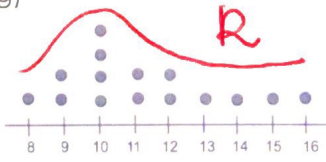
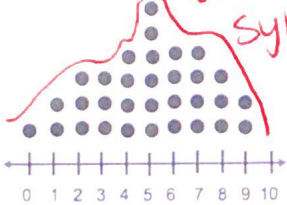
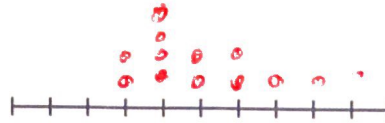


Topic	Things to Remember	Practice																				
1) Two Way Frequency Tables (day 1 notes)	<ul style="list-style-type: none">Joint frequencies: inside cellsMarginal frequencies: outside cells (totals)	a) Fill in the blanks below. <table><tr><th></th><th>High School Diploma</th><th>Bachelor's Degree</th><th>Master's/ Doctoral Degree</th><th>Total</th></tr><tr><th>Male</th><td>16</td><td>46</td><td>3</td><td>65</td></tr><tr><th>Female</th><td>12</td><td>51</td><td>3</td><td>66</td></tr><tr><th>Total</th><td>28</td><td>97</td><td>6</td><td>131</td></tr></table>		High School Diploma	Bachelor's Degree	Master's/ Doctoral Degree	Total	Male	16	46	3	65	Female	12	51	3	66	Total	28	97	6	131
		High School Diploma	Bachelor's Degree	Master's/ Doctoral Degree	Total																	
	Male	16	46	3	65																	
	Female	12	51	3	66																	
Total	28	97	6	131																		
<ul style="list-style-type: none">To create relative frequencies, divide each frequency by the grand total – round to the nearest hundredth	b) Create a two way relative frequency table based on the two way frequency table above. skip																					
<ul style="list-style-type: none">How many = whole numberProbability = decimal rounded to nearest hundredthPercent = percent (be sure to include % with answer)	Use either table above to answer the following questions. <table><tr><td>c) How many people have a high school diploma? 28</td><td>d) What percentage of people surveyed were female? 50.38%</td></tr><tr><td>e) What is the probability that a person has a Bachelor's Degree? 97/131 .74</td><td>f) What is the total number of people surveyed? 131</td></tr></table>	c) How many people have a high school diploma? 28	d) What percentage of people surveyed were female? 50.38%	e) What is the probability that a person has a Bachelor's Degree? 97/131 .74	f) What is the total number of people surveyed? 131																	
c) How many people have a high school diploma? 28	d) What percentage of people surveyed were female? 50.38%																					
e) What is the probability that a person has a Bachelor's Degree? 97/131 .74	f) What is the total number of people surveyed? 131																					
<ul style="list-style-type: none">Conditional probability key words: "if", "given that", "what percent of _____", etc.	Use either table above to answer the following questions. <table><tr><td>g) If a person is male, what is the probability that they have a Master's Degree? 3/65 .0461 or 4.61%</td><td>h) What percent of people with high school diplomas are male? 16/28 57.14%</td></tr><tr><td>i) What percent of people with Bachelor's Degrees are female? 51/97 52.58%</td><td>j) Given that a person is female, what is the probability that they have a high school diploma? 12/66 .1818 or 18.18%</td></tr></table>	g) If a person is male, what is the probability that they have a Master's Degree? 3/65 .0461 or 4.61%	h) What percent of people with high school diplomas are male? 16/28 57.14%	i) What percent of people with Bachelor's Degrees are female? 51/97 52.58%	j) Given that a person is female, what is the probability that they have a high school diploma? 12/66 .1818 or 18.18%																	
g) If a person is male, what is the probability that they have a Master's Degree? 3/65 .0461 or 4.61%	h) What percent of people with high school diplomas are male? 16/28 57.14%																					
i) What percent of people with Bachelor's Degrees are female? 51/97 52.58%	j) Given that a person is female, what is the probability that they have a high school diploma? 12/66 .1818 or 18.18%																					

