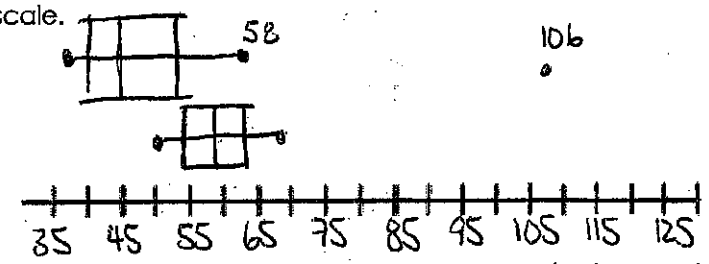
GSE Algebra I | **Unit 6: Describing Data** *Study Guide* | **Name:**

What you need to know & be able to do	Things to remember	Problem	Problem
Identify the measures of central tendency.	• Mean • Median • Mode	1? 36, 39, 58, 42, 106, 39, 48, 45 Mean ($\bar{x}$): 51.63 Median: 43.5 Mode: 39	**2.** 50, 55, 60, 58, 62, 57, 68, 51, 63 Mean ($\bar{x}$): 58.22 Median: 58 Mode: none
Identify the measures of spread.	• Q1 • Q3 • IQR • Minimum • Maximum • Range • MAD	**3.** (Use the same #s from 1) Q1: 39 Q3: 53 IQR: 53-39=14 Min: 36 Max: 106 MAD: 15.19 Range: 106-36 =70	**4.** (Use the same #s from 2) Q1: 53 Q3: 62.5 IQR: 62.5-53=9.5 Min: 50 Max: 68 MAD: 7.10 Range: 68-50 =18
Construct a box-and-whisker plot.	• First dot: Min • First Line: Q1 • Middle Line: Median • Third Line: Q3 • Last dot: Max • Outlier: Q1 – 1.5(IQR) Q3 + 1.5(IQR)	**5.** Using the data from #1 and 2, give the 5-number summaries. Remember to label the type of statistic. **6.** Construct 2 box and whisker plots. Remember to label your scale. **7.** Are there any outliers? Show your work! Yes! Q3+1.5IQR 53+1.5(14)=74 106 is an outlier. **8.** Which data set had the higher median? data set 2 **9.** Which data set has the greater IQR? data set 1 **10.** Which data set had the lower maximum? data set 2 **11.** In what span of numbers did the top 50% of data fall in data set 1? 43.5 - 106 **12.** How would you describe the shape of data set 2? Symmetric, unimodal	

Statistic	Min	Q1	Med.	Q3	Max
Data 1	36	39	43.5	53	106
Data 2	50	53	58	62.5	68



Construct a probability table.

- Joint Probability: Individual Cell/Table Total
- Marginal Probability: Row or Column Total/ Table Total
- Conditional Probability: Individual Cell/Row or Column Total

Complete the table to answer the following questions.

	Football	Basketball	Soccer	Total
Males	48	35	17	100
Females	22	38	40	100
Total	70	73	57	200

13. What is the probability that a randomly chosen female likes soccer? Is this conditional, marginal, or joint frequency?

$\frac{40}{100} = 40\%$ conditional

14. What is the probability that someone likes basketball? Is this conditional, marginal, or joint frequency?

$\frac{73}{200} = 36.50\%$ marginal

15. Given that a person likes football, what is the probability they are male? Is this conditional, marginal, or joint frequency?

$\frac{48}{70}$ = conditional

Find the line of best fit.

- $y = ax + b$
- r = correlation coefficient (if close to 0 bad fit; if close to 1 or -1 good fit.)

y / x						
Price	4.00	5.50	3.50	8.00	5.50	7.00
# of Sandwiches	68	55	85	22	64	28

16. Determine the line of best fit. $y = -0.07x + 9.25$ $r = -.968$

Is this model a good fit for the data?

