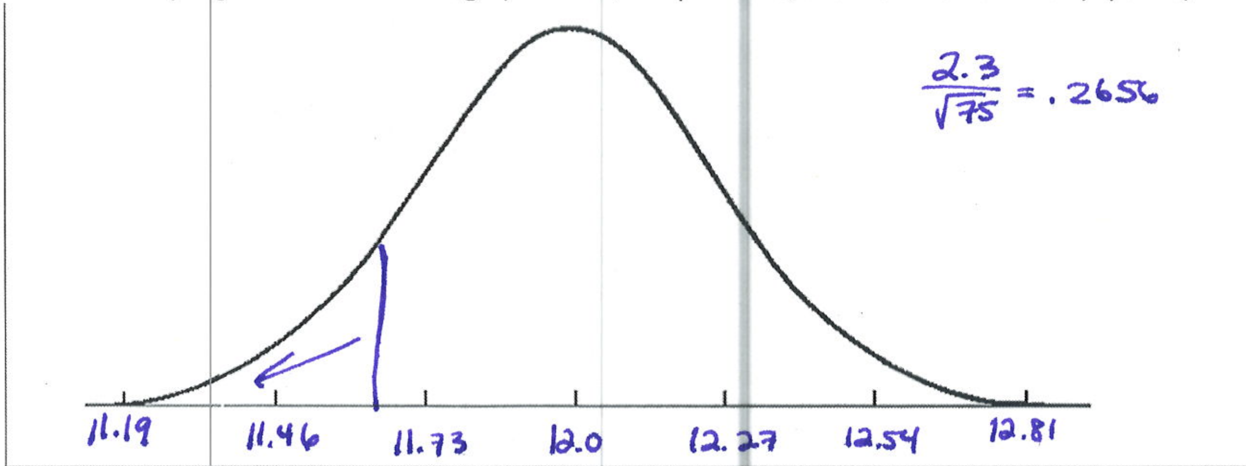
MAT 223, Final Exam, Spring 2017 Name KEY

Instructions: Show all work. Use exact answers unless specifically asked to round. Explain thoroughly using complete sentences. If you use your calculator to perform statistical tasks, say which command/operations you are using and what you entered into your calculator, and what you got back to show work. If you do not show work and the answer is incorrect, no credit will be awarded.

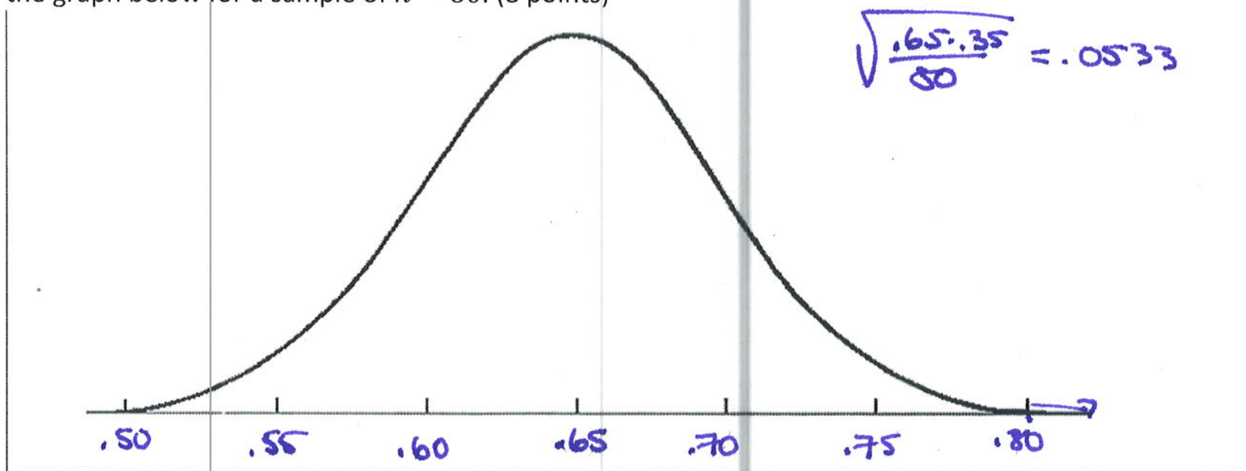
1. Draw the sampling distribution on the graph below with $\mu = 12.0, \sigma = 2.3$ when $n = 75$. (7 points)