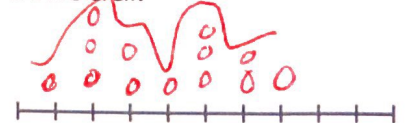
<p>2) Measures of Center and Spread (day 4 notes)</p>	<ul style="list-style-type: none"> Measures of Center: mean, median, mode, Q1, Q3 Measures of Spread: range, IQR, MAD When finding measures of center/spread by hand, order from least to greatest first 	<p>For the data sets below, find the mean, median, mode, Q1, Q3, range, and IQR.</p> <div> <div> <p>a) 30, 27, 24, 32, 40, 26, 37, 31, 27</p> <p>Mean = 30.44</p> <p>Median = 30</p> <p>Mode = 27</p> <p>Q1 = 26.5</p> <p>Q3 = 34.5</p> <p>Range = 40-24 = 16</p> <p>IQR = 34.5-26.5 = 8</p> </div> <div> <p>b) 3, 9, 5, 6, 1, 7, 1, 9</p> <p>Mean = 5.125</p> <p>Median = 5.5</p> <p>Mode = 1</p> <p>Q1 = 2</p> <p>Q3 = 8</p> <p>Range = 8</p> <p>IQR = 6</p> </div> </div>	
	<ul style="list-style-type: none"> The smaller the MAD, the more consistent the data. 	<p>c) Data set A has a MAD of 12.5 and data set B has a MAD of 8.92. Which data set is more consistent?</p> <p>Data Set B</p> <p>d) Class A's last quiz grades had a MAD of 4.25 and Class B's last quiz grades had a MAD of 4.29. Which class was less consistent?</p> <p>Class B</p>	
<p>52+10.5 x 62.5 45-10.5 x 34.5 yes! 25</p>	<ul style="list-style-type: none"> Outliers are data values that are very small or very large compared to all other data values. <p>Q1 45 Q3 52</p> <p>IQR = 7 10.5</p> <p>1.5 · 7 =</p>	<p>Determine whether the data sets below have outliers; if they do, identify them. Then identify and find the best measure of center and spread.</p> <div> <div> <p>e) 50, 45, 49, 53, 51, 52, 40, 46, 48, 58, 25</p> <p>Measure of Center: median</p> </div> <div> <p>f) 0, 12, 2, 6, 18, 3, 5, 11, 36, 9</p> <p>Measure of Center: median</p> <p>12-3 = 9 12+13.5 = 25.5</p> <p>9 · 1.5 = 13.5 x 72.5 = 36</p> <p>yes! 36</p> </div> </div>	
<p>3) Box Plots (day 5)</p>	<ul style="list-style-type: none"> Symmetric and Uniform = use the mean Skewed left/right Use median as best measure of center 	<p>Describe the distributions below. Then identify and find the best measure of center and spread.</p> <div> <div> <p>g)</p>  <p>Measure of Center: median</p> </div> <div> <p>h)</p>  <p>Measure of Center: mean</p> </div> </div>	

- Unimodal = one mode
- Bimodal = two modes

i) Draw a dot plot that is unimodal.

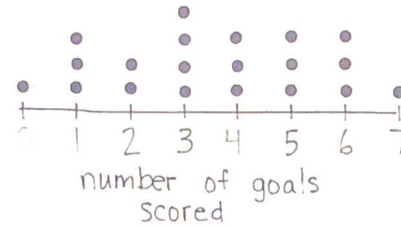


j) Draw a dot plot that is bimodal.



- At least = that # or more
- At most = that # or less

Use the dot plot below to answer the following questions.



m) How many people scored less than 6 goals?

16

k) How many people scored at least 3 goals?

14

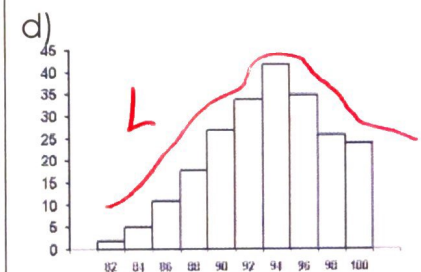
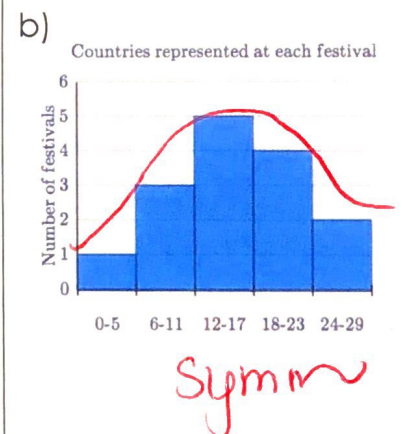
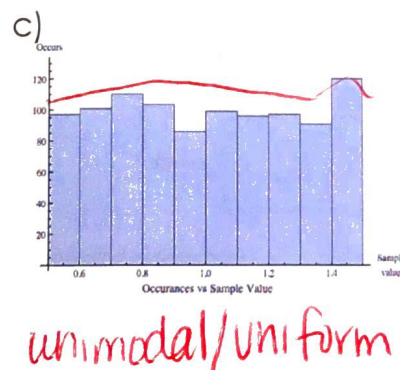
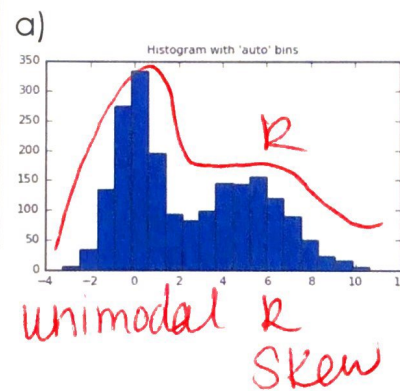
l) How many people scored between 2 and 5 goals?