yes, very strong positive correlation

A. How many sandwiches would you need to buy for them to be 2.00 each? Is this interpolation or extrapolation?

$2.00 = -.07x + 9.25$ $x = 103.6$

$-7.25 = -.07x$ You would need to buy at least 104 sandwiches

B. How many sandwiches would you need to buy for them to be 5.00 each? Is this interpolation or extrapolation?

$5.00 = -.07x + 9.25$ $x = 60.7$

$-4.25 = -.07x$ You would need to buy at least 61 sandwiches

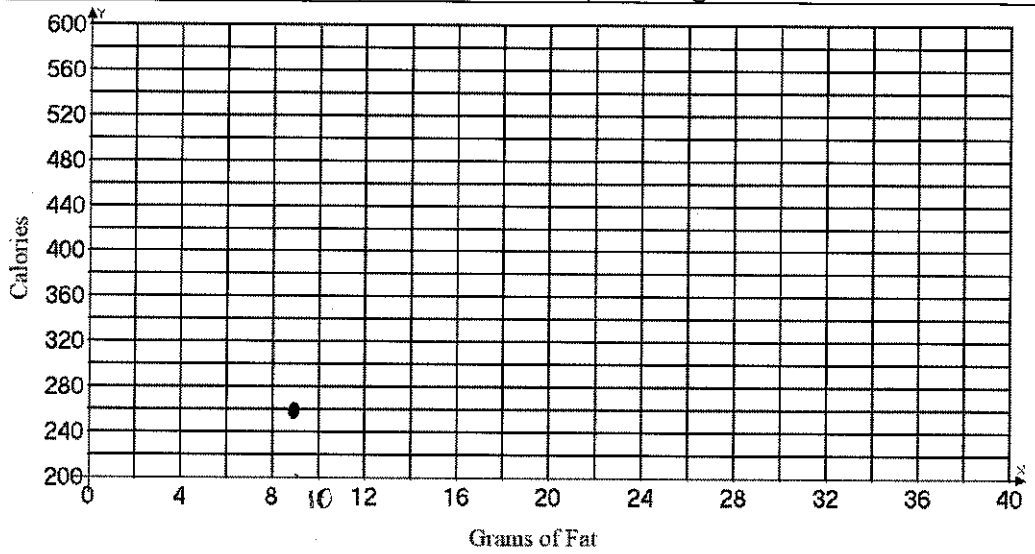
C. What would you expect the price per sandwich to be if you bought 10?

$y = -.07(10) + 9.25 =$ \$8.55 per sandwich

17.

Sandwich	Total Fat (g)	Total Calories
Hamburger	9	260
Cheeseburger	13	320
Quarter Pounder	21	420
Quarter Pounder with Cheese	30	530
Big Mac	31	560
Arch Sandwich Special	31	550
Arch Special with Bacon	34	590
Crispy Chicken	25	500
Fish Fillet	28	560
Grilled Chicken	20	440
Grilled Chicken Light	5	300

Create a scatterplot using the table above.



a. Calculate the correlation coefficient. What does it mean in terms of the data?

$r = .975$ there is a very strong positive correlation between fat and calories

b. Calculate the line of best fit. Describe the meaning of the parameters.

$y = 11.73x + 193.85$

$a = 11.73$ means for every gram of fat there's 11.73 calories

$b = 193.85$ - y-int, for 0 fat it still has 193.85 calories

c. How ~~many~~ calories would a sandwich have if it only had 3g of fat? Is this interpolation or extrapolation?

$y = 11.73(3) + 193.85 = 229.04$ calories

d. How many calories would a sandwich have if it has 23g of fat? Is this interpolation or extrapolation?

$y = 11.73(23) + 193.85 = 463.64$ calories

e. How many grams of fat would a sandwich have if it has 750 calories?

$750 = 11.73x + 193.85$

$556.15 = 11.73x$

$x = 47.41$ g of fat

Determine if the situation has a positive, negative, or no correlation and if there is causation.

- Positive: Both items are increasing/decreasing
- Negative: one item increases as the other decreases
- No Correlation: No relationship
- Causation: One item causes the other.

For #18-21, describe whether you think the two variables would have a positive, a negative, or no correlation. If there is a correlation, is there also causation?

18. Practicing Free Throws vs. Free Throw Percentage

Positive
yes, causation

19. Colors of the Sky vs. Time of Day

no correlation

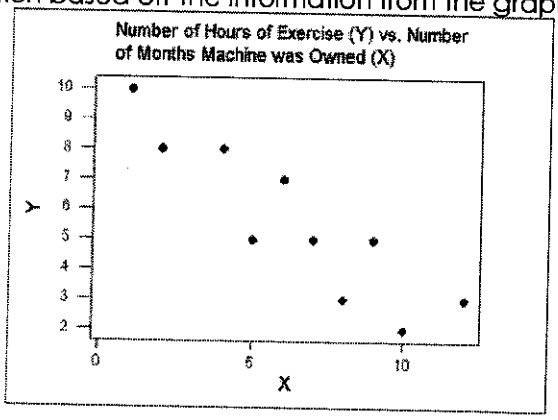
20. Weight vs. Amount of Exercise

negative (usually)
yes, causation

21. Number of Followers on Twitter vs. Number of Friends on Facebook

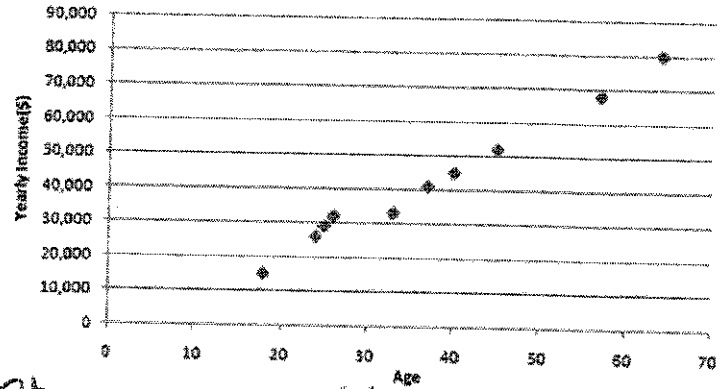
positive
no causation

22. Estimate the correlation coefficient and describe the correlation based off the information from the graph.



Weak negative $r \approx -.5$ $\approx -.4$

23. Estimate the correlation coefficient and describe the correlation based off the information from the graph.



Strong positive $r \approx .8$ $\approx .9$