Module 8H Study Guide

Name: Key

Use the following categorical data table to answer questions #1 – 5

The Pew Research Center asked a random sample of 2024 adult cell phone owners from the United States which type of cell phone they own: iPhone, Android, or other (including non-smart phones). Here are the results, broken down by age category.

	18-34	35-54	55+	Total
iPhone	169	171	127	467
Android	214	189	100	503
Other	134	277	643	1054
Total	517	637	870	2024

	18-34	35-54	55+	total
iphone	8.3%	8.4%	6.3%	23.1%
Android	10.6%	9.3%	4.9%	24.9%
other	6.6%	13.7%	31.8%	52%
total	25.5%	31.4%	43.%	100%

1. Explain what the number 127 represents in this table as a joint frequency.
   127 people have an iphone and are 55+ years old.

2. In the space to the right of this frequency table, create a relative frequency table by entire table.

3. What percent of 55+ people use an iPhone? Show how to calculate this from **each** table.
   $\frac{\text{iphone}}{55+} = \frac{127}{870} = 14.6\%$ $\frac{6.3\%}{43\%} = 14.6\%$

4. Write a conditional frequency statement about the "other" phone type.
   Of the other phone type, 31.8% are 55+ owners.

5. True or False (explain your answer):

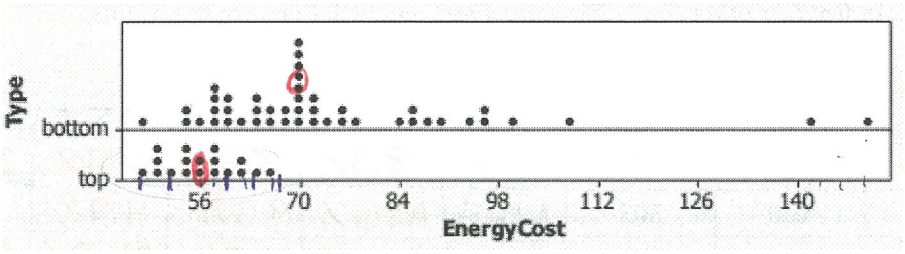
   a. Approximately 42.5% of Android users are in the age group 18-34.
   $\frac{18\text{-}34}{\text{Android}} = \frac{214}{503} = 42.5\%$ TRUE

   b. Approximately 8.4% of 35-54 year olds use an iPhone.
   $\frac{\text{iphone}}{35\text{-}54} = \frac{171}{637} = 26.8\%$ FALSE

   c. 35-54 year olds tend to have an iPhone more than 18-34 year olds.
   $\frac{\text{iphone}}{35\text{-}54}$ vs $\frac{\text{iphone}}{18\text{-}34}$
   $\frac{171}{637}$ $\frac{169}{517}$
   26.8% 32.7%
   FALSE.
   32.7% of 18-34 yr olds have an iphone, whereas, only 26.8% of 35-54 yr olds have an iphone.

Use the following information and dot-plot to answer question #6-9:

How do the annual energy costs (in dollars) compare for refrigerators with top freezers and refrigerators with bottom freezers? The data below is from the May 2010 issue of Consumer Reports.



6. Which type of freezer has more variability in its annual energy cost? Explain how you can tell.

The bottom freezer has more variability because it costs between $54 – $144.

7. Describe the shape of this distribution for each type of freezer.

Bottom: Skewed Right, unimodal
top: trimodal, fairly normal.

8. For each freezer, circle the dot that would represent the median cost.

bottom total: 46
top is total

9. Without doing any calculations, which type of freezer would have a mean that is larger than its median? Explain.

bottom freezer would have a larger mean because it has two large outliers.

John has taken 5 tests so far this semester in his math class. The mean of these tests is 88 and standard deviation is 5.

10. If on his sixth test he earns a 88, what would happen to the mean and to the standard deviation? Explain.

The mean would stay the same and the standard deviation would decrease.

11. If on his sixth test, he instead earns a 50, what would happen to the mean and to the standard deviation? Explain.

The mean would decrease, standard deviation would increase because 50 is so far below the original mean.

12. List the possible test scores that John could earn on his sixth test that would increase the mean but decrease the standard deviation.

He can earn anything greater than 88 but less than 5 units away. 89, 90, 91, 92

Tania's five test scores are: 78, 86, 60, 94, 100.

13. Find the mean and standard deviation of Tania's test scores.

$\mu = \frac{60+78+86+94+100}{5}$, $\mu = 83.6$

$\sigma = 13.9$

$\sigma = \sqrt{\frac{(60-83.6)^2+(78-83.6)^2+\ldots+(100-83.6)^2}{5}}$

Use the following data about the amount of fat in different McDonald's sandwiches to answer questions #14 - 17:

Chicken or Fish Sandwiches	
Sandwich Name	**Fat (g)**
Filet-O-Fish	19
McChicken	16
Premium Crispy Chicken Classic Sandwich	22
Premium Crispy Chicken Club Sandwich	33
Premium Crispy Chicken Ranch Sandwich	27
Premium Grilled Chicken Classic Sandwich	9
Premium Grilled Chicken Club Sandwich	20
Premium Grilled Chicken Ranch Sandwich	14
Southern Style Crispy Chicken Sandwich	19

total: 9

Beef Sandwiches	
Sandwich Name	**Fat (g)**
Big Mac	29
Cheeseburger	12
Daily Double	24
Double Cheeseburger	23
Double Quarter Pounder with Cheese	43
Hamburger	9
McDouble	19
McRib	26
Quarter Pounder Bacon and Cheese	29
Quarter Pounder Bacon Habanero Ranch	31
Quarter Pounder Deluxe	27
Quarter pounder with Cheese	26

total 12

14. Find the 5-number summary for both "Chicken or Fish" and "Beef" sandwiches.

9, 14, 16, 19, 19, 20, 22, 27, 33
min Q1: 15 median Q3 = 24.5 max

5 # summary: 9, 15, 19, 24.5, 33

9, 12, 19, 23, 24, 26, 26, 27, 29, 29, 31, 43
min Q1 = 21 Median 26 Q3 = 29 max

5 # summary: 9, 21, 26, 29, 43

15. Does either group contain any outliers? Use the 1.5 x IQR rule. Show what you found as the "magic fences" for each group.

