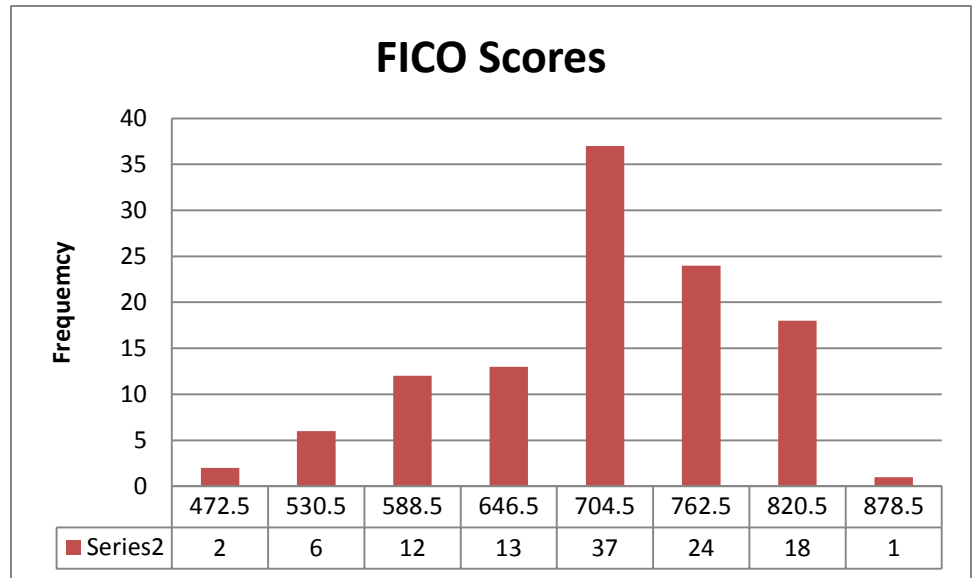


Instructions: Answer each question as thoroughly as possible. Be sure to use correct notation. To earn full credit, you must show work or the display from your TI-83/84 calculator screens. Round answers to three decimal places where appropriate (where the rounding rules discussed in class do not apply). Formulas are at the back. Note: The review is more comprehensive than the actual exam and will likely take longer to complete than the real test even working straight through.

1. In order to determine the average FICO score of Americans, a simple random sample of 100 scores were taken as reported by Experian. Use the data shown in the chart below to answer the following questions.



- a. Population
- b. Sampling Method
- c. Data Variable
- d. Is the data qualitative or quantitative?
- e. Is the data nominal, ordinal, interval or ratio?
- f. Class Width
- g. Sample Size
- h. Mean
- i. Median
- j. Modal Class

k. If a trade publication reported only the mean FICO score, would this be misleading? Why or why not?

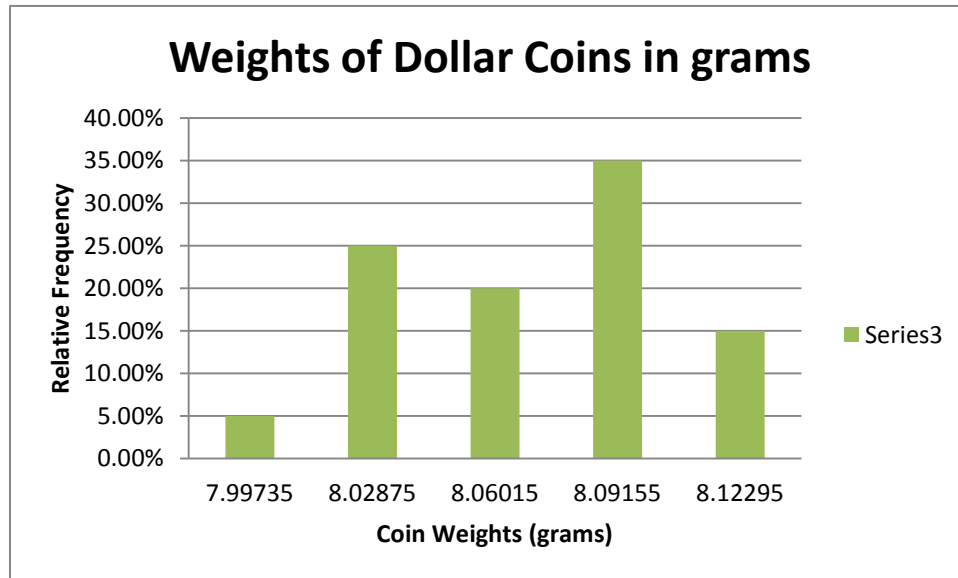
l. Standard Deviation

m. What does the standard deviation tell you about the data?