What is the probability that the mean of a sample from this distribution will be less than 11.5?

normalcdf(-E99, 11.5, 12, .2656) = .02988

2. The population being sampled has a proportion of $p = 0.65$. Draw the sampling distribution on the graph below for a sample of $n = 80$. (8 points)



What is the probability that a sample of that size having a proportion greater than 80%?

normalcdf(.8, E99, .65, .0533) = .0024

3. The mean number of pets per student for a random sample of 47 students is 2.5 pets. If the standard deviation for the sample is 1.1 pets, find the confidence interval for the sample. (6 points)

T-Interval Stats

$\bar{x} = 2.5$

$S_x = 1.1$

$n = 47$

C-Level: .95

(2.177, 2.823)

4. If the sample above was reduced to 30 students, how does that affect the confidence interval? (3 points)

it will get wider

(2.0893, 2.9107)

5. Find the t-distribution values with the indicated properties. (3 points each).

a. $P(t > 1.36, df = 4)$

tcdf(1.36, E99, 4) = .1227

b. $P(t < -0.83, df = 11)$

tcdf(-E99, -.83, 11) = .2121

c. If we are using a t-distribution curve with $df = 21$, and the probability associated with it is 0.1173, what is the value of t that corresponds to this probability?

invT(.1173, 21) = −1.2238

6. If you want to calculate the mean height of women to 0.1 inches, knowing the standard deviation of the population is 3.1 inches, with 95% confidence, how large a sample size is needed? (5 points)

$$MofE = \frac{z_{\alpha/2}\sigma}{\sqrt{n}} \Rightarrow \sqrt{n} = \frac{z\sigma}{MofE}$$

$$n = \frac{z^2\sigma^2}{(MofE)^2} = 3691.77$$

$$\boxed{n = 3692}$$

7. Find the 80% confidence interval for defective chips for a sample of 200 computer chips if the defective rate of those chips is 4%. (5 points)

1 Prop Z Int

$x = .04 * 200 = 8$

$n = 200$

C-level: .8

(.02224, .05776)

8. If we wanted to calculate a 99% confidence interval with the margin of error being 1%, what is the sample size that would be needed in the above scenario? (5 points)

$$n = p(1-p)\frac{z^2}{E^2}$$

$n = 2537$

$$.04 * .96 \left(\frac{2.57}{.01}\right)^2 = 2536.28...$$

9. A teacher believes students learning math in elementary school do better when they use physical learning aids to understand concepts. She conducts a study to test her beliefs. (6 points)
   a. Describe a Type I error in the context of this study.

   The learning aids do not make a difference but you conclude that they do.

   b. Describe a Type II error in the context of this study?

   The learning aids make a difference but you conclude they do not.

10. A study is conducted and a P-value of 0.02 is found. What does this mean? Does your interpretation change if the α-level is 0.01? (5 points)

the P-value means there is a 2% chance that the results of the sample were obtained from the assumed population purely by chance. The significance level of 0.01 just means that this result is not unlikely unenough to reject our initial assumptions.

11. A national business magazine reports that the mean retirement age for women executives is 61.0. A woman's right's organization believes this value does not accurately reflect the current trend in retirement. To test this, the group polled a simple random sample of 95 recently retired women executives and found that they had a mean retirement age of 61.5. Assuming the population standard deviation is 2.5 years, is there sufficient evidence to support the organization's belief at the $\alpha = 0.05$ level? (8 points)

$H_0: \mu = 61$
$H_a: \mu \neq 61$

Z-test Stats
$\mu_0 = 61$
$\sigma = 2.5$
$\bar{x} = 61.5$
$n = 95$

$z = 1.949$
$p = .05125$
fail to reject Ho
this is not sufficient evidence to think retirement age has changed

a. Does your conclusion change if the sample size is 200? (3 points)

