.



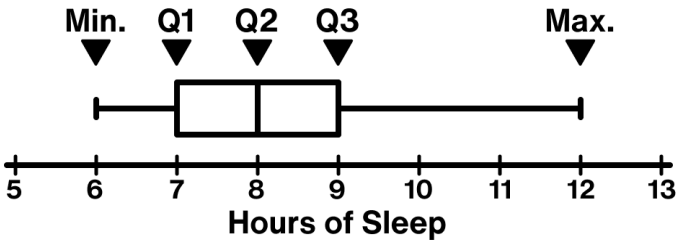
Quartiles (Q1, Q2, Q3) divide a data set into four sections.

- Quartile 1 is the median of the lower half of the data.
- Quartile 2 is the median of the entire data.
- Quartile 3 is the median of the upper half of the data.

This data set shows the number of hours 15 students slept on a school night. The first, second, and third quartiles are labeled.

The quartiles, along with the minimum and maximum values, can be used to create a *box plot*.

This box plot visualizes the number of hours each student slept on a school night.



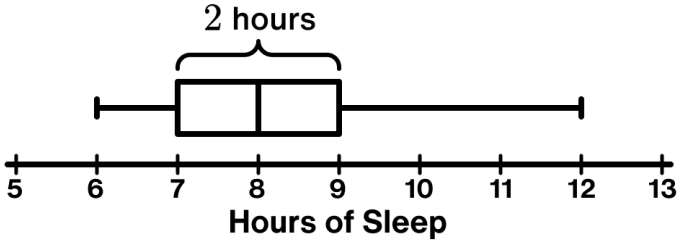
The spread of a data set can also be measured by the *interquartile range (IQR)*.

The IQR is the difference between Q1 and Q3.

It is where the middle half of the data lies.

The IQR of this data is 2 because $9 - 7 = 2$.

The middle half of the data lies within 2 hours.



desmos

Unit 6.8, Family Resource

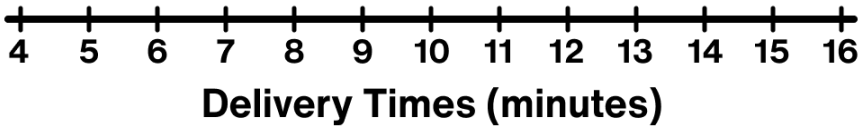
Try This at Home

Visualizing Data

The owner of a pizza shop wanted to know more about how long it took to deliver their pizzas. One day, they recorded the time, in minutes, of 10 pizza deliveries. They organized their data into a table.

5	7	10	16	9	12	9	10	11	9

1.1 Create a dot plot of the delivery times.



1.2 Which statement best describes the data set?

A. The center is around 3 and the spread is small.

B. The center is around 3 and the spread is large.

C. The center is around 9 and the spread is small.

D. The center is around 9 and the spread is large.

This histogram shows the delivery times for a restaurant in a day.

2.1 Dylan says that there were 5 deliveries that day.

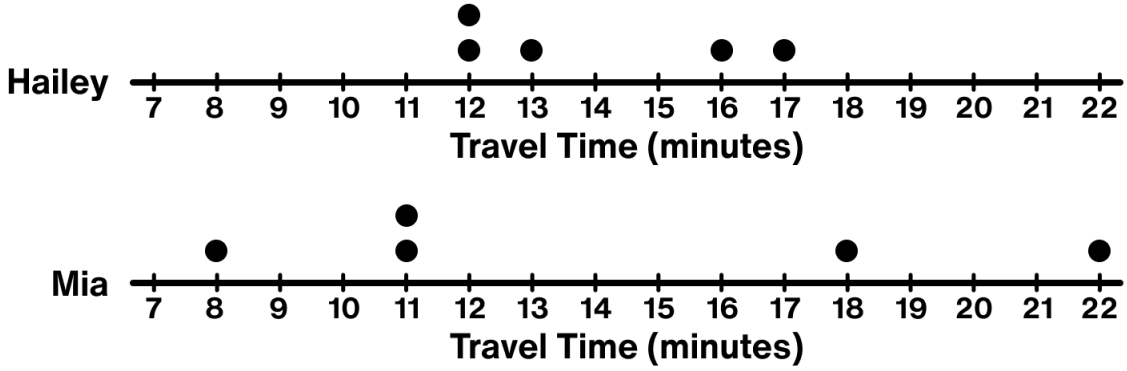
Do you agree with Dylan?

