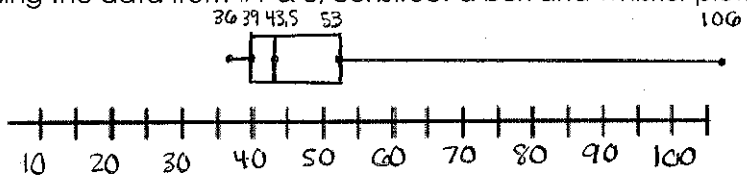


Name: _____

Date: _____

Use the following to review for you test. Work the Practice Problems on a separate sheet of paper.

What you need to know & be able to do	Things to remember	Problem	Problem
Identify the measures of central tendency.	<ul style="list-style-type: none"> • Mean (average) • Median (number that appears most) • Mode 	<p>1. 36, 39, 58, 42, 106, 39, 48, 45</p> <p>36 (min) 39 (Q1) 42 45, 48, 58 (median) 106 (max)</p> <p>mean = 51.63 median = 43.5 mode = 39</p>	<p>2. 50, 55, 60, 58, 62, 57, 68, 51, 63</p> <p>50 (min) 51 (Q1) 55, 57, 58 (median) 60, 62, 63, 68 (max)</p> <p>mean = 58.22 median = 58 mode = none</p>
Identify the measures of spread.	<ul style="list-style-type: none"> • Q1 • Q3 • IQR • Minimum • Maximum • Range • MAD 	<p>3. (Use the same #s from 1)</p> <p>Q₁ = 39 Q₃ = 53 IQR = 53 - 39 = 14 min = 36 max = 106 Range = 106 - 36 = 70</p> <p>SEE MAD work on next page</p>	<p>4. (Use the same #s from 2)</p> <p>Q₁ = 53 Q₃ = 62.5 IQR = 62.5 - 53 = 9.5 min = 50 max = 68 Range = 68 - 50 = 18</p> <p>SEE MAD work on next page</p>
Construct a box-and-whisker plot.	<ul style="list-style-type: none"> • First dot: Min • First Line: Q1 • Middle Line: Median • Third Line: Q3 • Last dot: Max • Outlier: Check Q1 - 1.5(IQR) or Q3 + 1.5(IQR) 	<p>5. Using the data from #1 & 3, construct a box and whisker plot.</p>  <p>6. Are there any outliers? Show your work!</p> <p>39 - 1.5(14) to 53 + 1.5(14) 18 to 74</p> <p>Yes, 106 is an outlier</p>	
Determine if the situation has a positive, negative, or no correlation and if there is causation.	<ul style="list-style-type: none"> • Positive: Both items are increasing/decreasing • Negative: one item increases as the other decreases • No Correlation: No relationship • Causation: One item causes the other. 	<p>7. Practicing Free Throws vs. Free Throw Percentage</p> <p>Positive Correlation and No Causation</p>	<p>8. Colors of the Sky vs. Time of Day</p> <p>No Correlation and No Causation</p>
		<p>9. Weight vs. Amount of Exercise</p> <p>Negative Correlation and Causation</p>	<p>10. Number of Followers on Twitter vs. Number of Friends on Facebook</p> <p>No Correlation and No Causation</p>

