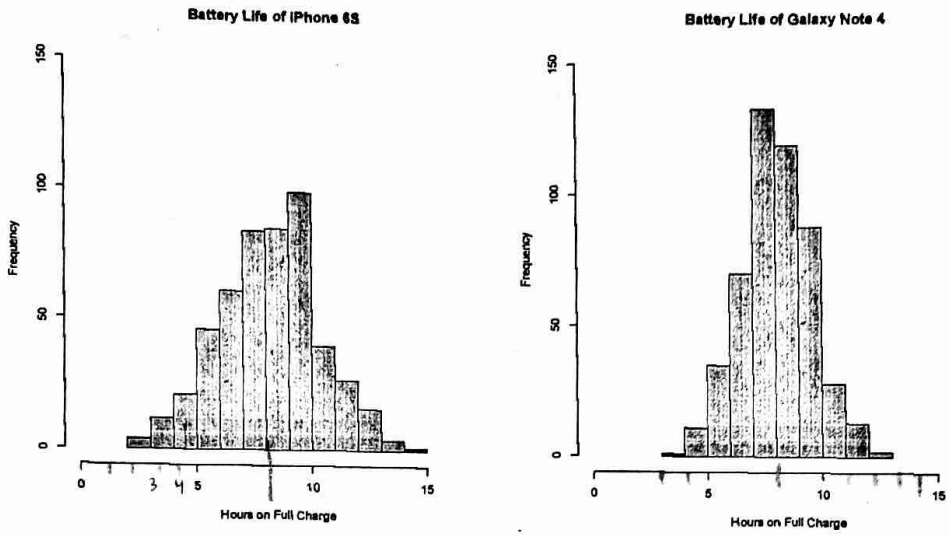
1. The following histograms are the battery lifespans from a full charge of a random sample of 500 iPhone 6S (left) and 500 Galaxy Note 4 (right). These histograms use the left hand rule so that the bin to the left of 5 corresponds to the number of phones that lasted 5 hours on a full charge. Both phones have a mean lifespan of exactly 8 hours.

$\bar{x} = 8$

8



(a) Which phone has the larger spread for battery lifespan? Explain in two or fewer sentences using specific values to measure spread to make your comparison. Be concise. [3]

2

iPhone 6S has the larger spread for battery life because the data is more spread from the center. The range for the iPhone is 15 − 3 = 12 hours and the range for the Galaxy is 13 − 4 = 9 hours.

(b) Assume the Empirical Rule holds true for both histograms. Between what two values does 68% of the data occur for the Galaxy Note 4? Estimate a single value for the standard deviation to complete this question. Show your work [3]

3

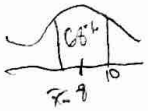
~~$\frac{15-3}{6} = \frac{12}{6} = 2$~~ ~~$8-2=6$~~ ~~$8+2=10$~~

$\frac{13-4}{6} = \frac{9}{6} = \frac{3}{2} = 1.5$ $8-1.5=6.5$ $8+1.5=9.5$

By the empirical rule, the standard deviation is approximately 1.5 for the Galaxy SD, ∴ approximately 68% of the data for the Galaxy lie between 6.5 and 9.5 hours for battery life.

(c) I am planning on buying one phone based simply off the time each phone stays alive. I plan on being out in LA everyday without recharging and the only criterion on picking a phone is that I want my phone to last 10 hours or more. Which of the phones do I prefer? Explain assuming the Empirical Rule holds true for both phones and/or using your knowledge of z-scores. Be concise. [3]

3



iphone $\frac{10-8}{2} = 1$

Gal. $\frac{10-8}{1.5} = 1.33$

$100\% - (50\% + 34\%) = 16$

You would prefer the iphone because about 16% of the data for iphones last 10 hours or more. However, the z-score for 10 hours of a Galaxy is 1.33, meaning that less than about 16% of the data for the Galaxy last 10 hours or more.

2. The following regression line looks at the variables of distance between the departure and arrival cities, and price of an airline ticket between the two cities. The mean price of a ticket is \$384.50 and the mean distance of a flight is 1525 miles. The standard deviation for price is \$57 and the standard deviation for distance is 215 miles.

8

$$\hat{Price} = 49 + 0.22 \cdot Distance$$

(a) Find the correlation for the linear relationship between distance and price. Also report the R^2 for the model. [3]

$m = r \frac{s_y}{s_x}$ $\quad 0.22 = r \frac{57}{215}$ $\quad r \approx 0.83$ $\quad R^2 = (0.83)^2 = 0.69$

3 The correlation between distance and price is 0.83.
R^2 is 0.69

(b) Interpret the slope in context in one sentence. [1]

1 For each additional mile, we expect the price to increase on average by \$0.22.

(c) The distance between Los Angeles and New York City is 2448.3 miles and the flight costs \$222. What is the residual for the flight from LA to NYC?[2]