2.2 How many deliveries were made in less than 10 minutes?

desmos

Unit 6.8, Family Resource

Hailey and Mia are curious about how long it takes them to travel to school. For one week, they decide to record their travel times. The dot plots show their data from the week.



Mean and MAD

3.1 What is the mean of Hailey's travel times?

3.2 What is the mean of Mia's travel times?

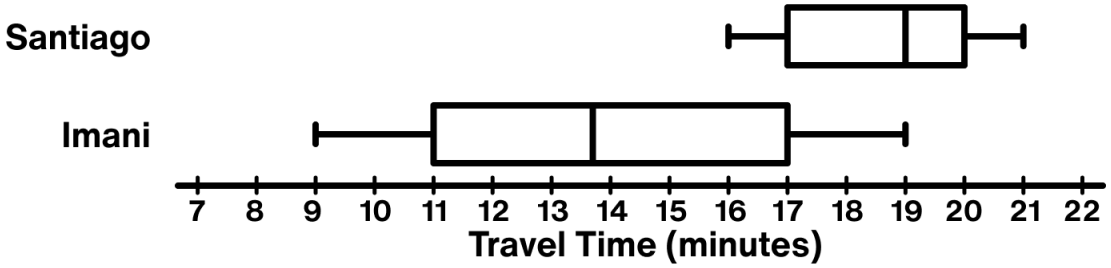
3.3 Without calculating, whose data set has a higher MAD? Explain your thinking.

Median and IQR

4.1 What is the median of Hailey's travel times?

4.2 What is the median of Mia's travel times?

Two new students recorded their travel times and visualized their data as box plots.



5.1 Label the first, second, and third quartiles of Santiago's box plot with Q1, Q2, and Q3.

5.2 What is the IQR of Santiago's data?

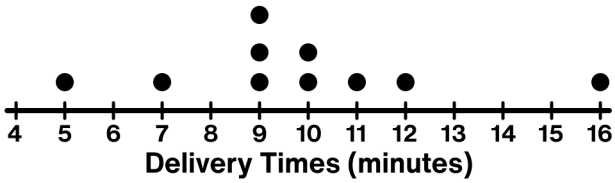
5.3 Who had a more consistent travel time to school? How do you know?

desmos

Unit 6.8, Family Resource

Solutions:

1.1



1.2 **D**

2.1 Disagree. *Explanations vary*. Dylan probably counted the number of bins. There were $2 + 6 + 8 + 5 + 1$ or 22 deliveries that day.

2.2 8 deliveries

3.1 14 minutes

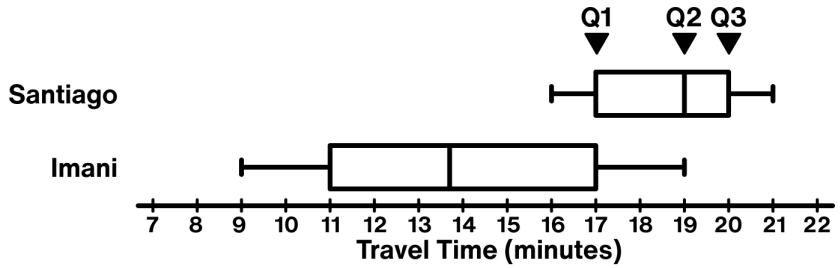
3.2 14 minutes

3.3 Mia. *Explanations vary*. Mia's data is more spread out.

4.1 13 minutes

4.2 11 minutes

5.1



5.2 3 minutes

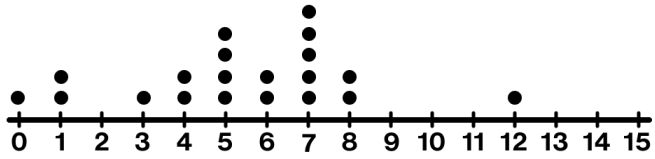
5.3 Santiago. *Explanations vary*. Santiago's IQR is smaller than Imani's, which means that Santiago's data is closer together.

desmos

Unit 6.8, Lesson 2: Notes Name _____

My Notes

This dot plot shows the number of books that students in a sixth-grade class read in one month.



