Name: Key **Math 127 Exam 1A Spring 2015**

Oath: ***"I will not discuss the exam contents with anyone until it is returned to me by my instructor"***

Sign Name: Key

Points are in parentheses for each question. Half point for clearly writing your name.

Students may use:

1. **A handheld calculator that is not your cell phone and not the computer calculator.**
2. **One sheet of hand-written or typed notes on 8.5" by 11" paper. No photos, scans, or copies of any posted materials (like answer keys or completed worksheets). No pink sheets.**
3. **The StatCrunch webpage and datasets.**

This exam must be completed in one sitting, but it is untimed.

Good luck.

ZZZ - Retired -

1. Load up the "**Calendar Year 2015 Large Survey**" dataset. Show fraction and percentage rounded to two decimal places on all the categorical variable questions. Round all summary statistics to two decimals if necessary.

1a. (2) What percentage of the ***"Males"*** use "***Pinterest***" in any capacity?

4/41 = 9.76% 10/102 = 9.8%

1b. (2) What percentage of all respondents "***Text and Drive***" all the time?

12/124 = 8.06% 30/333 = 9.01%

1c. (2) What percentage of all respondents are motivated by "***Knowledge***" or "***Expression***"?

56/125 = 44.8% 158/334 = 47.31%

1d. (2) What percentage of those who would "***Google***" a date are "***Female***"?

47/59 = 79.66% 129/156 = 82.69%

1e. (2) **How many** students have a "***Commute***" of at least 40 minutes? ~~10~~ 34 Students

1f. (2) Show calculation. Determine the lower fence for "***Age***".

$LF = Q_1 - 1.5(IQR)$

$= 18 - 1.5(6) = 9$

$= 19 - 1.5(5) = 11.5$

1g. (2) How many students are official high outliers for "***Number of Tattoos***"? ~~17~~ 56 Students

1h. (2) What is the median age of the students who smoke? ~~25.5~~ 24

1i. (2) What is the best measure of center for "***Talk Radio / Podcasts***"? Give its name and value.

Median = 1 Hour

1j. (2) What is the best measure of spread for "***TV Hours***"? Give its name and value.

IQR = ~~2 Hours~~ 8 Hours

1k. (2) How many children does the oldest student have? ~~4~~ 2

1l. (2) How many students failed to answer "***Is Contraception Wrong***"? ~~3~~ 5

1m. (2) Convert Professor Kupe's "***Ideal Children***" to a *z*-score. He's in row 1. Show calculation.