lower: $15 - 1.5(24.5 - 15) = 0.75$

upper: $24.5 + 1.5(24.5 - 15) = 38.75$

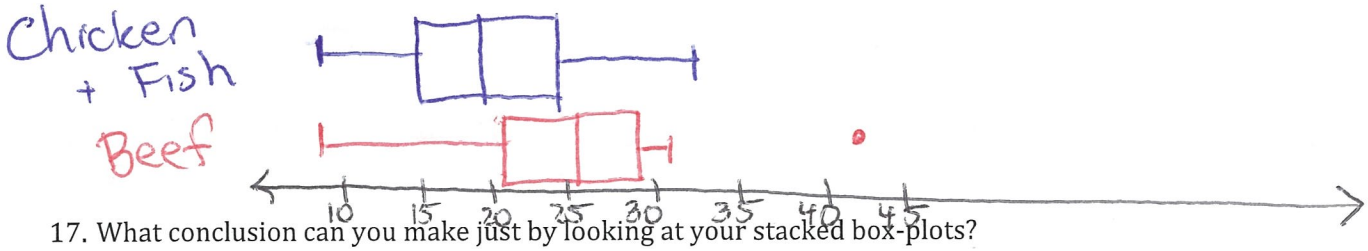
★ No outliers

lower: $21 - 1.5(29 - 21) = 9$

upper: $29 + 1.5(29 - 21) = 41$

★ 43 is an outlier

16. Construct stacked box-plots for these two groups (this means use one number line, draw both box-plots above that number line). Be sure to label the box-plots by group.



17. What conclusion can you make just by looking at your stacked box-plots?

Beef has more variability in its amount of fat.

Chicken + Fish has a normal distribution.

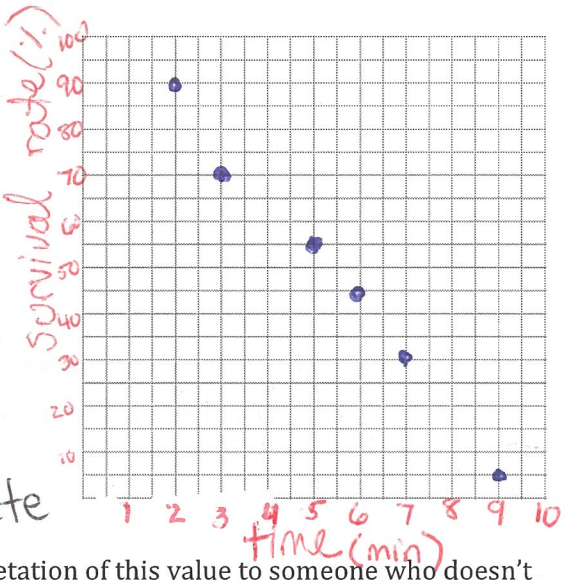
Use the following information for problems #18 – 21:

Studies have shown that people who suffer sudden cardiac arrest have a better chance of survival if a defibrillator shock is administered very soon after cardiac arrest. How is survival rate related to the time between when cardiac arrest occurs and when the defibrillator shock is delivered? The following data give:

x =time from cardiac arrest and shock (min) y = survival rate (as a %)

x	2	3	5	6	7	9
y	90	70	55	45	30	5

18. In the space to the right, create a scatterplot of these data.



19. Describe the correlation between x and y.

There is a strong negative linear relationship between time from cardiac arrest + shock and survival rate.

20. The linear regression equation for this data is:
$y = -11.5x + 110.5$
What does the slope mean in this context?

For each minute after time from cardiac arrest + shock, the survival rate decreases by 11.5%.

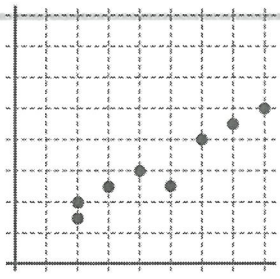
21. The correlation coefficient, $r = -0.99$. Explain the interpretation of this value to someone who doesn't know any statistics, using the scatter plot and your answer to #19.

r = -.99 means that the data points have a strong linear relationship with a negative slope.

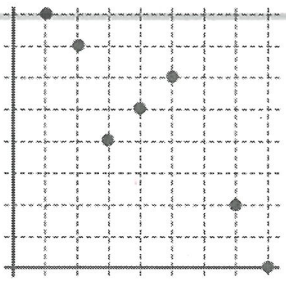
22. Match the graph to the correlation coefficient, r.

Correlation	0.72	−0.90	0.96	−0.42
Graph	C	B	A	D

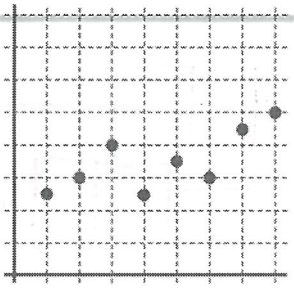
Graph A



Graph B



Graph C



Graph D