Change $n = 200$ $z = 2.828$ $p = .004677...$ yes, it matters here, reject Ho

b. What if the sample size is 30? (3 points)

Change $n = 30$ $z = 1.095$ $p = .2733$ fail to reject Ho for any significance level

c. What if $\alpha = 0.10$? or $\alpha = 0.01$? (3 points)

under initial conditions, we would reject Ho for $\alpha = .10$

$\alpha = .01$ would not alter conclusion

d. What does your conclusion mean in context? (3 points)

under initial conditions,

there is not sufficient evidence to conclude the age women executives retire has changed

12. A scientist believes that 12 out of every 100 workers suffers from extreme sensitivity to dust in the workplace. 197 workers are surveyed and it was found that 15% of them had extreme sensitivity to dust at work. Is this sufficient evidence to prove the hypothesis at the $\alpha = 0.10$ level? (8 points)

1PropZTest $p_0 = .12$

$x = .15 * 197 = 29.55$ - round to 30

$n = 197$

prop p_0

$z = 1.39$

$p = .0816$

$H_0: p = .12$

$H_a: p > .12$

reject H_0

this is sufficient evidence to think it is higher than .12

13. If we wished to analyze the scenario above with a confidence interval instead of a hypothesis test, find the corresponding confidence interval and the conclusion to the test (in context). (5 points)

1 Prop Z Int

$X = 30$

$n = 197$

C-Level = .90 (corresponds to $\alpha = .1$)

(.11018, .19439) .12 is inside the interval so we conclude this is not sufficient evidence to think they are different

14. Use the two-way table below to answer the questions that follow. (3 points each)

Gender	Preferred Program			
	Dance	Sports	Movies	Total
Women	16	6	8	30
Men	2	10	8	20
Total	18	16	16	50

a. What is the probability that a randomly selected person from this sample is a man?

$\frac{20}{50} = \frac{2}{5} = 40\%$

b. What is the probability that a randomly selected person from this sample prefers to watch dance?

$\frac{18}{50} = \frac{9}{25} = 36\%$

c. What is the probability that a randomly selected person from this sample is both a man and prefers to watch dance?

$$\frac{2}{50} = \frac{1}{25} = 4\%$$

d. What is the probability that a randomly selected person from this sample is either a man or prefers to watch dance?

$$\frac{20}{50} + \frac{18}{50} - \frac{2}{50} = \frac{36}{50} = \frac{18}{25} = 72\%$$

e. What is the probability that a randomly selected person from this sample is a man given that they watch dance?

$$\frac{2}{18} = \frac{1}{9} = 11\%$$

f. Are gender and preferred program independent? Why or why not? Show mathematical calculations to justify your answer.

$P(M) \cdot P(D)$ $\qquad$ $P(M \& D)$

$$\frac{2}{5} \cdot \frac{9}{25} = .144 \neq \frac{1}{25} = .04$$

dependent

15. A committee is to consist of 8 faculty selected from a staff of 28. 20 of the faculty are women, and 8 are men. (3 points each)

a. What is the probability that the entire committee will be men?

$$\frac{\binom{8}{8}}{\binom{28}{8}} = 3.217 \times 10^{-7}$$

b. What is the probability that the committee will consist of exactly 4 women and 4 men?

$$\frac{\binom{20}{4}\binom{8}{4}}{\binom{28}{8}} = .1091$$

c. What is the probability that the committee will have at least two women?

no women $\qquad$ one woman

$$\frac{\binom{20}{0}\binom{8}{8} + \binom{20}{1}\binom{8}{7}}{\binom{28}{8}} = 5.18 \times 10^{-5}$$

$$1 - 5.18 \times 10^{-5} = .999948$$

16. A charity holds a raffle and sells 2500 tickets for $5.00 each. They are awarding a first place prize of $2000, a second place prize of $1500, and four third place prizes of $500. Find the expected value of purchasing a ticket. Interpret the result in the context of the problem. [Hint: it may help to construct a table.] (8 points)

$1995 * \frac{1}{2500} + 1495 * \frac{1}{2500} + 495 * \frac{4}{2500} - 5 \cdot \frac{2494}{2500} = -2.8$

for each ticket purchased, one can expect to lose $2.80

17. A normal distribution has a mean of 75 and $\sigma = 12$. Use the Empirical Rule to answer the questions that follow. (3 points each)

a. What is the probability that a value will fall between 51 and 99?