$z = \frac{y - \bar{y}}{s} = \frac{0 - 2.51}{2.169} = -1.158$

$\frac{0 - 2.45}{1.69} = -1.4497$

2. Graph Reading (circle best choice)

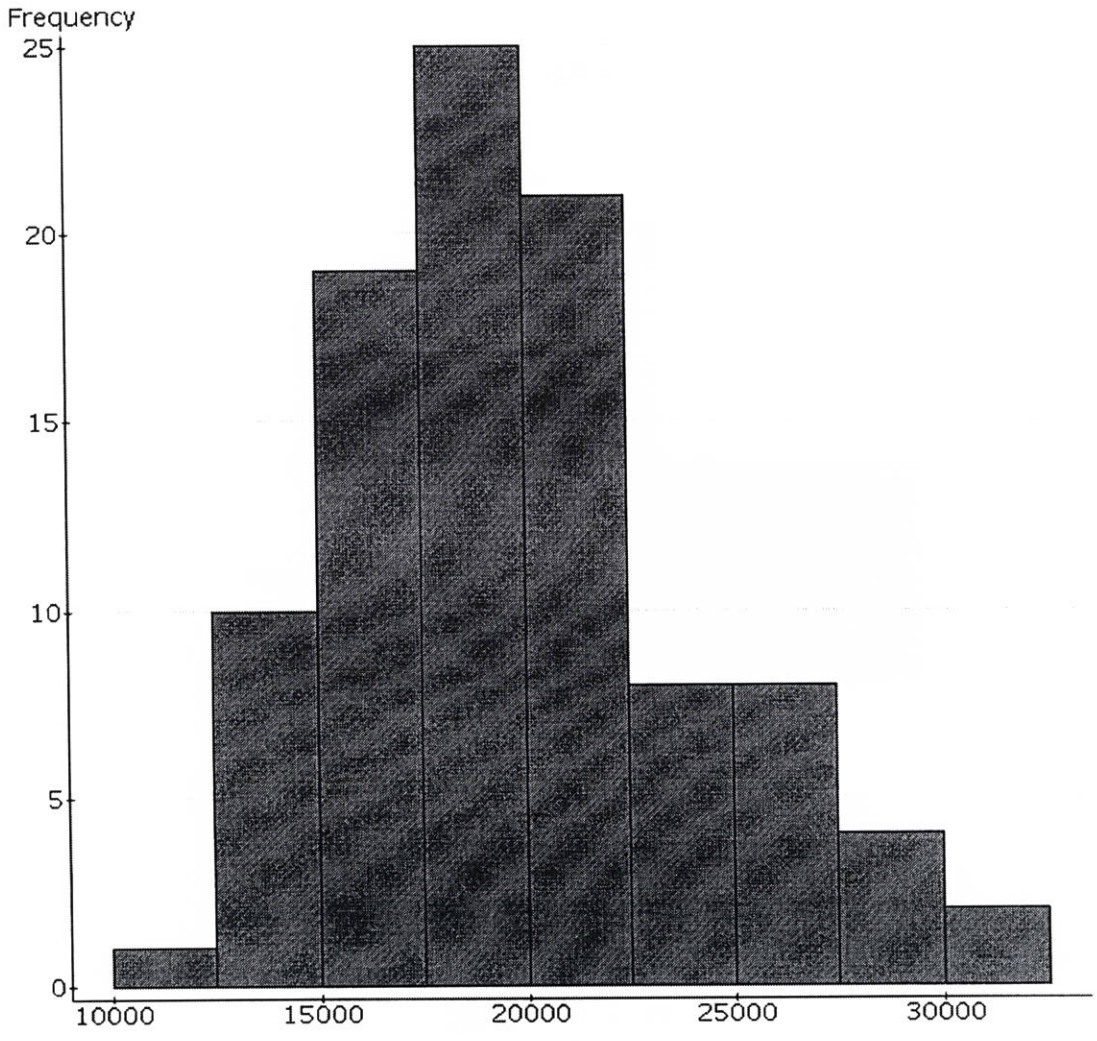
2a. (2) How many total observations were used to create this graph? 25 98 32,500

2b. (2) The standard deviation is _____. 42.39 4157.43 36,881.06

2c. (2) The mean is greater than the median. True False

2d. (2) This graph is a _____. Bar Plot Histogram

2e. (2) The range is _____. 10,000 19833 32,500



3. Graph Reading (compute value or circle best choice). For this boxplot, all values used to create it were unique. No trickery, no pile ups of data at any value.

