NEW JERSEY CENTER
FOR TEACHING & LEARNING

Algebra I

Data & Statistical Analysis

2015-11-25

www.njctl.org

Table of Contents

Click on the topic to go to that section

Measures of Central Tendency: Mean, Median, Mode & Additional Measures of Data

Return to Table of Contents

Measures of Central Tendency Key Terms

Mean - The sum of the data values divided by the number of items; average

Median - The middle data value when the values are written in numerical order

Mode - The data value that occurs the most often

Other data related terms:

Minimum - The smallest value in a set of data

Maximum - The largest value in a set of data

Range - The difference between the greatest data value and the least data value

Outliers - Numbers that are significantly larger or much smaller than the rest of the data

Minimum and Maximum

14, 17, 9, 2, 4, 10, 5

What is the minimum in this set of data?

What is the maximum in this set of data?

Outliers

Outliers - Numbers that are relatively much larger or much smaller than the data

Which of the following data sets have outlier(s)?

A. 1, 13, 18, 22, 25

B. 17, 52, 63, 74, 79, 83, 120

C. 13, 15, 17, 21, 26, 29, 31

D. 25, 32, 35, 39, 40, 41

1 Which of the following data sets have outlier(s)?

A. 13, 18, 22, 25, 100

B. 17, 52, 63, 74, 79, 83

C. 13, 15, 17, 21, 26, 29, 31, 75

D. 1, 25, 32, 35, 39, 40, 41

2 The data set: 1, 20, 30, 40, 50, 60, 70 has an outlier which is _____ than the rest of the data.

- A higher
- B lower
- C neither

What is the Median?

Given the following set of data, what is the median?

10, 8, 9, 8, 5

Who remembers what to do when finding the median of an even set of numbers?

Find the Median

When finding the median of an even set of numbers, you must take the mean of the two middle numbers.

Find the median

12, 14, 8, 4, 9, 3

3 Find the median: 5, 9, 2, 6, 10, 4

- A 5
- B 5.5
- C 6
- D 7.5

4 Find the median: 15, 19, 12, 6, 100, 40, 50

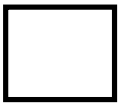
- A 15
- B 12
- C 19
- D 6

5 Find the median: 1, 2, 3, 4, 5, 6

- A 3 & 4
- B 3
- C 4
- D 3.5

What is the Range

Given a maximum of 17 and a minimum of 2, what is the range?



6 Find the range: 4, 2, 6, 5, 10, 9

- A 5
- B 8
- C 9
- D 10

7 Find the range, given a data set with a maximum value of 100 and a minimum value of 1.

8 Find the range for the given set of data:
13, 17, 12, 28, 35

Find the Mode

Find the mode

10, 8, 9, 8, 5

Find the mode

1, 2, 3, 4, 5

What can be added to the set of data above, so that there are two modes? Three modes?

9 What number can be added to the data set so that there are 2 modes: 3, 5, 7, 9, 11, 13, 15 ?

- A 3
- B 6
- C 8
- D 9
- E 10

10 What value(s) must be eliminated so that data set has 1 mode: 2, 2, 3, 3, 5, 6 ?

11 Find the mode(s): 3, 4, 4, 5, 5, 6, 7, 8, 9

- A 4
- B 5
- C 9
- D No mode

Finding the Mean

To find the mean of the ages for the Apollo pilots given below, add their ages. Then divide by 7, the number of pilots.

Apollo Mission	11	12	13	14	15	16	17
Pilot's age	39	37	36	40	41	36	37

$$\text{Mean} = \frac{39 + 37 + 36 + 40 + 41 + 36 + 37}{7} = \frac{266}{7} = 38$$

The mean of the Apollo pilots' ages is 38 years.

Find the Mean

Find the mean

10, 8, 9, 8, 5

12 Find the mean

20, 25, 25, 20, 25

13 Find the mean

14, 17, 9, 2, 4,10, 5, 3

14 The data value that occurs most often is called the

- A mode
- B range
- C median
- D mean

15 The middle value of a set of data, when ordered from lowest to highest is the _____.

- A mode
- B range
- C median
- D mean

16 Find the maximum value: 15, 10, 32, 13, 2.

- A 2
- B 15
- C 13
- D 32

17 Identify the set(s) of data that has no mode.

- A 1, 2, 3, 4, 5, 1
- B 2, 2, 3, 3, 4, 4, 5, 5
- C 1, 1, 2, 2, 2, 3, 3,
- D 2, 4, 6, 8, 10

18 Find the range: 32, 21, 25, 67, 82.

19 Identify the outlier(s): 78, 81, 85, 92, 96, 145.

20 If you take a set of data and subtract the minimum value from the maximum value, you will have found the _____.

- A outlier
- B median
- C mean
- D range

Find...

Find the mean, median, mode, range and outliers for the data below.

High Temperatures for Halloween

Year	Temperature
2003	91
2002	92
2001	92
2000	89
1999	96
1998	88
1997	97
1996	95
1995	90
1994	89
1993	91
1992	92
1991	91

High Temperatures for Halloween

88 89 90 91 92 93 94 95 96 97

Mean

Median

Mode

Range

Outliers

High Temperatures for Halloween

Year	Temperature
2003	91
2002	92
2001	92
2000	89
1999	96
1998	88
1997	97
1996	95
1995	90
1994	89
1993	91
1992	92
1991	91

Hint

Find the mean, median, mode, range and outliers for the data.