n. Shape?

o. Explain how the descriptive statistics confirm the shape.

2. In order to determine the average weight of dollar coins produced by the US Mint, random sample of 20 coins in circulation were taken from a single two banks in the Midwest. Use the data shown in the chart below to answer the following questions.



a. Population

b. Sampling Method

c. Data Variable

d. Is the data qualitative or quantitative?

- e. Is the data nominal, ordinal, interval or ratio?
- f. Class Width
- g. Sample Size
- h. Mean
- i. Median
- j. Modal Class
- k. If a trade publication reported only the mean weight of dollar coins, would this be misleading? Why or why not?

- l. Standard Deviation
- m. What does the standard deviation tell you about the data?

- n. Shape?
- o. Explain how the descriptive statistics confirm the shape.

- p. Use the information in the chart to build a pie chart for the data. Is this a better or worse method of visualizing the data? Explain your answer.

3. The data below was collected from 40 women measuring the sizes of their waists in centimeters. Use the information to answer the questions that follow.
- 67.2, 82.5, 66.7, 93.0, 82.6, 75.4, 73.6, 81.4, 99.4, 67.7, 100.7, 72.9, 85.0, 85.7, 126.0, 74.5, 74.5, 94.0, 92.8, 105.5, 75.5, 126.5, 70.0, 98.0, 104.7, 67.8, 99.3, 91.1, 74.5, 95.5, 79.5, 69.1, 105.5, 78.8, 85.7, 92.8, 72.7, 75.9, 68.6, 68.7

4. Use the following data on the length of R movies to construct a stem-and-leaf plot. Use the graph to answer the questions below.
- R: 110, 96, 120, 113, 108, 94, 90, 154, 127, 150, 136, 119.
- Shape of the graph.
 - Percentile rank of 130 minutes.
 - Write a statistically true statement about the length of R movies.
 - Give the five number summary of the data above.
 - Use that information to construct a box plot of the data.
 - Are there any outliers?
 - Suppose we also included data from PG-13 movies, PG movies and G movies. What type of graph would be best used to compare the results?

5. Use the data below on Brown and Yellow M&Ms plain candy weights (grams) to construct a stem-and-leaf plot comparing the two data sets. Use the graph to answer the following questions.

Yellow: 0.883, 0.769, 0.859, 0.784, 0.824, 0.858, 0.848, 0.851

Brown: 0.696, 0.876, 0.855, 0.806, 0.840, 0.868, 0.859, 0.982

a. What is the shape of each graph?

b. What is the percentile rank of 0.830 grams for each color of M&M candy?

c. Write a statistically true statement about the claim that both candies weigh the same.

d. Which color candy has a more consistent weight?

e. Use your TI-83/84 to construct a boxplot for each data set. Note any outliers on each.

6. A sample of 1504 American adults were recently surveyed by Pew Research (<http://www.people-press.org/2013/11/01/gop-deeply-divided-over-climate-change/>) to determine how party affiliation affected views of climate change. The data from the survey is briefly summarized below.

	Humans primary cause	Other causes primary/not occurring
Democrat/Lean Democratic	482	271
Republican/Lean Republican	172	579

- a. What is the probability that a randomly selected American agrees that humans are causing climate change?
- b. What is the probability that a randomly selected American is both Republican (or Republican leaning) and believes that humans are not causing climate change?
- c. What is the probability that a randomly selected American agrees that climate change is caused by humans, given that they are Republican (or Republican leaning)?
- d. What is the probability that 4 randomly selected Americans will agree that humans are causing climate change given that they are Democrats (or Democratic leaning)? (without replacement)
- e. Construct a hypothesis test to determine if party affiliation affects one's views on climate change. Use a 1% level of significance.
- i. H_0 :
- H_1 :
- ii. Test Statistic:

- iii. P-Value:
- iv. Reject or Fail to Reject the null hypothesis.
- v. State your conclusion in the context of the problem.