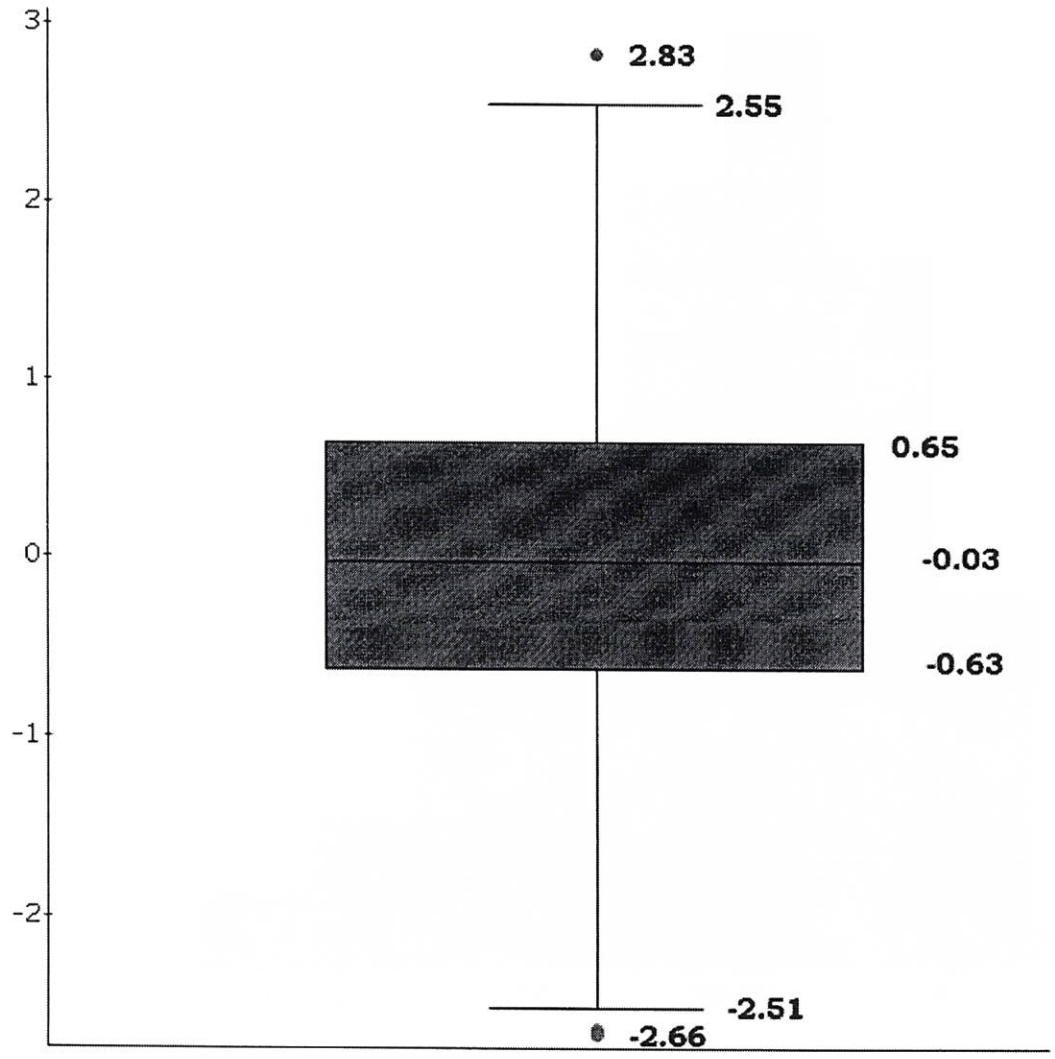
3a. (2) Compute the interquartile range: $IQR = 0.65 - (-0.63) = 1.28$

3b. (2) Compute the range: $Range = 2.83 - (-2.66) = 5.49$

3c. (2) The 79th percentile _____. (Must exceed 0.65) Must exceed 2.55

3d. (2) The 49th percentile _____. Could be positive (Must be negative)

3e. (2) The upper fence is 2.55. True (False)



4. To research the effect of wearing sunglasses when playing poker, a player ran an experiment over the course of one year. Each time he went to the casino, he flipped a coin and if it came up heads, he wore sunglasses that day. He kept track of his daily net winnings for the year and compared the mean profits with and without sunglasses. He played four times each week, twice during the week and both on Saturday and Sunday.

Because the players vary wildly between "***during the week***" and "***on the weekend***", he noted this down in his records. If during the week, he wore sunglasses the first day, then he didn't the second day. Same thing for the weekend.

4a. (2) Experimental units: Pokers Sessions

4b. (2) Experimental factor: Sunglasses - Y/N

4c. (2) Blocking factor: DAY OF WEEK

4d. (2) Response variable: NET WINNINGS

5. Identify the official sampling methodology for each scenario (simple random, stratified, cluster, systematic, convenience, census, multistage).

