

Name: _____ Period _____

Unit 4A:

Describing Data

Resources and Information at:

www.tinyurl.com/9layson



Unit 4A: Describing Data

Unit Essential Question: How do we summarize, represent, interpret, and compare data?

Concept:

Concept:

Concept:

Lesson Essential Questions:

Lesson Essential Questions:

Lesson Essential Questions:

Vocabulary:

Vocabulary:

Vocabulary:

Additional Information/Resources:

www.usatestprep.com

school: newmanchesterga
code: newton58

Mean, median, and mode are called measures of central tendency. These measures are used to summarize or to compare data.

- Use **mean** when you need the **average** of a group of numbers. If one of the numbers in the group is much higher or lower than the other numbers, the mean is not the best choice to summarize data.
- Use **mode** when the data entries are not numbers.
- Use **median** when you want to find the **middle** value of a group of numbers or to break the data entries into two groups.

#1-4 Decide which measure of central tendency to use. Use that method to find your answer.

1. Seven of Devon's friends were born in November, September, June, June, April, January and June.
2. Rakesh watches television for the following amounts of time during the week: 2 hours, 3.5 hours, 1.5 hours, 4 hours, 2 hours, and 2 hours.
3. Daily earnings: \$60, \$65, \$38, \$78, \$80
4. Daily computer use: 64 minutes, 85 minutes, 55 minutes, 92 minutes, 120 minutes, 154 minutes, 171 minutes

WEIGHTS OF STUDENTS IN CLASS

Name	Weight (kg)
Martha	49
Mary	60
Tom	58
Gene	73
Chip	67
Meg	60
Dottie	63
Sean	80
Kim	64
Alli	68
Jesse	60

- Use the data at the right for Exercises 5-10
5. What is the mode?
 6. What is the mean?
 7. What is the median?

Suppose Sonya enrolls in the class and her weight is 51 kg. Without computing, answer these questions.

8. How will Sonya affect the new mean?
9. How will Sonya affect the new median?

Suppose Hector now joins the class. His weight is 70 kg.

10. When both Sonya and Hector join the class, what are the new mode, mean, and median?

Name _____

Date _____

OUTLIERS - Lesson 4.1

An outlier is a number that is quite different from the other values in a data set.

Randi recorded the number of hours she worked each week during July and August.

18 22 16 20 44 19 21 16

1. Which value is an outlier? _____
2. What is the mean of the data set? (Round to the nearest tenth.) _____
3. What is median of the data set? _____
4. What is the mean of the data set if the outlier is removed? (Round to the nearest tenth.) _____
5. How does the outlier affect the mean of this data set? _____
6. What is the median of the data set if the outlier is removed? _____
7. How does the outlier affect the median of this data set? _____

Marco recorded the number of cars that passed through a toll plaza each hour of his shift.

98 112 37 124 89 105 94 120

8. Which value is an outlier? _____
9. What is the mean of the data set? (Round to the nearest tenth.) _____
10. What is median of the data set? _____
11. What is the mean of the data set if the outlier is removed? _____
12. How does the outlier affect the mean of this data set? _____
13. What is the median of the data set if the outlier is removed? _____
14. How does the outlier affect the median of this data set? _____

OUTLIERS Lesson 4.1 HW

An outlier is a number that is quite different from the other values in a data set.

A video store owner recorded the number of DVDs that were rented each day during the first week of October.

Sunday - 98	Monday - 79	Tuesday - 83	Wednesday - 72
Thursday - 77	Friday - 96	Saturday - 155	

- Which value is an outlier? _____
- What is the mean of the data set? (Round to the nearest tenth.) _____
- What is the median of the data set? _____
- What is the mean of the data set if the outlier is removed? (Round to the nearest tenth.) _____
- What is the median of the data set if the outlier is removed? _____
- How does the outlier affect the mean and the median of this data set? _____

Building Skills

Decide which measure of central tendency to use. Use that method to find your answer.

- 24, 35, 31, 24, 25, 9, 27
- Length of movies: 117 minutes, 108 minutes, 115 minutes, 124 minutes, 118 minutes
- Cost of CDs: \$15.99, \$15.99, \$21.99, \$16.99, \$17.99, \$18.99
- Friends' favorite flowers: roses, carnations, mums, roses, daisies, roses

Problem Solving

Choose a measure of central tendency. Then solve.

- Olivia plays basketball. In the last four games, she scored 18 points, 16 points, 15 points, and 17 points.
- Phil is making CD recordings of his band. The CDs have 64 minutes, 48 minutes, 69 minutes, 66 minutes, and 65 minutes of music.