-2 & 2

95%

b. What is the probability that a value will fall between 51 and 87?

81.5%

$\frac{68+95}{2} = 81.5$

c. What is the probability that a value will be greater than 63?

-1

100 − 16 = 84%

d. What is the percentile of the value 87?

1

84th percentile

e. What is the z-score of the value 21.8?

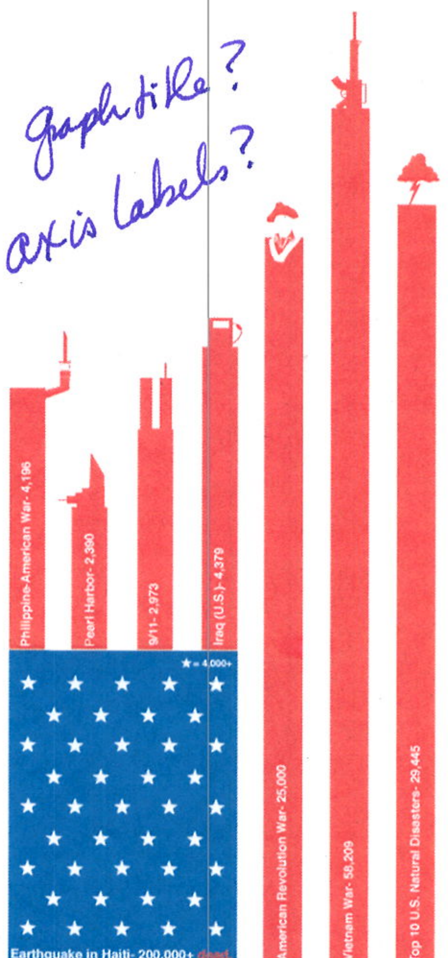
$\frac{21.8-75}{12} = -4.43$

18. If the ACT has a mean of 21.6 and a standard deviation of 5.2, above what score does the top 10% of scores fall? (5 points)

invNorm(.90, 21.6, 5.2) = 28.26

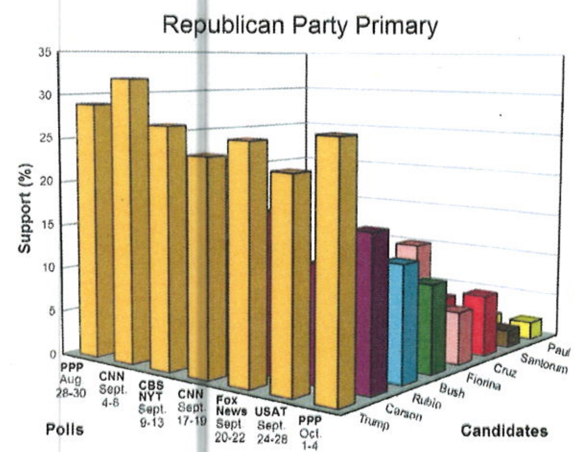
19. Some example bad graphs are shown below. Analyze the graphs and find at least 4 things that are wrong or misleading across all the graphs shown. (8 points)

graph title?
axis labels?



Zeros don't start at same place

what does the blue section mean?