$$3. \quad |36 - 51.63| = 15.63$$

$$|39 - 51.63| = 12.63$$

$$|39 - 51.63| = 12.63$$

$$|42 - 51.63| = 9.63$$

$$|45 - 51.63| = 6.63$$

$$|48 - 51.63| = 3.63$$

$$|58 - 51.63| = 6.37$$

$$|106 - 51.63| = 54.37$$

$$\begin{array}{r} + \\ \hline 121.52 \end{array} \text{ SUM}$$

$$\div 8$$

$$\textcircled{15.19} \text{ MAD}$$

$$4. \quad |50 - 58.22| = 8.22$$

$$|51 - 58.22| = 7.22$$

$$|55 - 58.22| = 3.22$$

$$|57 - 58.22| = 1.22$$

$$|58 - 58.22| = 0.22$$

$$|60 - 58.22| = 1.78$$

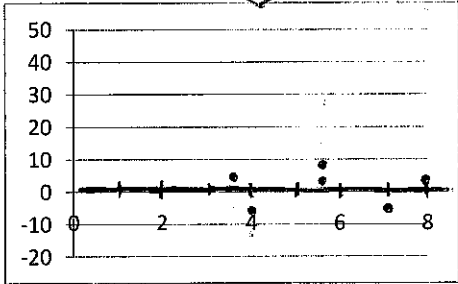
$$|62 - 58.22| = 3.78$$

$$|63 - 58.22| = 4.78$$

$$|68 - 58.22| = 9.78$$

$$\begin{array}{r} + \\ \hline 44.7 \end{array} \text{ SUM}$$

$$\textcircled{4.47}$$

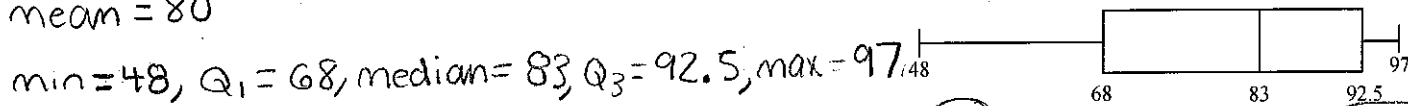
<p>Find the line of best fit.</p>	<ul style="list-style-type: none"> $y = ax + b$ $r =$ correlation coefficient (if close to 0 bad fit; if close to 1 or -1 good fit.) 	<p>11. Determine the line of best fit. Is this model a good fit for the data?</p> <p>$r = -0.97$ <i>Very strong negative correlation</i></p> <table border="1"> <tr> <td>Price</td> <td>4.00</td> <td>5.50</td> <td>3.50</td> <td>8.00</td> <td>5.50</td> <td>7.00</td> </tr> <tr> <td># of Sandwiches</td> <td>68</td> <td>55</td> <td>85</td> <td>22</td> <td>64</td> <td>28</td> </tr> </table> <p>$y = -13.72x + 130.28$ <i>YES it's a good fit.</i></p>	Price	4.00	5.50	3.50	8.00	5.50	7.00	# of Sandwiches	68	55	85	22	64	28														
Price	4.00	5.50	3.50	8.00	5.50	7.00																								
# of Sandwiches	68	55	85	22	64	28																								
<p><u>Scale</u> ←</p> <p>.00-.19 <i>very weak</i> .20-.39 <i>weak</i> .40-.59 <i>moderate</i> .60-.79 <i>strong</i> .80-1.0 <i>very strong</i></p> <p>Construct a residual plot and determine if the model is a good fit or not.</p>	<p>Correlation Coefficient</p> <ul style="list-style-type: none"> Find the predicted values. Actual minus predicted Plot the residuals If it makes a pattern it is NOT a good fit. No pattern is a good fit. 	<p>12. Using the line of best fit from #11, construct a residual plot.</p> <table border="1"> <tr> <th>Price</th> <th>Actual</th> <th>Predicted</th> <th>Residuals</th> </tr> <tr> <td>4.00</td> <td>68</td> <td>75.4</td> <td>-7.4</td> </tr> <tr> <td>5.50</td> <td>55</td> <td>54.82</td> <td>0.18</td> </tr> <tr> <td>3.50</td> <td>85</td> <td>82.26</td> <td>2.74</td> </tr> <tr> <td>8.00</td> <td>22</td> <td>20.52</td> <td>1.48</td> </tr> <tr> <td>5.50</td> <td>64</td> <td>54.82</td> <td>9.18</td> </tr> <tr> <td>7.00</td> <td>28</td> <td>34.24</td> <td>-6.24</td> </tr> </table> <p>GRAPH</p>  <p>GRAPH</p> <p><i>This is line of best fit above IS good because there is no clear pattern.</i></p>	Price	Actual	Predicted	Residuals	4.00	68	75.4	-7.4	5.50	55	54.82	0.18	3.50	85	82.26	2.74	8.00	22	20.52	1.48	5.50	64	54.82	9.18	7.00	28	34.24	-6.24
Price	Actual	Predicted	Residuals																											
4.00	68	75.4	-7.4																											
5.50	55	54.82	0.18																											
3.50	85	82.26	2.74																											
8.00	22	20.52	1.48																											
5.50	64	54.82	9.18																											
7.00	28	34.24	-6.24																											
<p>SKIP</p> <p>Find the exponential regression model.</p>	<ul style="list-style-type: none"> $y = a(b)^x$ $r =$ correlation coefficient (if close to 0 bad fit; if close to 1 or -1 then good fit.) 	<p>13. Determine the exponential regression model. Is this model a good fit for the data?</p> <table border="1"> <tr> <td>Year</td> <td>0</td> <td>2</td> <td>4</td> <td>7</td> </tr> <tr> <td>Revenue</td> <td>3</td> <td>4</td> <td>11</td> <td>25</td> </tr> </table>	Year	0	2	4	7	Revenue	3	4	11	25																		
Year	0	2	4	7																										
Revenue	3	4	11	25																										
<p>Construct a probability table.</p>	<ul style="list-style-type: none"> <u>Joint Probability</u>: Individual Cell/Table Total <u>Marginal Probability</u>: Row or Column Total/Table Total <u>Conditional Probability</u>: Individual Cell/Row or Column Total 	<p>Complete the table to answer the following questions.</p> <table border="1"> <tr> <td></td> <td>Football</td> <td>Basketball</td> <td>Soccer</td> <td></td> </tr> <tr> <td>Males</td> <td>48</td> <td>35</td> <td>17</td> <td>100</td> </tr> <tr> <td>Females</td> <td>22</td> <td>38</td> <td>40</td> <td>100</td> </tr> <tr> <td></td> <td>70</td> <td>73</td> <td>57</td> <td>200</td> </tr> </table> <p>14. What is the probability that a randomly chosen female likes soccer? $\frac{40}{200} = 0.2$ (same as 20%)</p> <p>15. What is the probability that someone likes basketball? $\frac{73}{200} \approx 0.37$ (same as 37%)</p> <p>16. Given that a person likes football, what is the probability they are male? $\frac{48}{70} \approx 0.69$ (same as 69%)</p>		Football	Basketball	Soccer		Males	48	35	17	100	Females	22	38	40	100		70	73	57	200								
	Football	Basketball	Soccer																											
Males	48	35	17	100																										
Females	22	38	40	100																										
	70	73	57	200																										

Name: _____ Date: _____

NOTE: You may use the "STAT" feature on your calculator.