Candy	Calories
Butterscotch Discs	60
Candy Corn	160
Caramels	160
Gum	10
Dark Chocolate Bar	200
Gummy Bears	130
Jelly Beans	160
Licorice Twists	140
Lollipop	60
Milk Chocolate Almond	210
Milk Chocolate	210
Milk Chocolate Peanuts	210
Milk Chocolate Raisins	160
Malted Milk Balls	180
Pectin Slices	140
Sour Balls	60
Taffy	160
Toffee	60

Calories from Candy

0 10 20 30 40 50 60 70 80 90 100 110 120 130 140 150 160 170 180 190 200 210

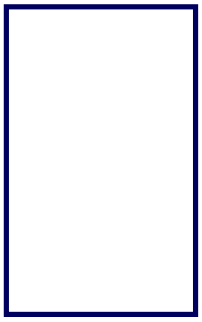
Mean

Median

Mode

Range

Outliers



Candy	Calories
Butterscotch Discs	60
Candy Corn	160
Caramels	160
Gum	10
Dark Chocolate Bar	200
Gummy Bears	130
Jelly Beans	160
Licorice Twists	140
Lollipop	60
Milk Chocolate Almond	210
Milk Chocolate	210
Milk Chocolate Peanuts	210
Milk Chocolate Raisins	160
Malted Milk Balls	180
Pectin Slices	140
Sour Balls	60
Taffy	160
Toffee	60

Central Tendency Application Problems

Return to Table of Contents

Application Problems - Method 1

Jae bought gifts that cost $24, $26, $20 and $18. She has one more gift to buy and wants her mean cost to be $24.
What should she spend for the last gift?

3 Methods :

Method 1: Guess & Check

Try $30

$\frac{24 + 26 + 20 + 18 + 30}{5} = 23.6$

Try a greater price, such as $32

$\frac{24 + 26 + 20 + 18 + 32}{5} = 24$

The answer is $32.

Application Problems - Method 2

Jae bought gifts that cost $24, $26, $20 and $18. She has one more gift to buy and wants her mean cost to be $24.
What should she spend for the last gift?

Method 2: Work Backward

In order to have a mean of $24 on 5 gifts, the sum of all 5 gifts must be $24 5 = $120.

The sum of the first four gifts is $88. So the last gift should cost $120 – $88 = $32.

24 5 = 120

120 – 24 – 26 – 20 – 18 = 32

Application Problems - Method 3

Method 3: Write an Equation

Let x = Jae's cost for the last gift.

$\frac{24 + 26 + 20 + 18 + x}{5} = 24$

$\frac{88 + x}{5} = 24$

88 + x = 120 (multiplied both sides by 5)

x = 32 (subtracted 88 from both sides)

Application Problems - Method 3

Your test scores are 87, 86, 89, and 88.
You have one more test in the marking period.

You want your average to be a 90.
What score must you get on your last test?

Answer

21 Your test grades are 72, 83, 78, 85, and 90. You have one more test and want an average of an 82. What must you earn on your next test?

22 Your test grades are 72, 83, 78, 85, and 90. You have one more test and want an average of an 85. Your friend figures out what you need on your next test and tells you that there is "NO way for you to wind up with an 85 average." Is your friend correct? Why or why not?

- Yes
- No

Consider the Data Set

Consider the data set: 50, 60, 65, 70, 80, 80, 85

The mean is:

The median is:

The mode is:

What happens to the mean, median and mode if 60 is added to the set of data?

Mean:

Median:

Mode:

Answer

Consider the Data Set

Consider the data set: 55, 55, 57, 58, 60, 63

- The mean is
- the median is
- and the mode is

What would happen if a value x was added to the set?

How would the mean change:

if x was less than the mean?

if x equals the mean?

if x was greater than the mean?

Consider the Data Set

Let's further consider the data set: 55, 55, 57, 58, 60, 63

- The mean is 58
- the median is 57.5
- and the mode is 55

What would happen if a value, "x", was added to the set?

How would the median change:

if x was less than 57?

if x was between 57 and 58?

if x was greater than 58?

Consider the Data Set

Consider the data set: 10, 15, 17, 18, 18, 20, 23

- The mean is 17.3
- the median is 18
- and the mode is 18

What would happen if the value of 20 was added to the data set?

How would the mean change?

How would the median change?

How would the mode change?

Answer

Consider the Data Set

Consider the data set: 55, 55, 57, 58, 60, 63

- The mean is 58
- the median is 57.5
- and the mode is 55

What would happen if a value, "x", was added to the set?

How would the mode change:

if x was 55?

if x was another number in the list other than 55?

if x was a number not in the list?

23 Consider the data set: 78, 82, 85, 88, 90. Identify the data values that remain the same if "x" is added to each value.

❑ A mean
❑ B median
❑ C mode
❑ D range
❑ E minimum

Data Displays

Return to Table of Contents

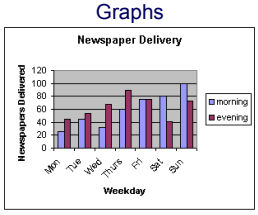
Data Display Examples

Tables

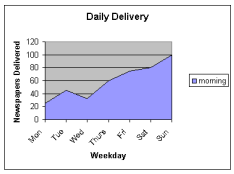
Ticket Sales for School Play

	Friday	Saturday	Sunday
7 PM	78	67	65
9 PM	82	70	30
Matinee	NA	35	82

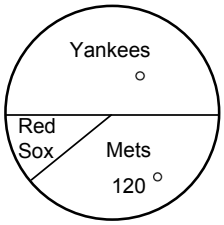
Graphs



Charts



24 In a recent poll in Syracuse, New York, 3,000 people were asked to pick their favorite baseball team. The accompanying circle graph shows the results of that poll.