7. A sample of 1660 American adults were recently surveyed by Business Insider-Survey Monkey (<http://www.businessinsider.com/poll-the-major-differences-between-how-single-men-and-women-approach-sex-2013-9>) to determine how men and women behaved on the first date. The data from the survey is briefly summarized below.

	Slept with someone on the first date	Never slept with someone on the first date
Men	333	406
Women	258	663

- a. What is the probability that a randomly selected American says they have slept with someone on their first date?
- b. What is the probability that a randomly selected American is both a woman and has slept with someone on a first date?
- c. What is the probability that a randomly selected American has slept with someone on a first date given that they are a woman?
- d. What is the probability that 3 randomly selected Americans will admit to sleeping with someone on a first date given that each is a man? (without replacement)

- e. Construction a hypothesis test to determine if gender plays a role in determine how likely someone is to be willing to sleep with a partner on a first date. Use a 5% level of significance.
- i. H_0 :

 - H_1 :
 - ii. Test Statistic:
 - iii. P-Value:
 - iv. Reject or Fail to Reject the null hypothesis.

 - v. State your conclusion in the context of the problem.
8. About 10% of people are left-handed. Suppose that we choose 40 people randomly.
- a. Find the mean.

 - b. What is the standard deviation?

 - c. What is the probability that exactly 5 people in the sample will be left-handed?
9. On average, suppose that 42 people arrive at a drive-thru window during lunch hour, between 11:30 and 12:30 p.m. The arrival times follow a Poisson distribution.
- a. What is the mean number of cars between 11:45 and 12:00 p.m.?

 - b. Calculate the probability that at least 15 new cars will arrive between 11:45 and 12:00 p.m.

10. According to a Harris Interactive Poll (<http://www.harrisinteractive.com/NewsRoom/HarrisPolls/tabid/447/mid/1508/articleId/970/ctl/ReadCustom%20Default/Default.aspx>), about 1-in-5 people have tattoos. Suppose that we choose a sample of 88 people.
- Calculate the mean number of people in the sample with tattoos.
 - What is the standard deviation?
 - What is the probability that 30 people or more in the sample have tattoos?
11. The average lifetime of an LED light bulb is about 50,000 hours, with a standard deviation of around 20,000 hours.
- Determine the lifetime of a light bulb in the 10th percentile.
 - Is it unusual for an LED light bulb to last longer than 85,000 hours?
 - If 20 LED light bulbs are tested, what is the probability that the sample mean is greater than 60,000 hours?
12. The average IQ on all IQ tests is set at 100, but IQ tests differ by their standard deviations. Use your knowledge of a normal distribution to determine which of two students received a better score on their IQ test if the first student received a score of 135 on a test with a standard deviation of 15, and the second student received a score of 145 on a test with a standard deviation of 20.

13. The random number generator in your calculator generates numbers between 0 and 1 that are uniformly distributed. What is the probability that when it produces a random number that the value will be between 0.25 and 0.60?
14. Polling firm has conducted a survey and found that the Green Party candidate is receiving 52% of the vote and the Libertarian candidate is receiving 48% of the vote.
- If 1500 people were surveyed, calculate a 95% confidence interval for the Green Party candidate.
 - Based on this information, what is a statistically accurate claim about who is winning the election (if anyone) according to the poll?
 - If the polling firm wanted to relax their confidence to 90%, what would happen to the interval and explain why.
15. The average breaking length from 60 mph for domestic automobiles is given below. Use this data to answer the following questions.
- 131, 136, 129, 127, 146, 146, 155, 139, 133, 131, 131, 140, 137, 144, 133, 122, 133, 143, 144, 133, 134, 143, 129, 128, 140, 144, 146, 130, 139, 140, 135, 136
- Find the mean of the sample.
 - Find the standard deviation of the sample.

c. Use this data to construct a 95% confidence interval for the mean breaking distance of domestic automobiles.

d. If you were running a highway safety program, would a 95% confidence interval be enough to base safe following distance on? What would happen to the interval if you increased the confidence level to 99.9%?