- Identify the Five-Number Summary number for the data of Johnny's test scores and draw the Box & Whisker plot.

mean = 80
 92, 96, 97, 83, 92, 58, 93, 88, 77, 48, 65, 80, 71



What is the range? IQR? MAD? Range: $97 - 48 = 49$ IQR: $92.5 - 68 = 24.5$

Are there any outliers in the data set? MAD: $162 \div 13 = 12.46$
 $68 - 1.5(24.5)$ to $92.5 + 1.5(24.5)$
 31.25 to 129.25 **NO outliers.**

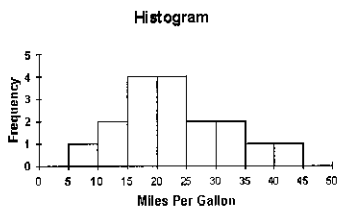
- The table gives the low temperatures in Chicago on eight randomly selected winter days. Which measure of central tendency probably gives the LEAST ACCURATE prediction of a "typical" low temperature on a Chicago winter day?

Chicago Lows							
17	25	28	12	16	55	18	22

mean = 24.125, median = 20
 $16.5 - 1.5(10)$ to $26.5 + 1.5(10)$
 1.5 to 41.5
 The mean is greatly affected by 55, which is an outlier.

- Describe the shape of the distribution (skew). Estimate the mode. Give a description of what this data could be representing.

skewed right
 mode = 10-15, 25-30, and 30-35



ex. car types that have the same mpg rating

- Construct a frequency table from the following information:
 A survey of 200 9th and 10th graders was given to determine what their favorite subject was. 72 said Math (50 which were freshmen), 38 said Social Studies (20 which were sophomores), and 40 freshmen and 50 sophomores said PE was their favorite.

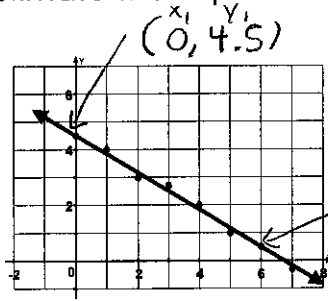
	Math	SS	PE	Total
9th	50	18	40	108
10th	22	20	50	92
Total	72	38	90	200

- Based on your table above, answer the following questions:
- Find the marginal relative frequencies (total row/column) 0.36, 0.19, 0.45, 0.54, 0.46
 - Find the joint relative frequencies (inside boxes) (TOP) 0.25, 0.09, 0.20 (BOTTOM) 0.11, 0.10, 0.25
 - What is the percent that a student surveyed is a freshman? $108/200 = 0.54$ (54%)
 - What is the probability that a student surveyed likes Math? $72/200 = 0.36$
 - P(9th | Math) means (in 9th grade and likes math) $50/72 \approx 0.69$

marginal
 conditional

5. For the given scatter plot, estimate the equation for the line of best fit and the correlation coefficient.

The line of best fit is a negative sloped line but the data is close to line so I would estimate r to be between -0.85 to -0.95 .

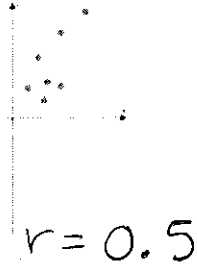
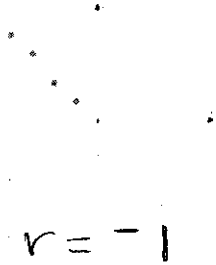


SLOPE FORMULA $m = \frac{y_2 - y_1}{x_2 - x_1}$

$m = \frac{0.5 - 4.5}{6 - 0}$
 $m = -0.67$

$y = -0.67x + 4.5$

6. Estimate the correlation coefficient for the following graphs.



7. Determine if the following situations represent a positive, negative, or no correlation.

- a) Number of hours studying for the SAT and your score. *positive*
- b) The distance you drive and the number of stars in the sky. *none*
- c) The temperature and the length of daylight hours for the day *positive*

8. Tell whether the following situations are causation: (yes or no)

- a) The number of boats on Lake Allatoona and the number of cars on the street *No*
- b) The hours you work and the money you make *Yes*
- c) The time spent studying and the A on the test *Yes*

9. The following table shows a person study hours versus their test scores.

Hours studied (x)	2	5	1	0	4	2	3
Grade on test (y)	77	92	70	63	90	75	84

- a) Use your calculator to find the line of best fit for the data above. $y = 6.09x + 63.93$
- b) What is the value of r ? 0.99 *strong* Is this a good fit? *Yes*
- c) Use the equation to predict the test grade for someone who studies 5 hours. 94.38
- d) Use the equation to predict the test grade for someone who studies 4.5 hours. 91.33