Number of Books Read This Month

1.1 There are 2 dots plotted at 4. What does this tell us about the situation?

1.2 Julian said that most students read 6 books or less this month. Do you agree with him? Explain your thinking.

A statistical question is a question that requires more than one piece of data to answer it.

Question A	Question B
What is the typical height of a building in NYC?	How tall is the Empire State Building?

2. Explain why Question A is a statistical question but Question B is not.

Summary

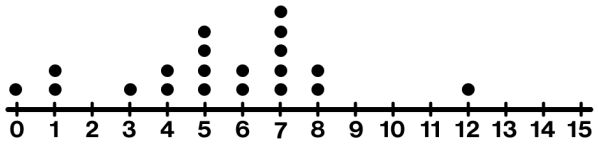
- ☐ I can describe and interpret a dot plot to help answer a statistical question.
- ☐ I can describe the advantages and disadvantages of using a dot plot to visualize data.

desmos

Unit 6.8, Lesson 2: Notes Name _____

My Notes

This dot plot shows the number of books that students in a sixth-grade class read in one month.



Number of Books Read This Month

1.1 There are 2 dots plotted at 4. What does this tell us about the situation?

This tells us that there are 2 students in the class who read 4 books this month.

1.2 Julian said that most students read 6 books or less this month. Do you agree with him? Explain your thinking.

I agree with Julian because 12 out of the 20 students read 6 books or less, and that is more than half.

A statistical question is a question that requires more than one piece of data to answer it.

Question A	Question B
What is the typical height of a building in NYC?	How tall is the Empire State Building?

2. Explain why Question A is a statistical question but Question B is not.

Question A is a statistical question because you will have to collect the heights of many buildings to determine what a typical height is. Question B is not statistical because you only need the height of one building to answer it.

Summary

- ☐ I can describe and interpret a dot plot to help answer a statistical question.
- ☐ I can describe the advantages and disadvantages of using a dot plot to visualize data.

desmos

Unit 6.8, Lesson 5: Notes Name _____

My Notes

Here is a histogram of minimum wages in the western United States.

1.1 How many states have a minimum wage **less** than $\$10.00$?

1.2 Michigan has a minimum wage of $\$9.65$.
Which bin should it go into? Put a star on that rectangle.

1.3 Adriana says that this histogram represents the minimum wages of 4 different states.

What advice would you give Adriana to better understand the histogram?

Summary

- ☐ I can describe and interpret a histogram that represents a data set.
- ☐ I can compare and contrast dot plots and histograms to visualize data.

desmos

Unit 6.8, Lesson 5: Notes Name _____

My Notes

Here is a histogram of minimum wages in the western United States.

1.1 How many states have a minimum wage **less** than $\$10.00$?

15 **states**

1.2 Michigan has a minimum wage of $\$9.65$.
Which bin should it go into? Put a star on that rectangle.

$\$8.00$ **to less than** $\$10.00$ **(rectangle with the star)**

1.3 Adriana says that this histogram represents the minimum wages of 4 different states.

What advice would you give Adriana to better understand the histogram?

The 4 rectangles do not each represent a state, they represent bins that organize the data. The height of each rectangle tells you how many states are in each bin.

Summary

- ☐ I can describe and interpret a histogram that represents a data set.
- ☐ I can compare and contrast dot plots and histograms to visualize data.

desmos

Unit 6.8, Lesson 7: Notes Name _____

My Notes

1. In your own words, explain what the *mean* tells you about a data set.

Six friends played games together at the arcade.
Here are the number of tickets that each friend won.