(4)

READING AND INTERPRETING DATA IN CHARTS, TABLES, AND PLOTS

A stem-and-leaf plot shows numerical data in an organized chart. In this example of weights in pounds of some golden retriever dogs, the stem is the tens digit and the leaf is the ones digit.

Use this stem-and-leaf plot to answer questions 1 - 5.

- How many dogs' weights are noted in the plot? _____
- How many dogs have weights greater than 65 pounds but less than 75 pounds? _____
- The mode is the number that appears most often in the data set. What is the mode? _____
- The range is the difference between the least and greatest numbers in the data set. What is the range? _____
- The median is the value that appears in the middle of the data set when the values are arranged from least to greatest. What is the median? _____

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

WEIGHTS OF GOLDEN RETRIEVERS (lb)

Mr. Ruiz recorded his students' scores on a Science test in the table below. Display this data in a stem-and-leaf plot.

89	76	100	98	96
75	82	88	86	100
62	79	77	80	94
100	86	74	66	98

STEMS	LEAVES

Use your stem-and-leaf plot to answer the following questions.

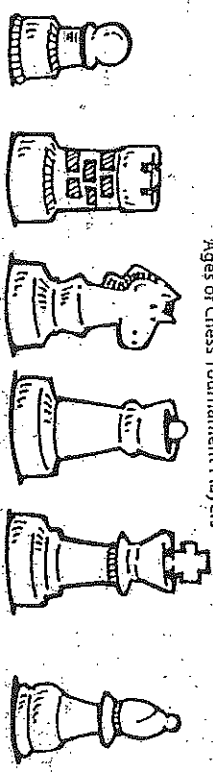
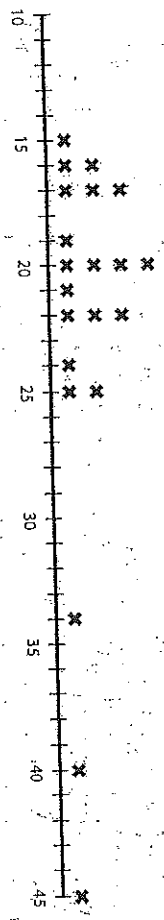
- What is the mode? _____
- What is the range? _____
- What is the median? _____

(5)

Read a Line Plot

(Dot Plot)

A dot plot is a graph that shows frequency of data along a number line. An outlier is a number or numbers that are much larger or smaller than other values.

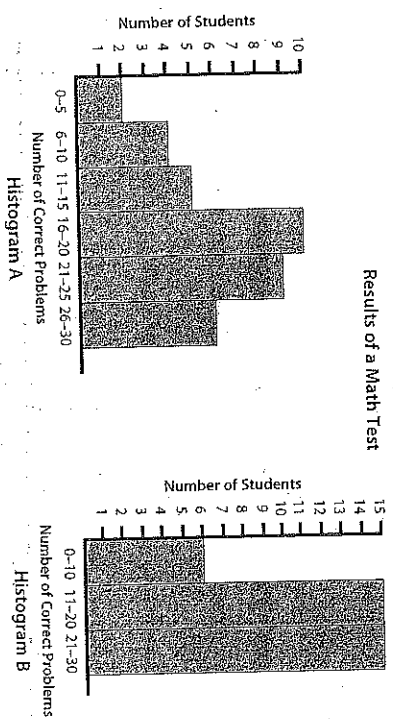


- 1) How many chess players are represented in the dot plot? _____
- 2) Which age occurs the most frequently (mode)? _____
- 3) What are the outlier(s) on this graph? _____ Why might these outlier(s) occur? _____
- 4) What's the average age of the chess players? _____
- 5) If you toss out the outlier(s), how will this change your average? _____
- 6) What is the median? _____ What is the range? _____
- 7) Describe the variability of this plot, mention any clusters or gaps. _____
- 8) Describe the distribution of the data. Is it skewed? _____

(6)

Read a Histogram

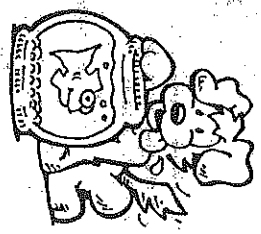
A histogram looks similar to a vertical bar graph. A histogram has one variable, and data is sorted into categories of this one variable. The range of your variable is important. If the range is too wide or too narrow, it may be difficult to interpret the data.