A 300
B 500
C 1200
D 1800

From the New York State Education Department. Office of Assessment Policy, Development and Administration. Internet. Available from www.nysedregents.org/IntegratedAlgebra; accessed 17, June, 2011.

Frequency Tables and Histograms

Return to Table of Contents

Vocabulary

A frequency table shows the number of times each data item appears in an interval.

To create a frequency table, choose a scale that includes all of the numbers in the data set.

Next, determine an interval to separate the scale into equal parts.

The table should have the intervals in the first column, tally in the second and frequency in the third.

Time	Tally	Frequency
10-19	IIII	4
20-29		0
30-39	~~IIII~~	5
40-49	IIII	4
50-59		0
60-69	III	3

Frequency Table

The following are the test grades from a previous year.

Organize the data into a frequency table.

95 85 93
77 97 71
84 63 87
39 88 89
71 79 83
82 85

Determine Range, Scale and Interval

95 85 93
77 97 71
84 63 87
39 88 89
71 79 83
82 85

Step 1: Find the range of the data then determine a scale and interval.

Hint: Divide the range of data by the number of intervals you would like to have and then use the quotient as an approximate interval size.

Determine Range, Scale and Interval

Answer

RANGE: 97 – 39 = 59

SCALE: 59/6 = 9.5555 so 10 would be the size of the intervals

INTERVALS: 30-39, 40-49, 50-59, 60-69, 70-79, 80-89, 90-99

Step 1 ... nine a scale and interval

Hint: Div... of intervals you would like to have and then use the quotient as an approximate interval size.

Create a Frequency Table

95 85 93
77 97 71
84 63 87
39 88 89
71 79 83
82 85

Test Grades

Grade	Tally	Frequency

Move to see answer

Create a Frequency Table

Length of Time Walking

15 30 15 45
45 30 30 60
30 60 15 30
45 45 60 15

Answer

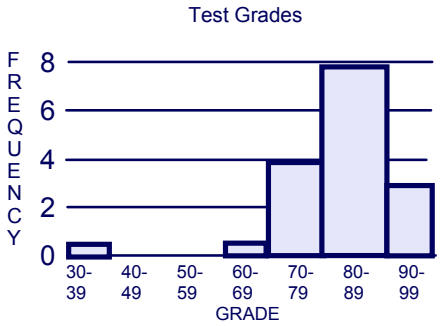
Move to see answer

60-69 III 3

Histogram

A histogram is a bar graph that shows data in intervals.

Since the data is shown in intervals,there is no space between the bars.



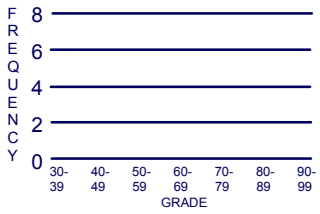
Create a Histogram

95 85 93
77 97 71
84 63 87
39 88 89
71 79 83
82 85

Test Grades

Grade	Tally	Frequency
30-39	I	1
40-49		0
50-59		0
60-69	I	1
70-79	IIII	4
80-89	IIII III	8
90-99	III	3

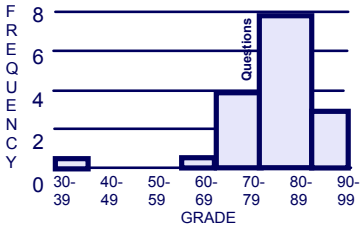
Test Grades



Note: Frequency tables and histograms show data in intervals

Questions

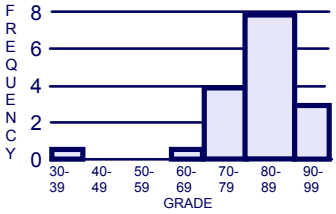
Test Grades



1. How many students scored an A?
2. How many students scored an 87?
3. How are histograms and bar graphs alike?
4. How are histograms and bar graphs different?
5. Why are there no spaces between the bars of a histogram?

Histograms

Test Grades



Notice that the test scores are closely grouped except one.

In statistics when a value is much different then the rest of the data set it is called an outlier.

25 In the following data what number is the outlier?
{ 1, 2, 2, 4, 5, 5, 5, 13 }

26 In the following data what number is the outlier?
{ 27, 27.6, 27.8 , 27.8, 27.9, 32 }

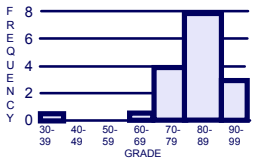
27 In the following data what number is the outlier?
{ 47, 48, 51, 52, 52, 56, 79 }

Create a Frequency Table and Histogram

Example
Test Scores

95	85	93
77	97	71
84	63	87
39	88	89
71	79	83
82	85	

Grade	Tally	Frequency
30-39	I	1
40-49		0
50-59		0
60-69	I	1
70-79	IIII	4
80-89	IIII III	8
90-99	III	3

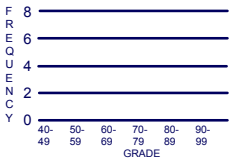


Create a Frequency Table and Histogram

Test Scores

87	53	95
85	89	59
86	82	87
40	90	72
48	68	57
64	85	

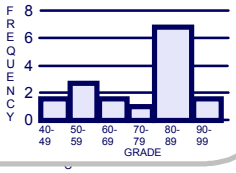
Grade	Tally	Frequency



Create a Frequency Table and Histogram

Answer

Grade	Tally	Frequency
40-49	II	2
50-59	III	3
60-69	II	2
70-79	I	1
80-89	IIII II	7
90-99	II	2