7	3	4	6	8	2

2.1 Calculate the mean number of tickets for this data. Show your calculations.

2.2 What does the mean tell us about this situation?

3. Describe how to determine the mean of any data set.

Summary

- ☐ I can describe what the *mean* of a data set is.
- ☐ I can calculate the mean of a data set.

desmos

Unit 6.8, Lesson 7: Notes Name _____

My Notes

1. In your own words, explain what the *mean* tells you about a data set.

   The *mean* is like an average. It tells you how much data would be in each group if all of the data were shared equally.

Six friends played games together at the arcade.
Here are the number of tickets that each friend won.

7	3	4	6	8	2

2.1 Calculate the mean number of tickets for this data.
Show your calculations.

$$\frac{7+3+4+6+8+2}{6} = 5$$

The mean is 5 tickets

2.2 What does the mean tell us about this situation?

The mean in this situation tells us how many tickets each friend would get if they shared all of the tickets equally.

3. Describe how to determine the mean of any data set.

To find the mean of any data set, you should add all of the data together, then divide by the total number of data points.

Summary

- ☐ I can describe what the *mean* of a data set is.
- ☐ I can calculate the mean of a data set.

desmos

Unit 6.8, Lesson 9: Notes Name _____

My Notes

1. Does the *mean absolute deviation (MAD)* tell us about the **center** of a data set or the **spread** of a data set.
   Circle your answer.

   Center Spread

Here is Marco's work for calculating the MAD of a data set.

Data Values	5	7	7	9	12
Absolute Deviation (distance from __8__) mean	3	1	1	1	4

2. Explain what Marco did to calculate the MAD.

3+1+1+1+4=10

10÷5=2

MAD: 2

3. Calculate the mean and MAD of this data set.
   Use the table to help you organize your thinking.

Data Values	1	4	5	6
Absolute Deviation (distance from ____) mean				

Mean: MAD:

Summary

- ☐ I can describe what the *mean absolute deviation* (MAD) is.
- ☐ I can calculate the mean absolute deviation of a data set.

desmos

Unit 6.8, Lesson 9: Notes Name _____

My Notes

1. Does the *mean absolute deviation* (MAD) tell us about the **center** of a data set or the **spread** of a data set.
   Circle your answer.

   Center | Spread

Here is Marco's work for calculating the MAD of a data set.

Data Values	5	7	7	9	12
Absolute Deviation (distance from __8__ mean)	3	1	1	1	4

2. Explain what Marco did to calculate the MAD.

   3+1+1+1+4=10

   10÷5=2

   MAD: 2

   Explanations vary. **Marco found the mean of the deviations. He added all of the deviations together and then divided by the number of data points.**

3. Calculate the mean and MAD of the data set.
   Use the table to help you organize your thinking.

Data Values	1	4	5	6
Absolute Deviation (distance from __4__ mean)	3	0	1	2

Mean:

$$\frac{1+4+5+6}{4} = 4$$

Mean: 4

MAD:

$$\frac{3+0+1+2}{4} = 1.5$$

MAD: 1.5

Summary

- ☐ I can describe what the *mean absolute deviation* (MAD) is.
- ☐ I can calculate the mean absolute deviation of a data set.

desmos

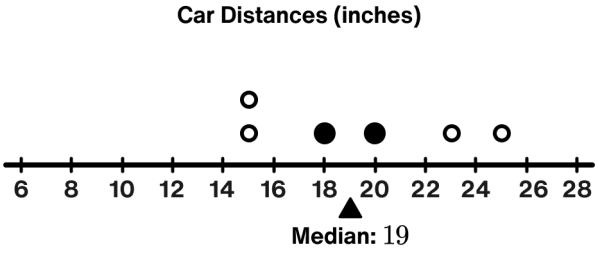
Unit 6.8, Lesson 11: Notes Name _____

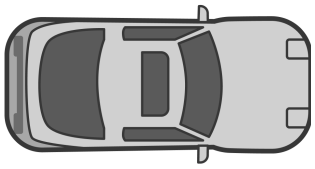
My Notes

1. In your own words, what is the *median* of a data set?

2. What is the median of this data set?

3. Explain why the median of Amoli's car distances is 19 inches.





4. Write your own set of at least 6 distances that have a median of 18 inches.

Summary

- ☐ I can describe what the median is.
- ☐ I can determine and interpret the median of a data set.