3D can't be read

Size of bars do not correspond to #'s displayed

20. Use the data set below to answer the questions that follow.

20	0	72	64	69	60	0	14	27	71
7	5	67	65	27	58	10	14	19	5
50	26	71	9	26	22	30	56	6	16
30	29	37	22	7	24	17	36	15	21
27	7	15	9	8	12	6	98	23	8

a. Construct a frequency table using 7 classes. (5 points)

Class	Frequency	Relative Frequency (Percents, nearest tenth)
Total		

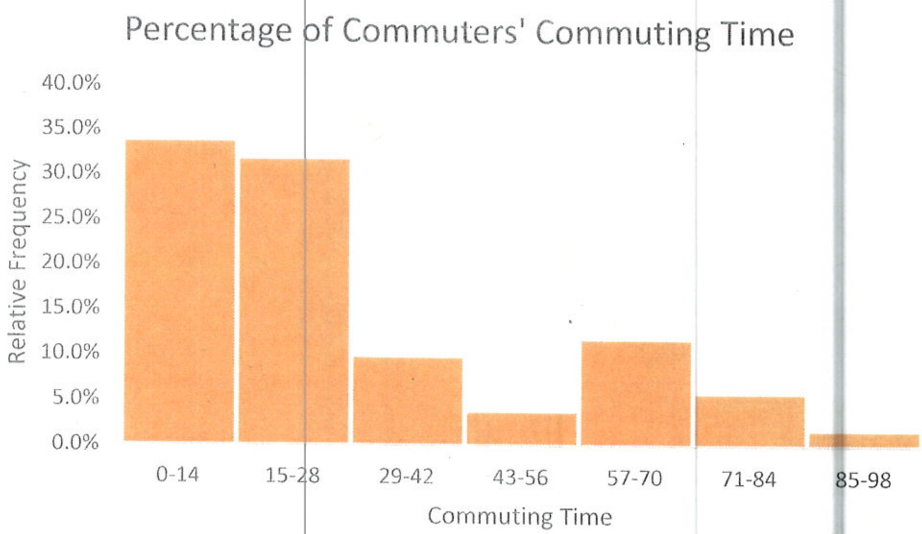
b. Draw a histogram of the data. (5 points)

See next page

c. Find the five-number summary of the data. (5 points)

Class	Frequency	Relative Frequency
0-14	17	34.0%
15-28	16	32.0%
29-42	5	10.0%
43-56	2	4.0%
57-70	6	12.0%
71-84	3	6.0%
85-98	1	2.0%

Min	0
Q1	9
Median	22
Q3	37
Max	98



d. Draw a boxplot with outliers. (8 points)

See next page

e. What percentile does a 12-minute commute represent? (3 points)

$\frac{15}{50}$ = 30^{th} percentile

f. What value represents the 90^{th} percentile? (3 points)

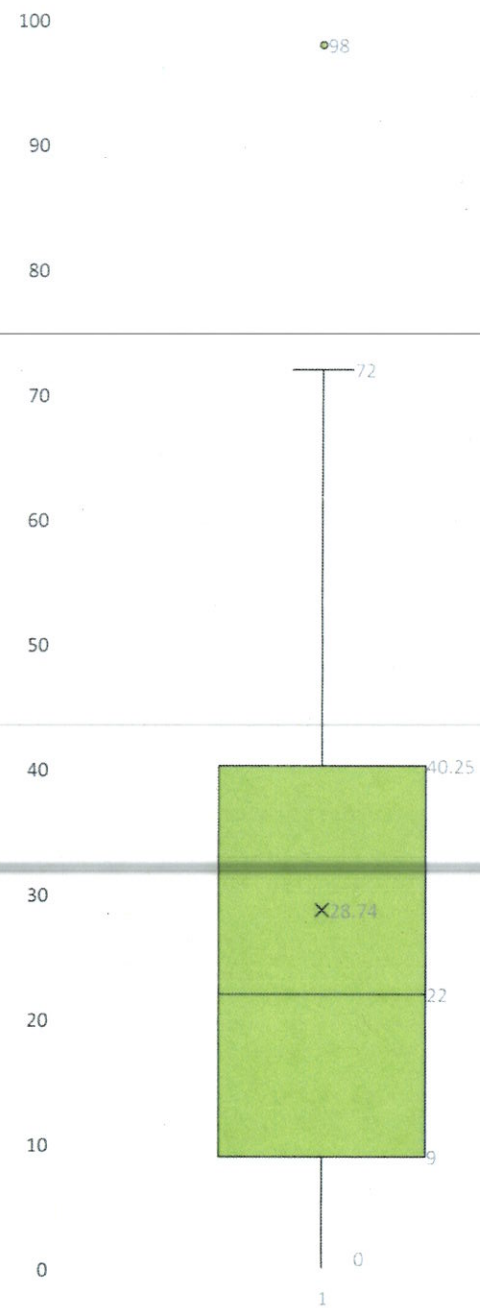
45^{th} value or 67 minute commute

g. What is the mean and standard deviation of the data? (6 points)