16. A TV station is thinking about developing a program for television. In order to guarantee enough audience for the program, the station needs to show their backers they will have enough viewers to justify the cost of producing the pilot. The backers set a threshold of 20% of the surveyed audience being interested in seeing the program before they will give the money for the pilot. A survey of 100 randomly selected viewers of the station determines that 35% of the population says they would be interested in watching the show based on a plot synopsis. Conduct a hypothesis test to determine if this is enough to earn the backing of investors.

a. State the null and the alternative hypothesis.

b. Test statistic

c. Sketch the sampling distribution, clearly labeling the critical values and the rejection region.

d. Reject or Fail or to Reject H_0 .

e. State your conclusion in the context of the problem.

c. Reject or Fail to Reject H_0 .

d. State the conclusion in the context of the problem.

19. The data below gives budgets and gross (both in millions) for 20 PG-13 movies. Determine if there is a correlation between money spent on a movie and money earned.

116	70	75	52	120	125	150	30	225	70
103	66	121	116	101	213	290	100	322	19
80	40	70	300	113	68	72	160	29	132
117	48	228	373	380	118	114	120	120	234

a. State the null and alternative hypothesis to test for a significant correlation.

b. Sketch the rejection region for a 5% level of significance and label the critical value.

c. Use the test statistic to determine whether there is a correlation between money spent and money generated on PG-13 movies.

d. State the regression equation that most closely models this data.

e. Use your equation to predict the likely amount of money to be earned at the box office if \$100 million is spent on a new PG-13 movie.

Stat 1450 Formulas

Unit 2

$$\text{Class Width} = \frac{\text{max} - \text{min}}{\# \text{classes}}$$

$$\text{Class Midpoint} = \frac{\text{lower} + \text{upper}}{2}$$

Unit 3

Range = maximum – minimum

$$\text{Coefficient of Variation} = \frac{S}{\bar{x}} \cdot 100\%$$

Chebyshev's Theorem:

At least $1 - \frac{1}{k^2}$ of the data lie within k standard deviations of the mean.

$$\text{IQR} = Q_3 - Q_1$$

$$\text{z-score} = \frac{x - \text{mean}}{\text{stdev}} = \frac{x - \mu}{\sigma}$$

Unit 4

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

$P(A \text{ and } B) = P(A) * P(B)$ if A and B are independent events

$P(A \text{ and } B) = P(A) * P(B|A)$ if A and B are dependent events

$$P(\text{complement of } E) = 1 - P(E)$$

$$P(\text{at least } 1) = 1 - P(\text{none})$$

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)}$$

Unit 5

$$\text{Binomial: } \mu = np \quad \sigma = \sqrt{npq} \quad q = 1 - p$$

$$\text{Poisson: } \sigma = \sqrt{\mu}$$

Unit 6

$$z = \frac{x - \mu}{\sigma}$$

$$z = \frac{(\bar{x} - \mu)}{\left(\frac{\sigma}{\sqrt{n}}\right)}$$

Unit 7

Proportions:

$$CI = \hat{p} \pm z \cdot \sqrt{\frac{\hat{p}\hat{q}}{n}} \quad \text{or} \quad \hat{p} - E \leq p \leq \hat{p} + E$$

$$n = \frac{z^2 \hat{p}\hat{q}}{E^2} = \hat{p}\hat{q} \left(\frac{z}{E}\right)^2$$

Means:

$$CI = \bar{x} \pm z \cdot \frac{\sigma}{\sqrt{n}} \quad \text{or} \quad \bar{X} - E \leq \mu \leq \bar{X} + E$$

$$n = \left(\frac{z^* \sigma}{E}\right)^2$$

Unit 8

Hypothesis Testing:

$$\text{Proportions} \quad z = \frac{(\hat{p} - p)}{\sqrt{pq/n}}$$

$$\text{Means} \quad z = \frac{(\bar{x} - \mu)}{\left(\frac{\sigma}{\sqrt{n}}\right)}$$