desmos

Unit 6.8, Lesson 11: Notes Name _____

My Notes

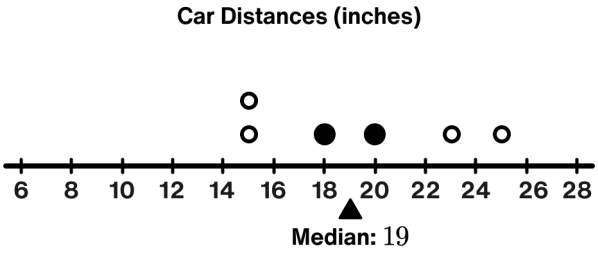
1. In your own words, what is the *median* of a data set?

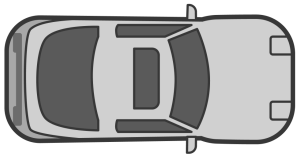
   ***Responses vary*. The median is the middle value of the data set when the values are written in order.**

2. What is the median of this data set?

   The median is 17**.**

3. Explain why the median of Amoli's car distances is 19 inches.





   ***Explanations vary*. When there's an even number of data points, the median is the average of the middle two numbers. The average of** 18 **and** 20 **is** 19**.**

4. Write your own set of at least 6 distances that have a median of 18 inches.

   Data sets vary. 5, 6, 9, 18, 19, 20, 21 **inches**

Summary

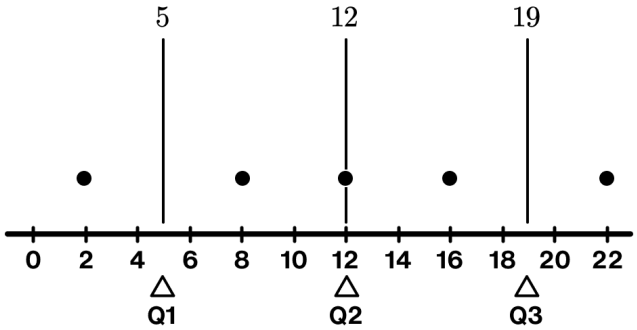
- ☐ I can describe what the median is.
- ☐ I can determine and interpret the median of a data set.

desmos

Unit 6.8, Lesson 13: Notes Name _____

My Notes

This dot plot shows the weight in pounds of 5 pumpkins.

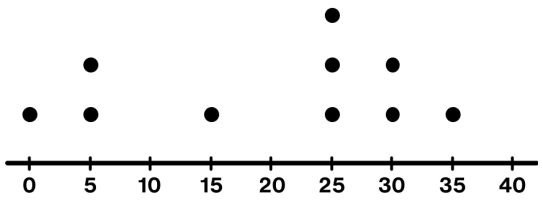


1.1 Explain why the median is 12.

1.2 Explain why Q1 is 5.

1.3 Explain why Q3 is 19.

2. Here is a data set with 10 data points. Determine the value of each of the following statistics.



Q1:

Median:

Q3:

Summary

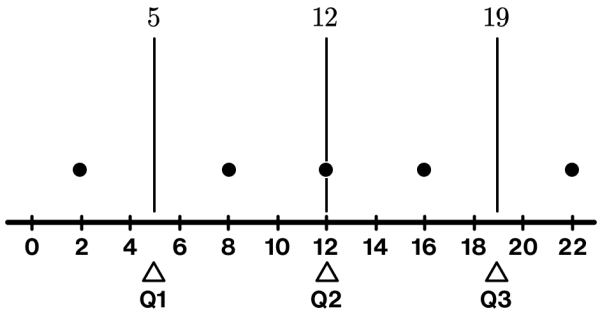
☐ I can determine and interpret the quartiles of a data set.

desmos

Unit 6.8, Lesson 13: Notes Name _____

My Notes

This dot plot shows the weight in pounds of 5 pumpkins.