$\bar{X} = 28.7$

$S = 23.93$

h. Describe the shape of the distribution. Based on that, which measure of central tendency would be best to use for this data? (3 points)


100
90
80
70
60
50
40
30
20
10
0
98
72
40.25
28.74
22
9
0
1

21. Use the data in the table below to answer the questions that follow.

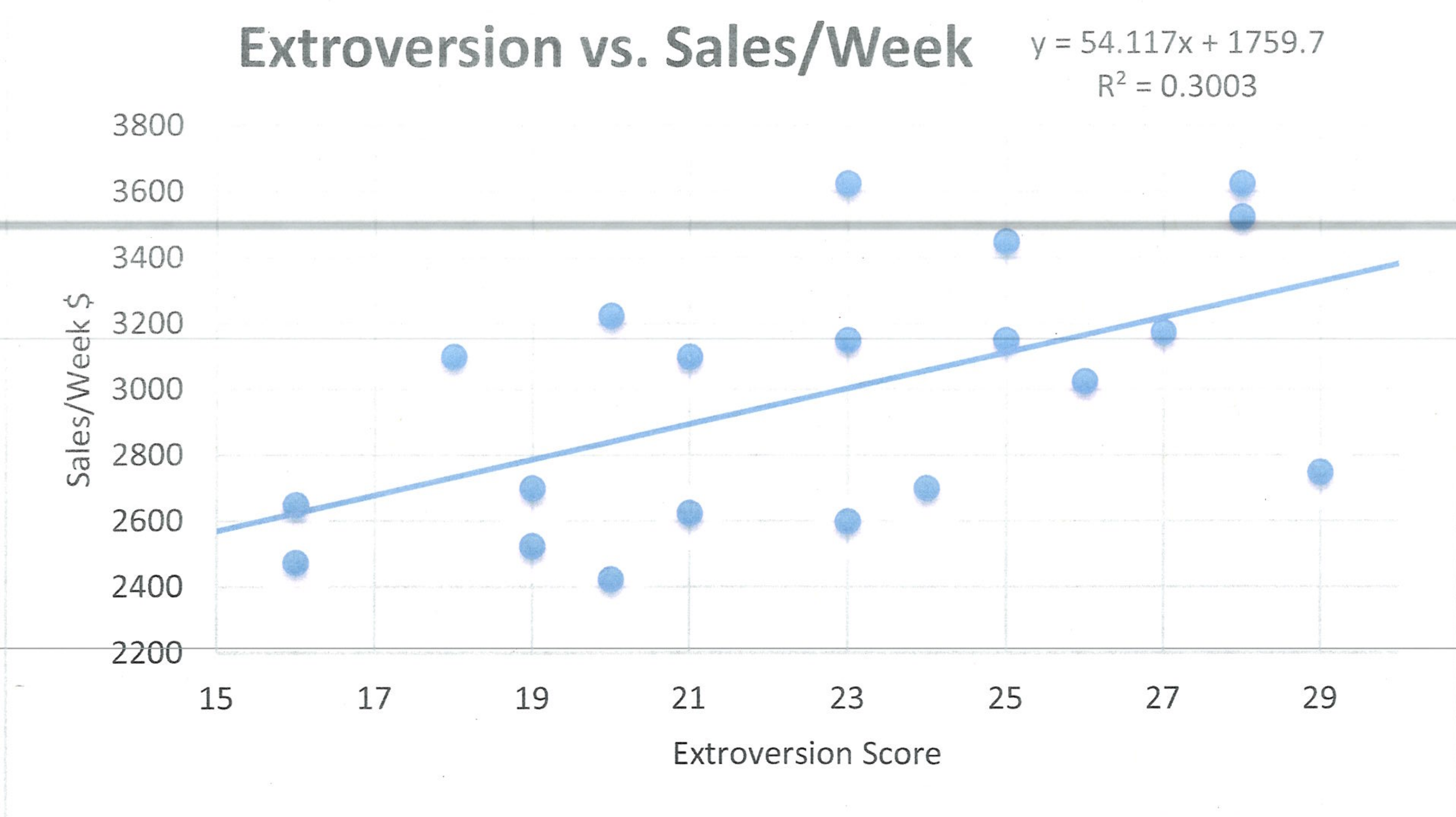
Sales Person	Intelligence	Extroversion	$ Sales/Week
1	89	21	$ 2,625
2	93	24	$ 2,700
3	91	21	$ 3,100
4	122	23	$ 3,150
5	115	27	$ 3,175
6	100	18	$ 3,100
7	98	19	$ 2,700
8	105	16	$ 2,475
9	112	23	$ 3,625
10	109	28	$ 3,525
11	130	20	$ 3,225
12	104	25	$ 3,450
13	104	20	$ 2,425
14	111	26	$ 3,025
15	97	28	$ 3,625
16	115	29	$ 2,750
17	113	25	$ 3,150
18	88	23	$ 2,600
19	108	19	$ 2,525
20	101	16	$ 2,650