12

4) Histograms

- Symmetric
- Uniform
- Skewed left/right

Describe the distributions below. Then name the best measure of center and spread.



- Start by sketching a histogram

Describe what the distribution for each scenario below would be.

e) A hard Biology test.

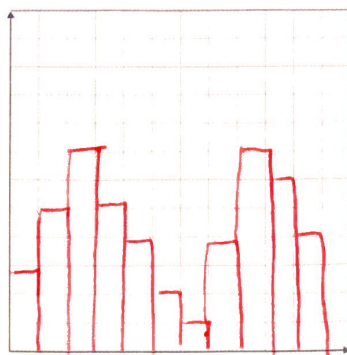


f) Outcomes of rolling a die 1000 times

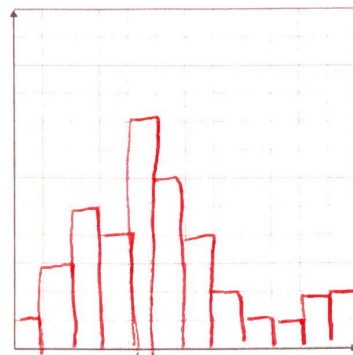
uniform

- Unimodal = one mode
- Bimodal = two modes

g) Draw a histogram that is bimodal.

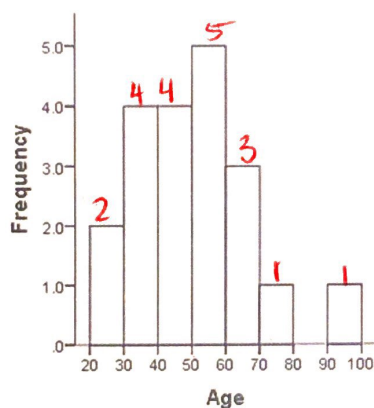


h) Draw a histogram that is unimodal.



- At least = that # or more
- At most = that # or less

Use the histogram below to answer the following questions.



i) How many people were surveyed?

20

j) How many people are over the age of 60?

5

k) How many people are at most 50?

10

n) How many people are between 20 and 40?

6

m) What age group is the most common?

50-60

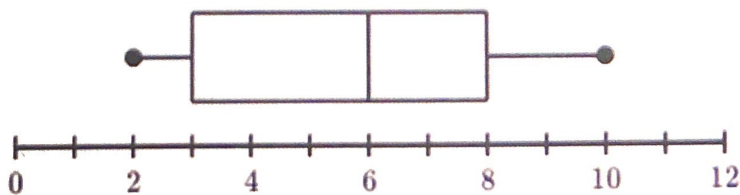
5) Box Plots
5) Box Plots
(day 7)
(day 7)

Five Number Summary:

- Minimum
- Q1
- Median
- Q3
- Maximum

a) Identify the five number summary of the box plot below. Be sure to include units.

Number of eggs laid



min: 2

Q₁ = 3

Q₃ = 8

Q₂ = 6

max = 10

- Range = max - min
- IQR = Q3 - Q1

b) Find the range of the box plot above.