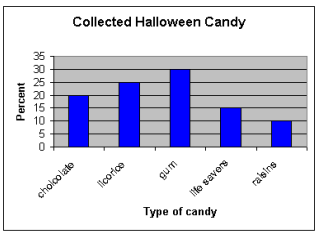


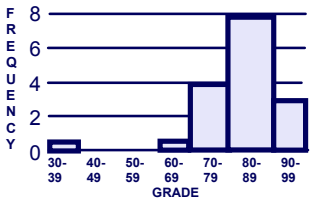
Compare and Contrast Bar Graphs and Histograms

Both compare data in different categories and use bars to show amounts.

Histograms show data in intervals, the height of the bar shows the frequency in the interval and there are no spaces between the bars.

Bar Graphs show a specific value for a specific category, and have a space between bars to separate the categories.





Stem and Leaf Plots

Return to Table of Contents

Stem-and-Leaf Plots

A type of graph that shows each data value and the number of occurrences. The leaf is the last digit and the stem consists of the remaining digits

Example:

List of math test grades for Bobby:
73, 42, 67, 94, 78, 99 84, 91, 82, 86, 94

First, order from least to greatest:
42, 67, 73, 78, 82, 84, 86, 91, 94, 94, 99

Bobby's Test Grades

4	2
6	7
7	3 8
8	2 4 6
9	1 4 4 9

Key: 4 | 2 = 42

Stem-and-Leaf Plots

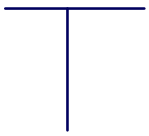
Create a stem-and-leaf for the data.

Remember:

- the leaf is the last digit and the stem consists of the remaining digits
- include a key

Answer

Daily Temperatures:
82, 95, 102, 78, 84, 96, 90, 80, 75, 101



Key:

Median

The median is the middle data value when the values are written in numerical order.

Remember! If a data set has an even number of values, the median is the mean of the two middle values.

Answer

Test	Grades
4	2
6	7
7	3 8
8	2 4 6
9	1 4 4 9

Key: 4 | 2 = 42

The median of this stem-and-leaf is 84.

28 What is the median of the data in the following stem-and-leaf plot?

Test	Grades
6	0 2 2
7	5 7 8 9
8	2 5 8
9	3 5 7

Key: 6 | 0 = 60

Answer

The Mean

The mean is the sum of the data values divided by the number of values.

Test	Grades
4	2
6	7
7	3 8
8	2 4 6
9	1 4 4 9

Key: 4 | 2 = 42

Step 1: Add all of the numbers

42 + 67 + 73 + 78 + 82 +84 + 86 + 91 + 94 + 94 + 99 = 890

Step 2: Divide the sum by the number of values (numbers)

890 divided by 11 = 80.9 So, 80.9 is the mean, also known as the average for this stem and leaf.

29

Test	Grades
6	0 2
7	5 7 8
8	2 5 8
9	3 5 7

Key: 6 | 0 = 60

29

Test Grade	
6	0
7	5
8	2
9	3 5

Answer

The Mode

The mode is the data value that occurs most often.

Remember!
The data set can have one mode, more than one mode or no mode.

Test Grades	
4	2
6	7
7	3 8
8	2 4 6
9	1 4 4 9

Key: 4 | 2 = 42

The mode of this stem-and-leaf is 94.

30

Test Grades	
6	0 2 2
7	5 7 8
8	2 5 8
9	3 5 7

Key: 6 | 0 = 60

30

Answer

31 Jorge made the accompanying stem-and-leaf plot of the weights, in pounds, of each member of the wrestling team he was coaching. What is the mode of the weights?

Stem	Leaf
10	9
11	
12	3 8
13	2 4 4 6 8
14	1 3 5 5 9
15	2 3 7 7 9
16	1 3 7 8 8 8 9
17	3 8

Key 16 | 1 = 161

A 145 C 152

B 150 D 168

32 The student scores on Mrs. Frederick's mathematics test are shown on the stem-and-leaf plot below.

Stem	Leaf
4	3
6	0 5 5 7 9
7	2 5 6 8 9 9 9
9	0 1 2 5 9

Key 4 | 3 = 43 points

Find the median of these scores.

The Stem and the Leaf

The *stem* is the first digit (the tens digit) which goes on the left.
The *leaf* is the second digit (the ones digit) which goes on the right.
Be sure to organize the leaves in numerical order.

Test Scores		
95	85	93
77	97	71
84	63	87
39	88	89
71	79	83
82	85	

Stem	Leaf

The Stem and the Leaf

Test Scores		
95	85	93
77	97	71
84	63	87
39	88	89
71	79	83
82	85	

Test Scores

Stem	Leaf
3	9
4	
5	
6	3
7	1179
8	23455789
9	357

Key 3|9 = 39

The Stem and the Leaf

Compare the stem and leaf plot to the frequency table from before.

Test Scores		
95	85	93
77	97	71
84	63	87
39	88	89
71	79	83
82	85	

Key 3|9 = 39

Stem	Leaf
3	9
4	
5	
6	3
7	1179
8	23455789
9	357

Grade	Tally	Frequency
30-39	I	1
40-49		0
50-59		0
60-69	I	1
70-79	IIII	4
80-89	III III	8
90-99	III	3

Example

Test Scores		
95	85	93
77	97	71
84	63	87
39	88	89
71	79	83
82	85	

Try This!
Create a Stem & Leaf Plot for the data.
Look at the example on the left for guidance.