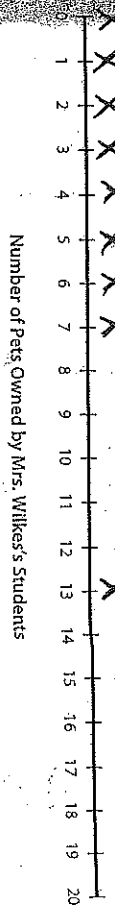
- 1a) How many students took the math test?** _____
- 1) How many students got between 16 and 20 problems correct? _____ Which histogram(s) gives you this answer? _____
 - 2) How many students got between 21 and 30 problems correct? _____ Which histogram(s) gives you this answer? _____
 - 3) Both histograms show the results of the same math test. How are the two histograms different? How are they the same? **Their distribution?** _____
 - 4) Which histogram gives more useful data? Explain. _____
 - 5) How could you change the histogram to better show those students who received a passing score? _____

Open Analysis - 5-6 © 2007 Creative Teaching Press

Plot a Spot



In Mrs. Wilke's class, students were asked to give the number of pets owned by their family. Their responses were recorded on the dot plot below. Use it to answer the following questions.



- How many students are represented in the dot plot? _____
- Which number of pets occurred most frequently (mode)? _____
- What are the outliers, if any, on this plot? _____
Explain why these outliers may have occurred? _____
- What is the average number of pets owned by the students? _____
- Toss out the outlier. How does that change your average? _____
- What is the median? _____ What is the range? _____
- Describe the variability of this plot, mention any clusters or gaps. _____
- Describe the distribution of the data. Is it skewed? _____

8

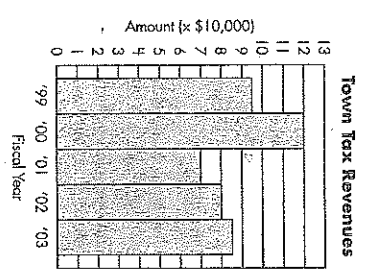
Lesson 4.2 HW Histograms + Stem-and-leaf Plots

A **histogram** is a type of bar graph. The categories are consecutive and the intervals are equal. Histograms are often used to analyze changes over a given time period.

Look at the histogram. What do the bars represent? Each bar represents an interval of 1 fiscal year. Five fiscal years are represented.

Which two consecutive years saw the greatest change in tax revenues? 2000-2001

How much did the town collect in tax revenues in 1999? (Remember that each line in the chart represents \$10,000.) \$95,000



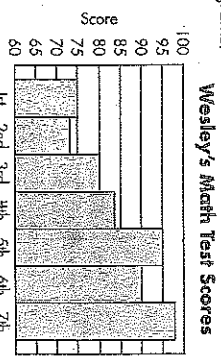
Answer the questions by interpreting data from the histograms.

- What does each bar represent? _____
- On which test did Wesley do best? _____
- What was the difference between Wesley's best score and his worst score? _____
- On which test did Wesley score 95 points? _____

5. What was Wesley's mean test score? _____

Use the stem-and-leaf plot of some poodles' weights, below, to answer questions 6 - 10.

- How many dogs' weights are noted in the plot? _____
- How many dogs have weights greater than 25 pounds but less than 35 pounds? _____
- What is the mode? _____
- What is the range? _____
- What is the median? _____



WEIGHTS OF POODLES (lb)

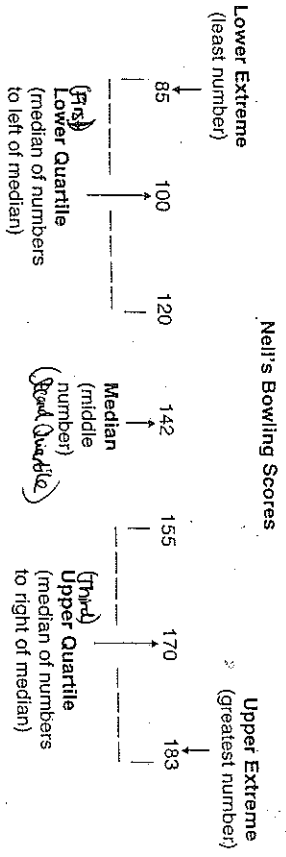
Stem	Leaf
2	6, 8, 8, 9
3	0, 2, 2, 3, 3, 3, 7, 8, 8
4	1, 4

9

Quartiles and Extremes

10.8
TAKE ANOTHER LOOK

Arranging data in increasing order can help you to better understand it.