10 - 2

8

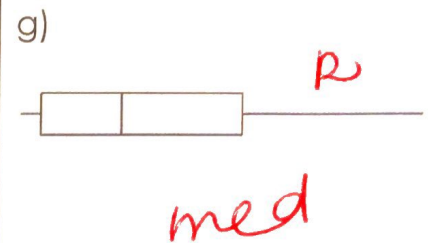
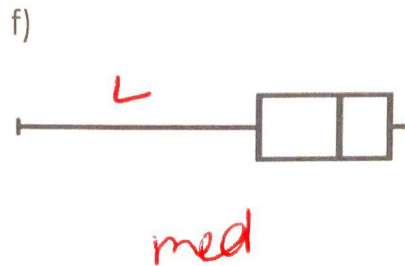
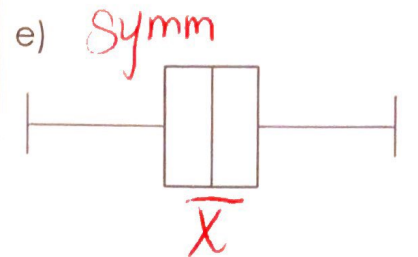
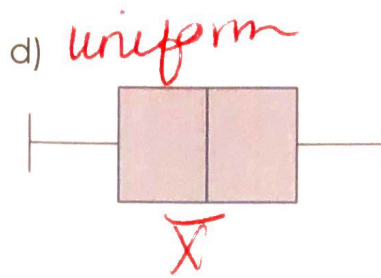
c) Find the IQR of the range above.

8 - 3

5

- Symmetric and Uniform = mean and range
- Skewed left/right = median and IQR

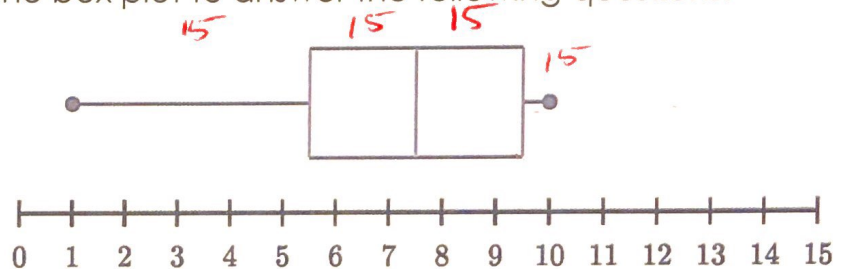
Describe the distributions below. Then identify the best measure of center and spread.



- Min = 0th percentile
- Q1 = 25th percentile
- Med = 50th percentile
- Q3 = 75th percentile
- Max = 100th percentile

Each "section" of the box plot represents 25% of the data.

60 people were surveyed on how many Harrison t-shirts they have. The results can be seen below. Use the box plot to answer the following questions.



h) What percent of people have at least 5.5 shirts?
75%

i) What percent of people have less than 7.5 shirts?
50%

j) How many people have more than 9.5 shirts?
15

k) How many people have between 5.5 and 9.5 shirts?
30

Mixed Review

6) Coach Webb is trying to decide which two of his point guards he wants to start for the first round of play-offs. The data below shows the numbers of points scored by Jace and Tyler from the past six games.

Jace: 11, 11, 6, 26, 6, 12

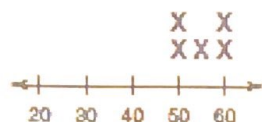
Tyler: 15, 12, 13, 10, 9, 13

Who do you think Coach Webb should select as a starting player and why?

Tyler is more consistent

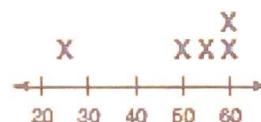
7) Which data set will have the greater mean absolute deviation? Why?

Set 1: 50, 50, 55, 60, 60



Mean = 55

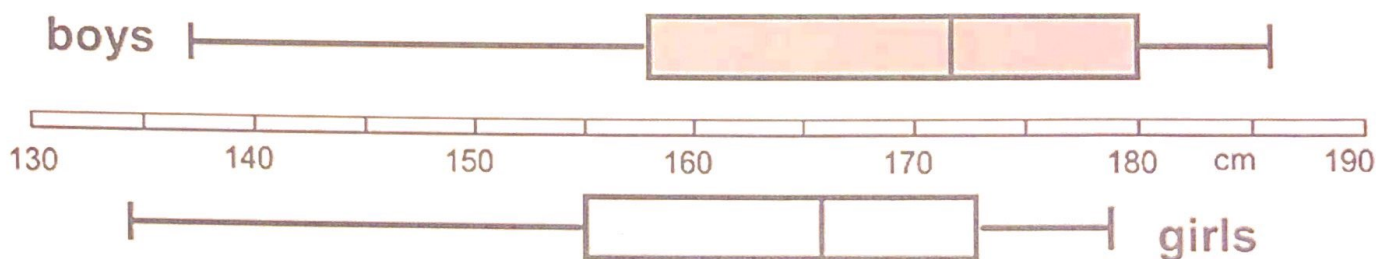
Set 2: 25, 50, 55, 60, 60



Mean = 50

Set 2 - more data farther from mean

8) Determine whether the statements below are true or false based on the box plots below.