a. Sketch a scatterplot of the extroversion vs. sales/week data. (5 points)

See next page

b. Find the linear regression equation that best fits the data. (5 points)

$y = 54.117x + 1759.7$


Extroversion vs. Sales/Week
y = 54.117x + 1759.7
R² = 0.3003
Sales/Week $
3800
3600
3400
3200
3000
2800
2600
2400
2200
15
17
19
21
23
25
27
29
Extroversion Score

c. What is the correlation? (3 points)

$r = .5480$

d. What is the coefficient of determination? (3 points)

$r^2 = .3003$

e. Interpret the slope in the context of the problem. (4 points)

For every point better on extroversion, on average one can expect to sell $54 more per week

f. What percent of the change in weekly sales can be explained by a change in extroversion score? (3 points)

30%

g. What does the y-intercept mean in the context of the problem? If it does not have meaning, explain why not. (4 points)

if you scored 0 on extroversion one could still expect to sell $1759.70 per week, on average (assuming 0 is a possible value on the test)

MAT 223, Formula Sheet for Final Exam

$$rel.frequency = \frac{count}{total}$$

$\mu = \bar{x} = \sum \frac{x_i}{n}$ $\sigma = \sqrt{\sum \frac{(x_i-\bar{x})^2}{N}}$ $s = \sqrt{\sum \frac{(x_i-\bar{x})^2}{n-1}}$

$IQR = Q_3 - Q_1$ outliers: $< Q_1 - 1.5IQR \ or > Q_3 + 1.5IQR$

Binomial distribution: $\binom{n}{x} p^x (1-p)^{n-x}$ $\mu = np$ $\sigma^2 = np(1-p)$

$$z = \frac{(x-\mu)}{\sigma} = \frac{x-\bar{x}}{s}$$

Standard errors: $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$

Counting: $\binom{n}{r} = \frac{n!}{r!(n-r)!}$ $P(n,r) = \frac{n!}{(n-r)!}$ spec. permt. $\frac{n!}{n_1!n_2!\ldots n_k!}$

Expected value: $\bar{x} = \mu = \sum X_i P(X_i)$

Sample sizes: $n > \hat{p}(1-\hat{p})\left(\frac{z_{\alpha/2}}{E}\right)^2$ $n > \left(\frac{z_{\alpha/2}\sigma}{E}\right)^2$

Confidence intervals: $\bar{x} \pm t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}$ $\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

Test statistics: $z \ or \ t = \frac{\bar{x}-\mu_0}{s/\sqrt{n}}$ $z = \frac{\hat{p}-p_0}{\sqrt{p_0(1-p_0)/n}}$

Common z-scores $z_{\alpha/2}$

Confidence level	$z_{\alpha/2}$
90%	1.645
95%	1.96
99%	2.575