Stem	Leaf
3	9
4	
5	
6	3
7	1179
8	23455789
9	357

Key 3|9 = 39

Example

Test Scores		
87	53	95
85	89	59
86	82	87
40	90	72
48	68	57
64	85	

Test Scores		
95	85	93
77	97	71
84	63	87
39	88	89
71	79	83
82	85	

Stem	Leaf
4	08
5	379
6	48
7	2
8	2556779
9	05

Key|40 = 40

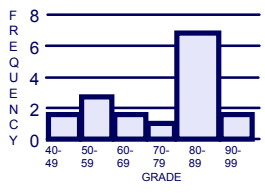
Answer

Stem and Leaf Plots

Stem and Leaf plots contain the information needed to make a histogram.

1. Compare the stem and leaf plot to the histogram. How are they alike? How are they different?
2. Can you make a stem and leaf plot from either a frequency table or histogram? Can you make a frequency table from a histogram?
3. How can you make a histogram from a stem and leaf plot? (Rotate the stem and leaf plot to demonstrate)

Stem	Leaf
4	08
5	379
6	48
7	2
8	2556779
9	05

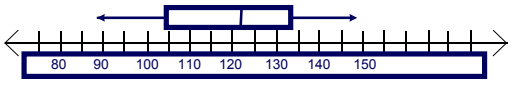


Box and Whisker Plots

Return to Table of Contents

Box and Whisker Plot

A box and whisker plot is a data display that organizes data into four groups



The median divides the data into an upper and lower half

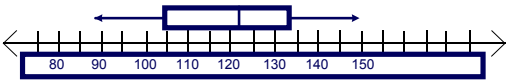
The median of the lower half is the lower quartile.

The median of the upper half is the upper quartile.

The least data value is the lower extreme .

The greatest data value is the upper extreme.

Box and Whisker Plot

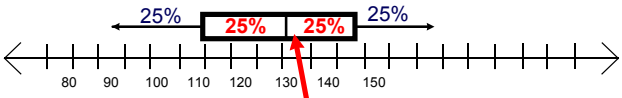


Drag the terms below to the correct position on the box and whisker graph.

median **lower quartile** **upper quartile**

upper extreme **lower extreme**

Box and Whisker Plot



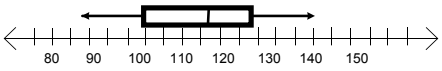
median

The entire box represents 50% of the data. 25% of the data lie in the box on each side of the median

Each whisker represents 25% of the data

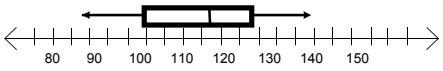
33 The lower extreme is

- A 87
- B 104
- C 122
- D 134



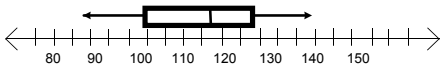
34 The median is

- A 87
- B 104
- C 122
- D 134



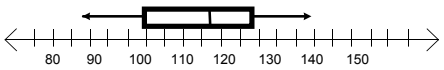
35 The lower quartile is

- A 87
- B 104
- C 122
- D 134



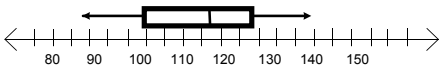
36 The upper quartile is

- A 87
- B 104
- C 122
- D 134



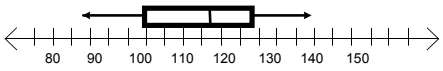
37 In a box and whisker plot, 75% of the data is between

- A lower extreme and the median
- B lower extreme and the upper extreme
- C lower quartile and the upper extreme
- D lower extreme and the upper quartile



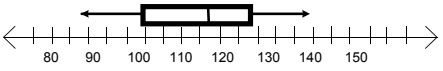
38 In a box and whisker plot, 50% of the data is between

- A lower extreme and the median
- B lower extreme and the upper extreme
- C lower quartile and the upper quartile
- D median and the upper extreme



39 In a box and whisker plot, 100% of the data is between

- A lower extreme and the median
- B lower extreme and the upper extreme
- C lower quartile and the upper quartile
- D median and the upper extreme



Find the Median

Steps for creating a box and whisker plot:

Find the median

8	8
9	667
10	1557
11	1259
12	2224589
13	236899
14	78

Find the Median

8	8
9	667
10	1557
11	1259
12	2224589
13	236899
14	78

median = 122

Then find the median of each half of the data

Find the Lowest and Highest Values

8	8
9	667
10	1557
11	1259
12	2224589
13	236899
14	78

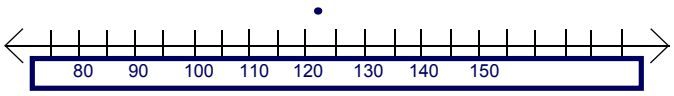
Lower Quartile = 105
Median = 122
Upper Quartile = 133

Then find the lowest and highest values

Draw the Plot

8	8
9	667
10	1557
11	1259
12	2224589
13	236899
14	78

Lower Extreme = 88
Lower Quartile = 105
Median = 122
Upper Quartile = 133
Upper Extreme = 148



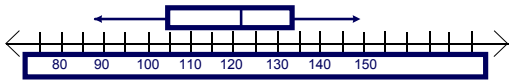
Create a box and whisker plot by plotting all
5 pieces of information. Then draw the plot.