Use this data for Exercises 1-5.

Weights of seven pumpkins (in pounds)
12 14 18 17 19 21 26

- The median of this data is _____.
- The upper quartile of this data is _____.
- The lower extreme of this data is _____.
- The lower quartile of this data is _____.
- The upper extreme of this data is _____.

Arrange the data in this table in increasing order. Then use it for Exercises 6 and 7.

	Test Scores
Boys	98 96 93 89 87 76
Girls	90 97 87 74 89 67

- The upper quartile for the boys' scores is _____.
- The lower extreme of girls' scores is _____ than the lower extreme of boys' scores. _____ or more.

Use with text pages 330-331.

R 111

10

Make a Box Plot

This is a list of the temperatures recorded periodically in April.

37, 44, 44, 45, 52, 54, 60, 56, 52, 53, 48, 46, 45, 46, 43

Temperatures in April



- Rewrite the numbers in order. What is the median? _____
- What are the high extreme and the low extreme? What is the range? _____
- What is the lower quartile? What is the upper quartile? _____
- What is the interquartile range? _____
- Use the data above to make a box plot on the number line above. Is the data skewed? _____
- What percentage of the temperatures were between 40 and 53? Above 44? _____

Lesson 4.3a HW

Quartiles and Extremes

NAME _____

1) 0 1 1 1 2 2 3 3 4 5

Lower Extreme _____
 Lower Quartile _____
 Median _____
 Upper Quartile _____
 Upper Extreme _____

2) 14 15 16 18 20 22 22 23

Lower Extreme _____
 Lower Quartile _____
 Median _____
 Upper Quartile _____
 Upper Extreme _____

3) 02 75 77 80 81 85 87 91 94

Lower Extreme _____
 Lower Quartile _____
 Median _____
 Upper Quartile _____
 Upper Extreme _____

4) 13 14 16 16 17 17 18 30

Lower Extreme _____
 Lower Quartile _____
 Median _____
 Upper Quartile _____
 Upper Extreme _____

5) 3 3 3 5 5 6 7 7 7 8 8

Lower Extreme _____
 Lower Quartile _____
 Median _____
 Upper Quartile _____
 Upper Extreme _____

6) 190 191 192 192 194 195 195 196

Lower Extreme _____
 Lower Quartile _____
 Median _____
 Upper Quartile _____
 Upper Extreme _____

7) 6 7 9 9 10 10 13 13 15

Lower Extreme _____
 Lower Quartile _____
 Median _____
 Upper Quartile _____
 Upper Extreme _____

8) 21 22 24 25 27 28 32 35

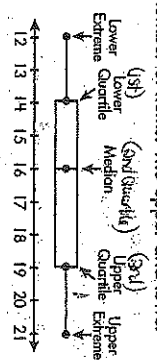
Lower Extreme _____
 Lower Quartile _____
 Median _____
 Upper Quartile _____
 Upper Extreme _____

13

Lesson 4.3a HW Box-and-Whisker Plots

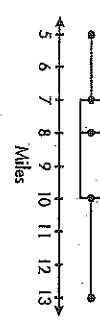
A **box-and-whisker** plot displays data along a number line. Quartiles are used to divide the data into four equal parts. Each quartile is 25% of the number of items. The upper and lower quartiles, representing 50% of the data, form the box. The upper extreme (highest value) and lower extreme (lowest value) form the whiskers.

This box-and-whisker plot represents the following data:
 12, 13, 14, 14, 15, 16, 17, 18, 19, 19, 21
 Upper Extreme: 21
 Lower Extreme: 12
 Middle Quartile (median): 16
 (1st) Lower Quartile (median of lower half): 14
 (3rd) Upper Quartile (median of upper half): 19
 Range (Difference between Upper and Lower Quartiles): 5

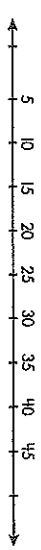


Use the box-and-whisker plots below to answer the following questions.

Miles Walked in Walk-a-Thon



- What is the median number of miles walked? _____
- What are the lowest and highest numbers of miles walked?
 lowest: _____ highest: _____
- If 126 people participated in the walkathon, how many people walked 7–10 miles? _____
- What percentage of the people walked more than 10 miles? _____
- The scores on a recent daily quiz were 10, 15, 20, 20, 30, 30, 40.
 What is the median of these scores? _____
- What is the lower quartile? _____
- What is the upper quartile? _____
- Using the number line below, draw a box-and-whisker plot for these scores.