5a. (2) Online dating habits among male college students were analyzed with a sample of size 200 from the male population of size 7000. The sample was chosen using every 35th name on the list of all 7000 students. **Sampling method:** Systematic

5b. (2) A company is being audited and the auditors will pull from the pool of 5500 accounts receivable records. The total pool can be broken down into groupings by dollar amount: Under \$1000, from \$1000 to \$5000, and over \$5000. The auditors will randomly select 100 records (using a computer program) from each grouping. **Sampling method:** Stratified

5c. (2) The county supervisors are curious if the rural residents of Cecil County would support the installation of a sewer system. The supervisors divide up the rural parts of the county into 195 zones, and then randomly select (using a computer program) six zones. Once the zones are selected, every household in those zones will be visited and surveyed. **Sampling method:** Cluster

5d. (2) The second day of Math 127, we visited the library and randomly selected a book. First, an aisle was randomly chosen, then a bookcase was randomly chosen, and finally a book was randomly chosen. **Sampling method:** Multistage

Sum is $5(7.47) = 37.35$

6a. (2) Invent a dataset with 5 unique values with a mean of exactly 7.47. No repeats. If your mean is off by just a little bit, points off.

1	2	3	4	27.35

6b. (2) Invent a dataset with 8 values where the IQR is larger than the third quartile.

-1	-1	-1	-1	1	1	1	1

Give the values: IQR = 2 Q_3 = 1

7. (5) Fire up the "**Titanic**" dataset. Argue if survival was independent of or dependent on age. Support with the proper conditional and / or marginal percentages.

Dependent!
57/109 = 52.3% of children survived
654/2092 = 31.3% of adults survived.
Big Diff. Based on age.

8. (5) "**IMDB Movie Ratings**" dataset. Describe the distribution of the "Year" for the movies in this sample. Use the values of the best summary statistics in your write up.

Unimodal Skewed left

Best: Median = 2000 IQR = 27
Not: Mean = 1991.2 St. Dev. = 22.05

Officially 88 low outliers
(Anything under 1940.5)

9. Open up the "**Marvel vs. DC at the Box Office**" dataset. The last four variables are the box-office revenues given in millions of US dollars.

9a. (2) Describe the "***Who***": Each Film

9b. (3.5) There are seven "***Whats***" in this dataset. Label as Q, C, or I.

C Company I Film Q Release (Cok)

Q Adjusted Q Worldwide Q Domestic Q Foreign

9d. (4) The "***Domestic***" revenue for Guardians of the Galaxy was $222.281 at the time of dataset creation. Is this unusual or rare in terms of *z*-scores? Show calculation and make a concluding remark.

0.439 *z*-score Unusual? Rare? Comment: No! Z-score is within ±2 St. Dev.

Show work:

$$Z = \frac{222.281 - 161.675}{138.152} = 0.439$$

9e. (5) Suppose the "***Adjusted***" column in the dataset is adjusted for inflation to bring all the entries up to 2015 values. All values in the column will be scaled to 105% of their current values. In other words, ***The Avengers*** would now be $1654.6404.

What would happen to the following summary numbers if the change was made all the movies?

Circle correct choice for each row:

Mean:	Increase by 5%	Remain unchanged
IQR:	Increase by 5%	Remain unchanged
Q_3:	Increase by 5%	Remain unchanged
Range:	Increase by 5%	Remain unchanged
Standard Deviation:	Increase by 5%	Remain unchanged

10. (2) Give the common language definition for standard deviation. No formulas.

"Like" the average / typical distance to the mean.

11. (2) Give the common language definition for a *z*-score. No formulas.

Number of standard deviations a data value lies away from the mean.

12. (2) True or False. An IQR can be negative.

13. (3) Fill in the six blanks in the table.

Contingency table results:
Rows: How Religious
Columns: Astrological Sign

	No	Yes	Total
Extremely religious	9	0	9
Not religious	29	13	42
Somewhat religious	40	10	50
Very religious	1	0	1
Total	79	23	102