$\hat{Price} = 49 + 0.22(2448.3) = 587.63$

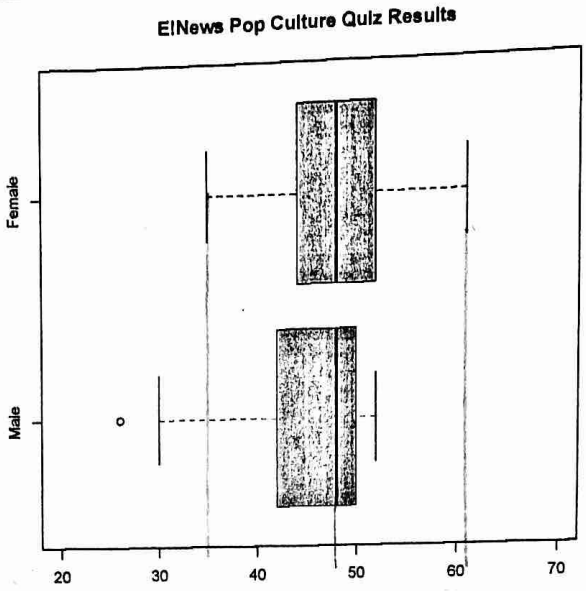
2 residual = observed $y - \hat{y} = \$222 - \$587.63 = -365.63$

The residual is -365.63

(d) In this scenario, would you prefer a positive or negative residual? Explain why in one sentence. [2]

2 You would prefer a negative residual, because this means you over estimated the cost and actually need to pay less than you had anticipated.

Staff members from E! News went out to Hollywood Boulevard to administer a 70 question pop culture quiz. Over a week the staff members were able to get 118 females and 97 males to take the quiz. The results for both the females and males are summarized in the box plots to the right. Each of the distributions is unimodal. Use this information to answer questions 1 and 2.



1. Which group has the higher mean score on the pop culture quiz?

   (a) Females
   (b) Males
   (c) They are the same.
   (d) Not enough information to tell.

2. Assume the Empirical Rule holds true for the female box plot. Between what two values does 95% of the data occur? You will need to estimate a a value for the standard deviation to complete this question.

   $\frac{61-35}{6} = \frac{26}{6} \approx 4.33$

   $48 - 2(4.3) = 39.4$

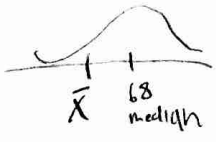
   $48 + 2(4.3) = 56.6$

   (a) 44 and 52
   (b) 35 and 61
   (c) 39 and 57
   (d) 30 and 70

3. Researchers conducted a study and determined that students who participate in sports are happier than students who do not. What can we conclude about this study?

   (a) Sports participation cause students to be happier.
   (b) Sports participation is only associated with students being happier.
   (c) Researchers cannot conclude anything about sports participation and the happiness of students because of possible confounding variables.

4. The distribution of a sample of 300 headstones from the Hollywood Forever cemetery is heavily left skewed. The median age of death in this sample is 68 years old. Which best describes the z-score for the median age of death in the Hollywood Forever cemetery?

   $z = \frac{0 - \bar{x}}{s}$

   (a) The z-score of the median age of death is positive.
   (b) The z-score of the median age of death is negative.
   (c) The z-score of the median age of death is zero.
   (d) There is not enough information to determine the anything about the z-score of the median age of death.

9. If you were comparing the differences between the favorite colors of males and females in this class, what is the best way to visualize the data?

(a) Scatterplot
(b) Segmented bar chart
(c) A single boxplot
(d) Side-by-side boxplot

A group of 500 patients who suffer from hypothyroidism, a condition in which your thyroid does not produce enough of certain hormones, were asked to participate in a study to determine the effectiveness of a new medication. The patients were randomly divided into two groups, one that was given the actual medication, and one that received a placebo pill. The results of the study are below. Use this information to answer questions 10-13.

	Medication	Placebo
Symptoms improved	205	140
Symptoms did not improve	65	90

10. What percent of patients who took the medication had improved symptoms?

(a) 41%
(b) 54%
(c) 65.2%
(d) 75.9%

$\frac{205}{205+65} = \frac{205}{270} = 0.759$

11. What percent of patients took the medication and had improved symptoms?

(a) 41%
(b) 54%
(c) 65.2%
(d) 75.9%

$\frac{205}{500} = 0.41$

% of patients who took placebo & had improved

$\frac{140}{250} = 0.56$

12. Was the new medication effective in treating hypothyroidism?

(a) Yes, a higher percent of patients who took the medication had improved symptoms than the patients who took the placebo.
(b) Yes, both groups had more patients with improved symptoms.
(c) No, the patients who took the placebo also had improved symptoms.
(d) No, this was not a controlled experiment.

13. Can we conclude that the improved symptoms were caused by the new medication?

(a) Yes, this is a controlled experiment. Since a higher percent of patients who took the medication had improved symptoms, we can conclude causation.
(b) Yes, this is a controlled experiment. We can always conclude causation with a controlled experiment.
(c) No, even though this is a controlled experiment, there was no difference between the treatment and control groups, so we cannot conclude causation.
(d) No, even though this is a controlled experiment, there might be a confounding factor since the placebo group had improved symptoms too.