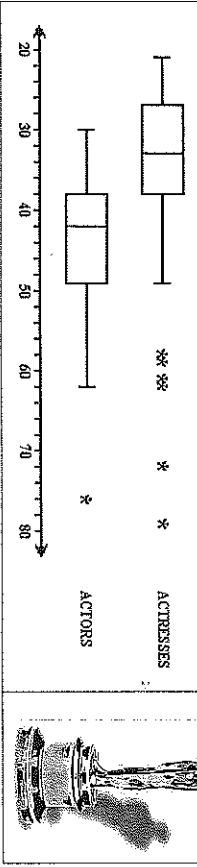


13

BOX-WHISKER PLOTS - THE OSCARS

EQ: How do we interpret a box-whiskers plot with and without outliers?

Every year since 1928, an Academy Award has been given to the best actor and best actress in a motion picture. The box-whisker plots below display the ages at which the 67 best actors and 67 best actresses won their Oscars. Outliers are asterisks.



1 List the outlier ages for...

a. ACTRESSES	b. ACTORS
--------------	-----------

2 Write the five number summary, IQR and range for actresses and actors w/out outliers.

ACTRESSES	MIN	Q ₁	MEDIAN	Q ₃	MAX	IQR	RANGE
ACTORS							

3 Write the five number summary, IQR and range for actresses and actors WITH OUTLIERS.

ACTRESSES	MIN	Q ₁	MEDIAN	Q ₃	MAX	IQR	RANGE
ACTORS							

4 Complete the chart below using the information above. SHOW WORK.

	DIFFERENCE OF RANGES	DIFFERENCE OF IQRs
ACTRESSES		
ACTORS		

5 What changed more... the range or the IQR. Explain why.

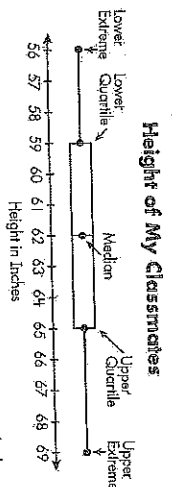
6 Marlee Matlin was the youngest actress to ever win an Oscar. How old was she? _____
 Marlon Brando and Richard Dreyfuss were the youngest actors to ever win an Oscar. How old were they? _____
 Henry Fonda was the oldest actor to ever win an Oscar. How old was he? _____

14

Lesson 4.3b Double Box-and-Whisker Plots

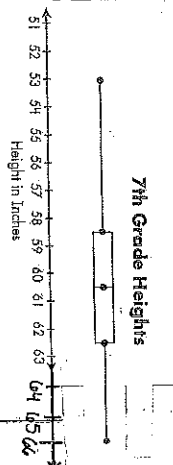
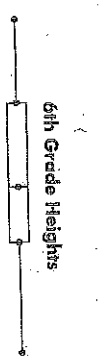
NAME _____

A **box-and-whisker plot** is a specific way to represent a set of data. The middle 50% of the data set is indicated by a rectangle, divided at the median of the data set. The lowest and highest 25% are indicated by "whiskers," or lines that branch out from either side of the rectangle. Mavis used the data she collected to create this box-and-whisker plot. It can be compared to the line plot on the previous page. A box-and-whisker plot does not show the number of data points. It cannot be used to find the mean or mode of the data.



What is the range? (Subtract the lower extreme from the upper extreme.) The range is 13 inches. Between which two numbers are the middle 50% of the students' heights? Look at the two ends of the box. The middle 50% are between 59 and 65. The interquartile range is 65 - 59, or 6.

Mavis polled 6th and 7th graders on their height. She made these box-and-whisker plots with the data. Interpret each one to answer the questions.



- What is the interquartile range in the 6th grade data? _____
- What is the median height in 6th grade? _____
- Between which two numbers are the lowest 25% of students' heights in 6th grade? _____
- What height is at the upper extreme in the 7th grade? _____
- What is the range of heights in 7th grade? _____
- What is the interquartile range in the 7th grade data? _____
- Is the data in either box-and-whisker plot skewed? _____

8. Compare the variability of both box-and-whisker plots. Are they the same? Does one have a higher variability than the other. Explain.

Spectrum Math
Grade 8
74

15

Lesson 4.4 - Two-Way Frequency Tables

Introduction

A _____ is a table of data that compares two variables (_____). It separates responses by a characteristic of the respondents.