Draw the Plot

Create a box and whisker plot by plotting all
5 pieces of information. Then draw the plot.

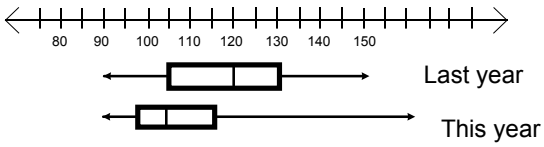
8	8
9	667
10	1557
11	1259
12	2224589
13	236899
14	78

Lower Extreme = 88
Lower Quartile = 105
Median = 122
Upper Quartile = 133
Upper Extreme = 148



40 Compare the two box and whisker plots.

Wrestling Team Weights

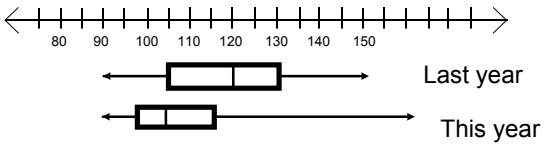


Last year's team had a greater median.

- True
- False

41 Compare the two box and whisker plots.

Wrestling Team Weights

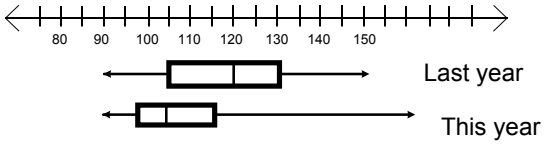


Both teams have about the same range.

- True
- False

42 Compare the two box and whisker plots.

Wrestling Team Weights

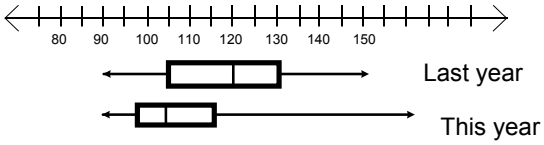


Last year's quartiles and median are lower than this year's.

- True
- False

43 Compare the two box and whisker plots.

Wrestling Team Weights



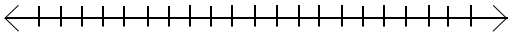
50% of the wrestlers weighed between 105 and 130 last year.

- True
- False

Try this!

Stem	Leaf
2	66689
3	56789
4	0011123458
5	02335567
6	12347
7	03

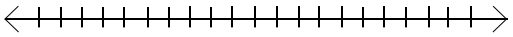
Lower Extreme =
Lower Quartile =
Median =
Upper Quartile =
Upper Extreme =



Try this!

Stem	Leaf
10	7
11	56
12	9
13	24
14	0224589
15	2344445

Lower Extreme =
Lower Quartile =
Median =
Upper Quartile =
Upper Extreme =

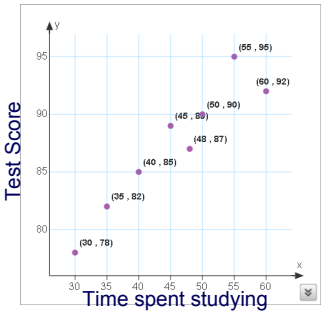


Scatter Plots and the Line of Best Fit

Return to Table of Contents

Scatter Plot

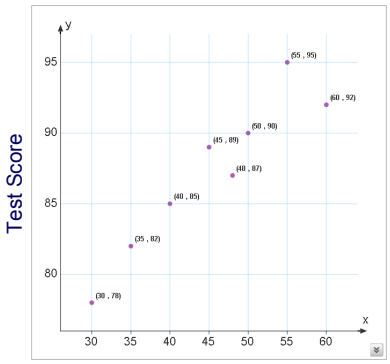
Time Studying	Test Score
45	89
30	78
50	90
60	92
40	85
48	87
55	95
35	82



What do you observe?

A scatter plot is a graph that shows a set of data that has two variables.

Scatter Plot

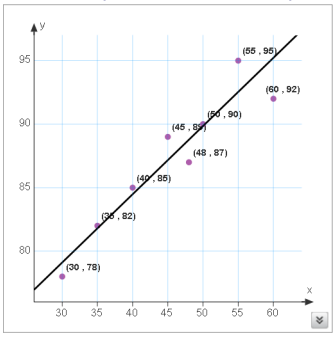


Predict the test score of someone who spends 52 minutes studying.

Predict the test score of someone who spends 75 minutes studying.

Draw a Line

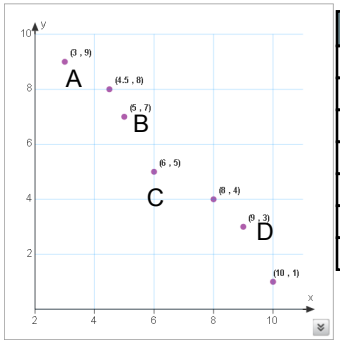
Notice that the points form a linear like pattern. To draw a line of best fit, use two points so that the line is as close as possible to the data points.



Our line is drawn so that it fits as close as possible to the data points. This line was drawn through (35,82) and (50,90).