a) The girls are taller on average. False

b) The shortest person is a boy. False

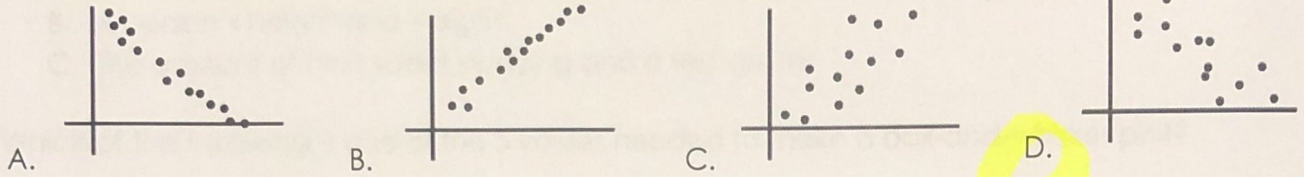
c) The tallest person is a boy. True

d) Both data sets are skewed to the left. True

e) Girls have a smaller IQR. ~~False~~ True

Practice MC/Review

1. Which of the following graphs has a **weak negative** correlation?



2. Which is the best-fitting line for the set of ordered pairs?

(1, 4.2) (2, 3.8) (3, 3.5) (4, 2.7) (5, 2.2)

A. $y = 4.81x - 0.51$

B. $y = -0.99x + 4.81$

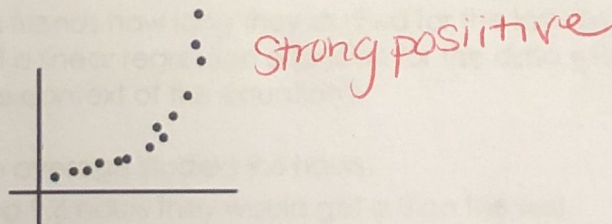
C. $y = -0.51x + 4.81$

D. $y = 0.51x - 0.99$

3. The events x and y have a correlation coefficient of $r = -0.08$. What is the relationship between x and y ?

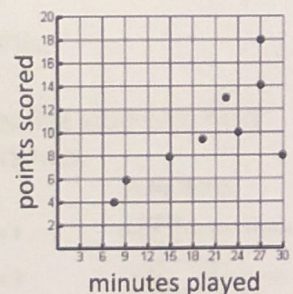
- A. The events have a strong negative linear relationship.
 B. The events have a strong positive linear relationship.
 C. The events have a weak positive linear relationship.
 D. There is very little or no correlation.

4. Given the scatter plot, describe the correlation.



5. The data below represents the amount of time of play and the number of points scored by one player in a recent basketball game. Describe the relationship in context.

*As minutes played ↑
 Points ↑*



Practice MC/Review

6. Between which of the following variables would you expect there to be a **negative** correlation?
- A. The outside temperature and the number of layers of clothing a person wears
 - B. A person's height and weight
 - C. The amount of time spent studying and a test grade
7. Which of the following is one of the 5 values needed to make a box-and-whisker plot?
- A. Mean
 - B. Median
 - C. Mode
 - D. Mean Absolute Deviation
8. Sarah's younger siblings are always bothering her when she does her homework, and she begins to wonder whether other students have the same trouble. She surveys 25 of her fellow juniors and records the number of siblings and the homework grades for each of them. She wants to see if there is a trend in the data she collected.

Which display should she use if she wants to evaluate if there is a correlation?

- A. Box Plot
- B. Histogram
- C. Scatter Plot
- D. Dot Plot

9. The table shows the sizes, in square feet, of a sample of eight houses from a neighborhood. House 5 is a renovated warehouse. What is the **range** of the sizes of the house?

House	1	2	3	4	5	6	7	8
Size	1025	1288	2344	988	12,985	1500	1077	2455

- A. 2957.75
- B. 1394
- C. 1348.5
- D. 11,997

10. George asked five of his friends how long they studied for the last math test and what grade they received. He found a linear regression equation for the data to be $y = 9.6x + 65.8$. What does the **9.6** mean in the context of this equation?