Unit 9

Two proportions:

$$\bar{p} = \frac{r_1 + r_2}{n_1 + n_2}$$

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{\frac{\bar{p}\bar{q}}{n_1} + \frac{\bar{p}\bar{q}}{n_2}}}$$

$$E = z_c \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$$

$$(\hat{p}_1 - \hat{p}_2) - E < (p_1 - p_2) < (\hat{p}_1 - \hat{p}_2) + E$$

Two means:

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$t = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}}$$

Unit 10

$$E = \frac{\text{Total of row of interest} \cdot \text{total of column of interest}}{n}$$

Unit 11

Correlation $t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$

TABLE A-3 *t* Distribution: Critical *t* Values

Degrees of Freedom	Area in One Tail				
	0.005	0.01	0.025	0.05	0.10
	Area in Two Tails		0.05	0.10	0.20
1	63.657	31.821	12.706	6.314	3.078
2	9.925	6.965	4.303	2.920	1.886
3	5.841	4.541	3.182	2.353	1.638
4	4.604	3.747	2.776	2.132	1.533
5	4.032	3.365	2.571	2.015	1.476
6	3.707	3.143	2.447	1.943	1.440
7	3.499	2.998	2.365	1.895	1.415
8	3.355	2.896	2.306	1.860	1.397
9	3.250	2.821	2.262	1.833	1.383
10	3.169	2.764	2.228	1.812	1.372
11	3.106	2.718	2.201	1.796	1.363
12	3.055	2.681	2.179	1.782	1.356
13	3.012	2.650	2.160	1.771	1.350
14	2.977	2.624	2.145	1.761	1.345
15	2.947	2.602	2.131	1.753	1.341
16	2.921	2.583	2.120	1.746	1.337
17	2.898	2.567	2.110	1.740	1.333
18	2.878	2.552	2.101	1.734	1.330
19	2.861	2.539	2.093	1.729	1.328
20	2.845	2.528	2.086	1.725	1.325
21	2.831	2.518	2.080	1.721	1.323
22	2.819	2.508	2.074	1.717	1.321
23	2.807	2.500	2.069	1.714	1.319
24	2.797	2.492	2.064	1.711	1.318
25	2.787	2.485	2.060	1.708	1.316
26	2.779	2.479	2.056	1.706	1.315
27	2.771	2.473	2.052	1.703	1.314
28	2.763	2.467	2.048	1.701	1.313
29	2.756	2.462	2.045	1.699	1.311
30	2.750	2.457	2.042	1.697	1.310
31	2.744	2.453	2.040	1.696	1.309
32	2.738	2.449	2.037	1.694	1.309
33	2.733	2.445	2.035	1.692	1.308
34	2.728	2.441	2.032	1.691	1.307
35	2.724	2.438	2.030	1.690	1.306
36	2.719	2.434	2.028	1.688	1.306
37	2.715	2.431	2.026	1.687	1.305
38	2.712	2.429	2.024	1.686	1.304
39	2.708	2.426	2.023	1.685	1.304
40	2.704	2.423	2.021	1.684	1.303
45	2.690	2.412	2.014	1.679	1.301
50	2.678	2.403	2.009	1.676	1.299
60	2.660	2.390	2.000	1.671	1.296
70	2.648	2.381	1.994	1.667	1.294
80	2.639	2.374	1.990	1.664	1.292
90	2.632	2.368	1.987	1.662	1.291
100	2.626	2.364	1.984	1.660	1.290
200	2.601	2.345	1.972	1.653	1.286
300	2.592	2.339	1.968	1.650	1.284
400	2.588	2.336	1.966	1.649	1.284
500	2.586	2.334	1.965	1.648	1.283
1000	2.581	2.330	1.962	1.646	1.282
2000	2.578	2.328	1.961	1.646	1.282
Large	2.576	2.326	1.960	1.645	1.282