Example 1:

Nail Service	Gender		Total
	Females	Males	
Manicure	58	31	89
Pedicure	77	27	104
Total	135	58	193

A _____ is the number of responses for a given characteristic.

- The _____ of a two-way frequency table are joint frequencies.
- In our table above, _____ and _____ are each joint frequencies.
- A _____ is the total number of times a response was given, or the total number of respondents with a given characteristic.

A _____ expresses a number of responses as a percentage of the total number of responses.

- It allows a _____ to be made for multiple responses in a single row, single column, or table.
- They are found by _____ the number of responses by the total number of responses.

C.R. Freq. - Gender

C.R. Freq. - Service

Nail Service	Gender	
	Females	Males
Manicure	$\frac{58}{135} = 43\%$	$\frac{31}{58} = 53\%$
Pedicure	$\frac{77}{135} = 57\%$	$\frac{27}{58} = 47\%$

C.R. Freq. - Total People

Nail Services	Gender	
	Females	Males
Manicure	$\frac{58}{193} = .30 = 30\%$	$\frac{31}{193} = .16 = 16\%$

Example 2:

Abigail surveys students in different grades, and asks each student which pet they prefer. Their responses are in the table to the right.

Grade	Preferred pet				Total
	Bird	Cat	Dog	Fish	
9	3	49	53	22	
10	7	36	64	10	
Total					

1. How many 9th graders preferred fish?
2. What does the joint frequency 7 represent?
3. What is the joint frequency of 9th graders who preferred cats?
4. Which pet was the most popular among all the students surveyed?
5. Complete the marginal frequencies for each type of pet and for each grade level. Which one is the smallest?
6. How many students are represented in the survey?
7. What would the conditional relative frequency be for 9th graders who preferred cats (with respect to 9th graders)?
8. What would the conditional relative frequency be for fish (with respect to everyone surveyed)?

Example 3:

Deshawn surveyed the seniors at his high school to find out who eats and doesn't eat breakfast regularly. Their responses are in the table at the right.

Breakfast habits	Gender		Totals
	Male	Female	
Eat Breakfast regularly	190	110	300
Do not Eat Breakfast regularly	130	165	295
Totals	320	275	595

1. How many females "eat breakfast regularly"?
2. 275 is a _____ frequency. It represents _____
3. How many people were included in this survey?

16

17

4. Complete the table by calculating the conditional relative frequencies with respect to gender.

Breakfast habits	Gender		Totals
	Male	Female	
Eat Breakfast regularly	190	110	300
Do not Eat Breakfast regularly	130	165	295
Totals	320	275	595

5. What can you conclude about the breakfast habits of males and females?

You Try:

Use the Two-Way Frequency table to the right to answer the following questions.

	Dance	Sports	TV	Total
Men	2	10	8	20
Women	16	6	8	30
Total	18	16	16	50

1. Looking at just the marginal frequencies (row & column totals) what can we conclude about the 3 activities?

- A. Dance is way more interesting.
- B. They have roughly equal appeal.
- C. Sports is the least chosen activity.
- D. TV is the preferred activity.

2. Looking at the joint frequencies we see that women show a strong preference for which activity? _____

3. What are the conditional relative frequencies (for both gender and hobby) for men who showed a preference for Sports?

C.R. Freq. for Gender: _____

C.R. Freq. for Hobby: _____

18

Lesson 4.4 HW – Two-Way Frequency Tables

Directions: Use the following two-way frequency table to answer questions 1-6.

Gender	Favorite Meal			Totals
	Pizza	Hamburger	Hot Dog	
Male	18	17	9	12
Female	14	7	13	19
Totals				

1. Complete the table by finding the marginal frequency for each food and each gender.
2. What is the joint frequency of males who prefer hot dogs? _____
3. What does the 19 in the table represent? _____
4. Which meal had the greatest marginal frequency? _____
5. How many total people were surveyed? _____
6. What is the conditional relative frequency for males who preferred spaghetti (with respect to the people who preferred it)? _____

Directions: Use the following two-way frequency table to answer questions 7-10.

Transportation	Gender		Totals
	Male	Female	
Walk	34	46	80
Car	28	17	45
Bus	15	12	27
Cycle	52	17	69
Totals	129	92	221

7. Identify one joint frequency from this table and describe its meaning. _____
8. Identify one marginal frequency from this table and describe its meaning. _____
9. Explain how you would calculate the conditional relative frequency for females who rode the bus (with respect to all the people who were surveyed). _____
10. Write one conclusion that you can make from this two-way frequency table. _____

19