- A. That his friends on average studied 9.6 hours.
- B. That if they studied 9.6 hours they would get a 0 on the test.
- C. That they should study 9.6 hours in order to get an A on the test.
- D. That for every hour they studied their grade would go up 9.6 points.

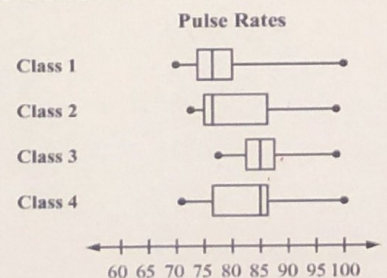
Mr. Murray recorded the pulse rates for each of the students in his classes after the students had climbed a set of stairs. He displayed the results, by class, using the box plots shown.

11. Which class had the **highest upper quartile**?

- A. Class 1
- B. Class 3
- C. Class 2
- D. Class 4

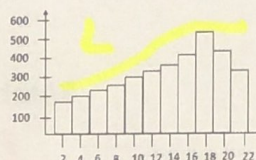
12. Which class had the **lowest IQR**?

- A. Class 1
- B. Class 3
- C. Class 2
- D. Class 4



Practice MC/Review

13. For which pair of variables would you most likely expect a **positive** correlation?
- Driving speed and time it takes to reach a destination
 - Years in school and annual salary**
 - Distance from school and how long it takes to get ready in the morning
 - Daily hours of electronic usage and test scores
14. What type of correlation does the following have? *The age of a person vs. The last four digits of their phone number.*
- Positive Correlation
 - Negative Correlation
 - No Correlation**
 - Skewed Right
15. Which is the best description of the distribution?
- Bimodal
 - Symmetric
 - Skewed Left**
 - Skewed Right



Time (minutes)	18	14	8	30	20
Distance	15	20	5	50	45

According to this data, if it takes a person 40 minutes to get to his/her desk, then how far do they park from their place of employment?

- A. 87 feet B. 48 feet C. 29 feet **D. 74 feet**

Open-Ended Response

Find the mean and the mean absolute deviation of the following data set.
Show ALL work!

18	7	8	6	0	0	2	9
44	33	18	20	26	26	24	17

17. Find the **mean**.

26

18. Find the **mean absolute deviation** (m.a.d.).

6.125

Practice MC/Review

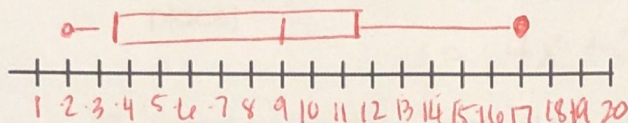
Several taste tests were conducted across the country to determine the consumer ratings of a new sugar-free ice cream. Consumers were asked to rate the level of sweetness on a scale of 1 to 20, with 20 being extremely sweet. The mean ratings of these samples were:

2	11	9	5	17	12	10	3	3	4	15	7	10
---	----	---	---	----	----	----	---	---	---	----	---	----

19. Identify the elements of the Five Number Summary for this data.

$$\text{min } 2 \quad Q_1 = 3.5 \quad Q_2 = 9 \quad Q_3 = 11.5 \quad \text{max} = 17$$

20. Create a Box-and-Whisker Plot to represent this data



21. What is the **Interquartile Range (IQR)** of this data?

$$11.5 - 3.5 = 8$$

22. Are there any outliers for this data? **Show your calculations.**

If so, identify them. If not, justify your conclusion.

no

$$3.5 - 12 = -8.5$$

$$11.5 + 12 = 23.5$$

A group of students were polled to find out how many were planning to major in a scientific field of study in college. The results of the poll are shown in the frequency table below.

	Science	Not Science	Total
Junior	150	210	360
Senior	112	200	312
Total	262	410	672

23. Find the **marginal frequencies** (Fill in the chart)

24. Out of the **juniors**, what **percent** are **not** planning to study a scientific field? (round to nearest percent)

$$210/360 = 58.3\% \quad (58\%)$$

25. Out of the **seniors**, what percent are pursuing a scientific field? (round to nearest percent)

$$112/312 = 36\%$$

$$35.89$$

Practice MC/Review

The data below represents the home values of a certain country from 1900 to 2000, based years **Let the year 1900 be $x = 0$** , and let x represent the number of years since 1900. The values are in thousands

Year	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000
Life exp.	30	35	37	30	35	38	47	53	58	65	70

26. What is the best fitting **linear** regression for the data? (round to the hundredths place)

$$y = .4x + 25$$

27. Based on the data, what will the home value be in **2020**?

\$73,000