44 Consider the scatter graph to answer the following: Which 2 points would give the best line of fit?

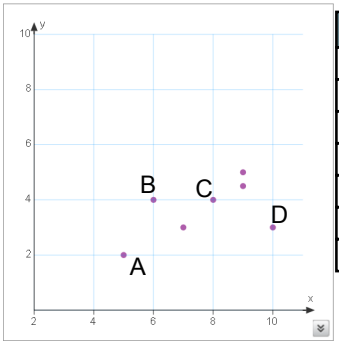
- A A and D
- B B and C
- C C and D
- D there is no pattern



X	Y
3	9
4.5	8
5	7
6	5
8	4
9	3
10	1

45 Consider the scatter graph to answer the following: Which 2 points would give the best line of fit?

- A A and D
- B B and C
- C C and D
- D there is no pattern



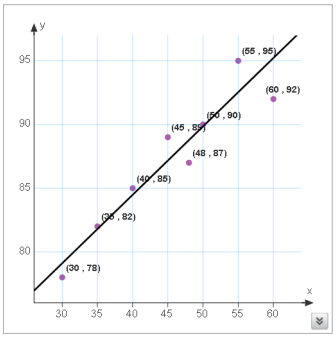
X	Y
5	2
6	4
7	3
8	4
9	4.5
9	5
10	3

Determining the Prediction Equation

Return to Table of Contents

Draw a Line

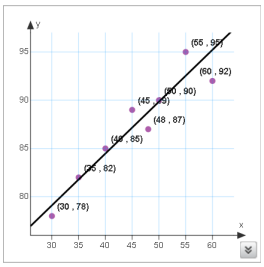
The points form a linear like pattern, so use two of the points to draw a line of best fit.



Our line is drawn so that it fits as close as possible to the data points. This line was drawn through (35,82) and (50,90).

Prediction Equation

Use the two points that formed the line to write an equation for the line.



Find m

$$m = \frac{90 - 82}{50 - 35}$$

$$m = \frac{8}{15}$$

Find b

$$90 = \frac{8}{15}(50) + b$$

$$90 = \frac{80}{3} + b$$

$$\frac{190}{3} = b$$

$$S = \frac{8}{15}t + \frac{190}{3}$$

Where S is the score for t minutes of studying.

This equation is called the Prediction Equation.

Extrapolation

Prediction Equations can be used to predict other related values.

$$S = \frac{8}{15}t + \frac{190}{3}$$

If a person studies 15 minutes, what would be the predicted score?

$$S = \frac{8}{15}(15) - \frac{190}{3} \approx 71.3$$

This is an extrapolation, because the time was outside the range of the original times.

Interpolation

If a person studies 42 minutes, what would be the predicted score?

$$S = \frac{8}{15}(42) - \frac{190}{3} \approx 85.7$$

This is an interpolation, because the time was inside the range of the original times.

What is Wrong?

Interpolations are more accurate because they are within the set.

The farther points are away from the data set the less reliable the prediction.

Using the same prediction equation, consider:

If a person studies 120 minutes, what will be there score?

$$S = \frac{8}{15}(120) + \frac{190}{3} \approx 127.3$$

What is wrong with this prediction?

What is the Prediction?

If a student got an 80 on the test,
What would be the predicted length of their study time?

$$80 = \frac{8}{15}t + \frac{190}{3}$$

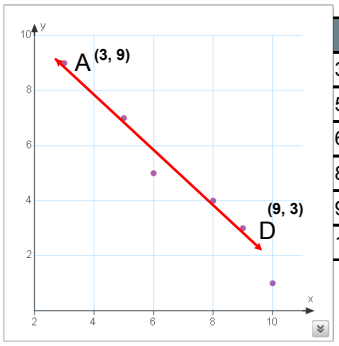
$$16.7 = \frac{8}{15}t$$

$$31.25 = t$$

The student studied about 31 minutes.

46 Consider the scatter graph to answer the following: What is the slope of the line of best fit going through A and D?

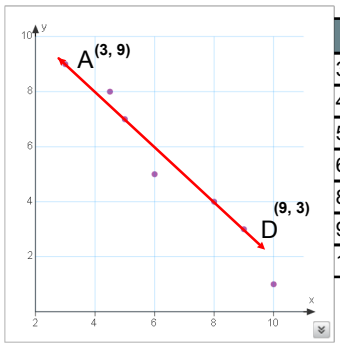
- A $\frac{3}{4}$
- B $\frac{-3}{4}$
- C 1
- D –1



X	Y
3	9
5	7
6	5
8	4
9	3
10	1

47 Consider the scatter graph to answer the following: What is the y-intercept of the line of best fit going through A and D?

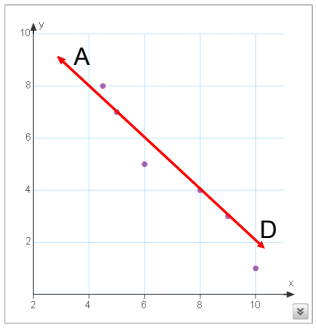
A 9
B 10
C 11
D 12



X	Y
3	9
4.5	8
5	7
6	5
8	4
9	3
10	1

48 Consider the scatter graph to answer the following The equation for our line is y = –1x + 12. What would the prediction be if x = 7? Is this an interpolation or extrapolation?

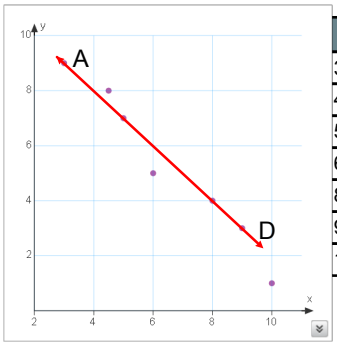
- A 5, interpolation
- B 5, extrapolation
- C 6, interpolation
- D 6, extrapolation



X	Y
3	9
4.5	8
5	7
6	5
8	4
9	3
10	1

49 Consider the scatter graph to answer the following: The equation for our line is y = –1x + 12. What would the prediction be if x = 14? Is this an interpolation or extrapolation?