1.1 Explain why the median is 12.

***Explanations vary*. The median is 12 because it is the middle value of the data set.**

1.2 Explain why Q1 is 5.

***Explanations vary*. The lower half of the data includes 2 and 8. The median of these numbers is 5.**

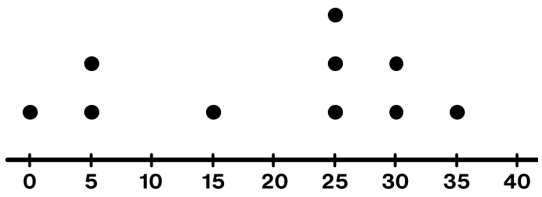
1.3 Explain why Q3 is 19.

***Explanations vary*. The upper half of the data includes 16 and 22. The median of these numbers is 19.**

2. Here is a data set with 10 data points. Determine the value of each of the following statistics.

Q1: 5

Median: 25



Q3: 30

Summary

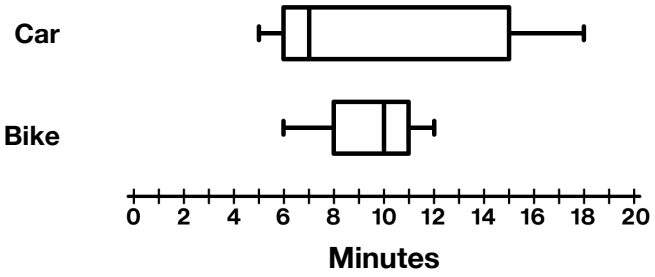
☐ I can determine and interpret the quartiles of a data set.

desmos

Unit 6.8, Lesson 14: Notes Name _____

My Notes

Here are box plots that show how long it takes for Jacy to get to school by car and by bike.



1. For each box plot, determine these statistics.

	Car		Bike
Q1		**Q1**	
Median		**Median**	
Q3		**Q3**	
Range		**Range**	
IQR		**IQR**	

2. If Jacy prefers a mode of transportation that's more predictable, should they go by car or by bike? Explain your reasoning.

Summary

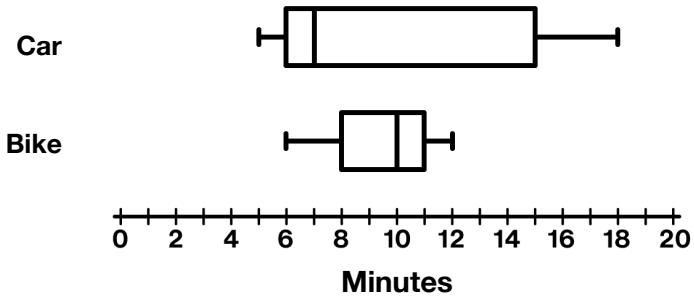
- ☐ I can create a box plot to visualize a data set.
- ☐ I can describe what *range* and *interquartile range (IQR)* are.
- ☐ I can determine the range and IQR of a data set.

desmos

Unit 6.8, Lesson 14: Notes Name _____

My Notes

Here are box plots that show times it takes for Jacy to get to school by car and by bike.



1. For each box plot, determine these statistics.

	Car		Bike
Q1	6 minutes	**Q1**	8 minutes
Median	7 minutes	**Median**	10 minutes
Q3	15 minutes	**Q3**	11 minutes
Range	13 minutes	**Range**	6 minutes
IQR	9 minutes	**IQR**	3 minutes

2. If Jacy prefers a mode of transportation that's more predictable, should he go by car or by bike? Explain your reasoning.

   By bike. *Explanations vary*. Jacy should go by bike because the range of times by car is a lot greater than the range of times by bike.

Summary

- ☐ I can create a box plot to visualize a data set.
- ☐ I can describe what *range* and *interquartile range (IQR)* are.
- ☐ I can determine the range and IQR of a data set.