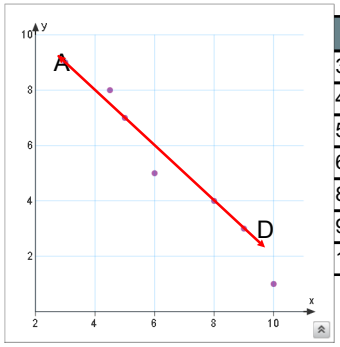
- A –4, interpolation
- B –4, extrapolation
- C –2, interpolation
- D –2, extrapolation



X	Y
3	9
4.5	8
5	7
6	5
8	4
9	3
10	1

50 Consider the scatter graph to answer the following: The equation for our line is y = –1x + 12. What would the prediction be if y = 11? Is this an interpolation or extrapolation?

- A 1, interpolation
- B 1, extrapolation
- C 2, interpolation
- D 2, extrapolation



X	Y
3	9
4.5	8
5	7
6	5
8	4
9	3
10	1

51 In the previous questions, we began by using the table at the right. Which of the predicted values (7,5) or (14, –2) will be more accurate and why?

- A (7,5); it is an interpolation
- B (7,5); there already is a 5 and a 7 in the table
- C (14, -2) it is an extrapolation
- D (14, -2); the line is going down and will become negative

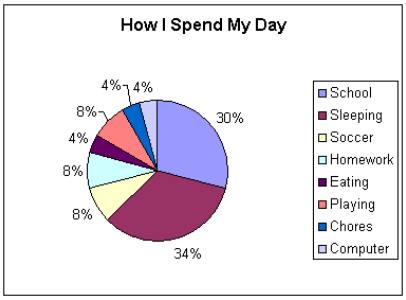
X	Y
3	9
4.5	8
5	7
6	5
8	4
9	3
10	1

Choosing a Data Display

Return to Table of Contents

Part to a Whole

A circle graph is used to illustrate a part to whole relationship



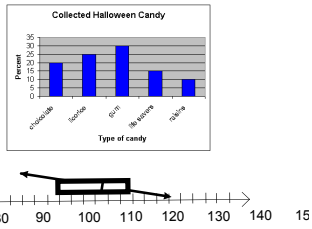
What You Have Learned

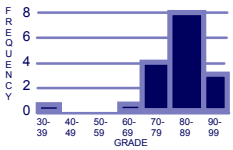
You have also learned about:

- Bar Graphs
- Histograms
- Frequency Tables
- Box and Whisker Graphs
- Stem and Leaf Graphs

Time	Tally	Frequency
10-19	IIII	4
20-29		0
30-39	IIII	5
40-49	IIII	4
50-59		0
60-69	III	3

6	0 2
7	5 7 8
8	2 5 8
9	3 5 7





52 Choose the best data display to show the number of hours of video games played each week for two months.

- A bar graph
- B histogram
- C circle graph
- D frequency table
- E stem-and-leaf
- F box-n-whisker

53 Choose the best data display to show the lower 25% of the scores on a math test.

- A bar graph
- B histogram
- C circle graph
- D frequency table
- E stem-and-leaf
- F box-n-whisker

54 Choose the best data display to show the number of students that earned an A, B, C, D & E on the last test.

- A bar graph
- B histogram
- C circle graph
- D frequency table
- E stem-and-leaf
- F box-n-whisker

55 Choose the best data display to show the percent of students that earned an A, B, C, D & E on the last test.

- A bar graph
- B histogram
- C circle graph
- D frequency table
- E stem-and-leaf
- F box-n-whisker

56 Choose the best data display to show the interval of grades for 50% of the students.

- A bar graph
- B histogram
- C circle graph
- D frequency table
- E stem-and-leaf
- F box-n-whisker

Misleading Graphs

Return to Table of Contents

Scale and Size

Changing the scale
The bigger the interval on the horizontal axis the more vertical the graph looks

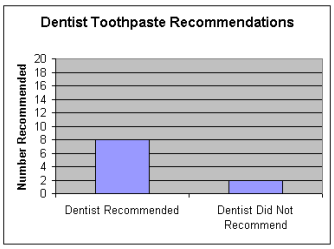
The bigger the interval on the vertical axis, the more horizontal the graph looks

Gap in the scale
Tricks the viewer into thinking that bars or lines begin at a value different from the actual value

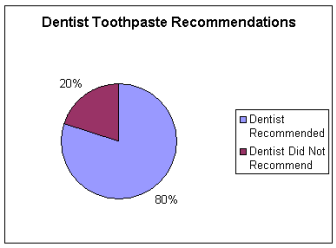
Changing the size of objects
The apparent sizes of objects may appear different from actual data values

57 You own a toothpaste company. Which graph would you select to use as an advertisement for your toothpaste?

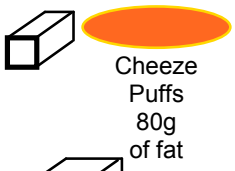
A

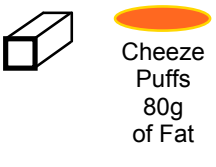


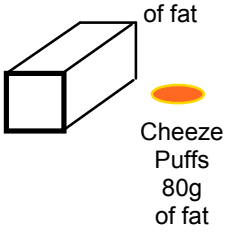
B



59 Super Crunch Cheeze Puffs is coming out with a new "healthy" snack food. Determine which of the graphs would best fit their "healthy" advertising needs.

A 

C 

B 

D