Print Name: Key Math 127 – Exam 1 – Fall 2017

Version Mario

Oath: ***"I will not discuss the exam contents with anyone on planet Earth until the answer key is posted to Blackboard."***

Sign Name: _____

The penalty for cheating on this Exam is a grade of 0% for Math 127 Exam 1.

Student Instructions

1. **This test is graded out of 100 points and counts for 20% of your Math 127 grade. Points are in parentheses for each question.**

2. **You can use a calculator, but you cannot use your phone. You can use the calculator on the computers if you wish.**

3. **You will need to use www.statcrunch.com. This is the only permitted webpage.**

4. **You are permitted to use one 8.5" by 11" sheet of notes, front and back. You will submit it with your test.**

   You may not use the pink sheet or copies of the pink sheet.

   You must produce (handwritten or typed up) your own sheet of notes.

   You may not use copies or scans of any instructor-created Math 127 content or answer keys.

5. **Show work or points will be deducted. If you only report an answer and it is wrong, you will receive no credit.**

1. (8) Identify the sampling method for each scenario. Pick from census, cluster, convenience, simple random, stratified, or systematic.

1a. Simple Random The StatCrunch University activity from class.

1b. Systematic Go to Professor Kupe's apartment building in Towson. It has 18 floors, and each floor has 20 apartments. Number the apartments 1, 2, 3, …, 359, 360. Take apartment #4 and then every 10th apartment until you reach the end of the list. Ask the residents in each apartment how they feel now that we've decided to remove the screens from the exterior windows.

1c. Cluster Break into Cecil College's registration office in the middle of the night. Obtain the list of all students who attend our school. Take every student whose last name begins with the letter "K" for the sample. By the way, the letter "K" was chosen randomly by a computer. Make a list of all the majors taken by the students in the sample.

1d. Convenience Go to the last Orioles home game in Baltimore. Buy your ticket, find your seat, and get comfortable. Survey the people around you in your section to ask them how they feel about gender neutral bathroom facilities.

2. (4) On a recent visit to Towson, Professor Kupe's Mom played Bejewed Blitz on her iPad each day. Here is the table for number of games each day:

Thursday	Friday	Saturday	Sunday	Monday
37	89	79	95	26

Show the calculations below to arrive at the lower and upper fences.

$Q_1 = 37$, $Q_3 = 89$, $IQR = 52$

$LF = Q_1 - 1.5(IQR) = 37 - 1.5(52) = -41$

$UF = Q_3 + 1.5(IQR) = 89 + 1.5(52) = 167$

zzz Retired

3. (4) In the **"Calendar Year 2017 Library Data"** dataset on StatCrunch, convert the ***"Weight"*** of ***"Drug Trafficking in the Americas"*** to a z-score. Show your calculation. **Hint:** Don't forget Edit → Find.

$$z = \frac{y - \bar{y}}{s} = \frac{1008 - 666.18}{395.74}$$

$$z = 0.864$$

ZZZ Retired

4. (4) In the "**Calendar Year 2017 Large Survey**" dataset, person #180 did not report a "***Salary***", but we learn her z-score was –0.73158. Show to calculation to solve backwards for her "***Salary***".

To Rounding

$$-0.73158 = \frac{y - 15981.69}{21845.47}$$

$$y = 15981.69 - 0.73158(21845.47) = \$0$$

5. (2) In the "**Calendar Year 2017 Large Survey**", how many students are outliers (think fences) for the variable "***College Credits***"? Answer only is OK.

Low Outliers: 0 High Outliers: 2

6. (2) Write a simple definition for median: It is the middle value for a numerically sorted list of values.

7. (2) Write a simple definition for standard deviation: SD is "like" the average distance to the mean.

8. (2) Write a simple explanation for how to take a simple random sample: Number your Population 1, 2, 3, ..., N. Use a Compute to generate n random numbers.

9a. (4) Fire up the "**Skyscrapers in the U.S.**" dataset. Identify each variable as **Q** = Quantitative, **C** = Categorical, or **I** = Identifier.

Rank: ? Building: I City: C Height: Q

Floors: Q Completed: Q Material: C Primary Use: C

9b. (2) How many buildings are missing their "***Material***"? 732

9c. (5) Describe the distribution of the variable "***Height***". Use only the best summary statistics for the situation. Bullet points are OK. Fully address outliers.

- Shape - Unimodal, Skewed Right
- Center = Median = 132.6 m
- Spread = IQR = 44.4 m
- Outliers: 93 high outliers - Any Building Taller than 225.9 or 226.3 m

10. (8) Legendary rock band Radiohead will run a designed experiment at their next 16 concerts. The goal is to maximize the amount of rock experienced at their shows.

Determined by a coin flip, half of the concerts will get an "***Encore***", and half will not.

Also determined by a coin flip, half the concerts will have "***Video Montage A***" playing behind the stage, and half the concerts will have "***Video Montage B***" playing.

You can presume the experiment will be balanced so that the four treatments are equally divided among the sixteen concerts.

Radiohead will determine the amount of rock at each concert by measuring the maximum crowd noise with a Larson Davis Model 831 Sound Level Meter.

For this experiment, identify the following:

10a. Experimental units: 16 Concerts

10b. Factor 1: Encore

10c. Factor 1's levels: Yes or No

10d. Factor 2: Video Montage

10e. Factor 2's levels: A or B

10f. Response variable: Crowd Noise MAX

10e. Is this experiment single blind, double blind, or neither? Neither

10f. Now open the "**Radiohead**" dataset on StatCrunch. Which of the four treatments had the highest mean "***Crowd Noise***"? Report the mean as well.

Best Treatment: Yes Encore, Video A Mean "***Crowd Noise***": 111.11

11. Use the "**Hip Surgery Outcomes**" dataset to answer the following problems.

Example answer for percentage problems: 27/55 = 0.4909 = 49.09%

11a. (2) Percentage of surgeries performed by "***Jackson***":

37/209 = 0.1770 = 17.70%

11b. (2) Percentage of surgeries that had "***Complications***":

28/209 = 0.1340 = 13.40%

11c. (2) Percentage of "***Jackson's***" surgeries that had "***Complications***":

9/37 = 0.2432 = 24.32%

11d. (2) Percentage of surgeries on patients in their 40's:

12/209 = 0.0574 = 5.74%

11e. (2) Percentage of surgeries on "***Hispanic***" "***Females***":

48/209 = 0.2297 = 22.97%

11f. (2) Percentage of surgeries with a "***Recovery***" length under two weeks

33/209 = 0.1579 = 15.79%

11g. (2) **Number of** surgeries performed by "***Bright***" on "***African American***" "***Females***": 9

11h. (2) Give the 90th percentile for "***Recovery (in days)***": (50) 26.4 Explain its meaning with a sentence or two:

20% of patients recover within (50) 26.4 days, while 80% have recoveries of at least (50) days.

26.4

11i. (2) Which "***Ethnicity***" tends to have the shortest "***Recovery***" times? Hispanic

11j. (2) "***Brand***" vs. "***Gender***". Independent or Dependent? No explanation needed on this one.

Answer: I

ZZZ Retired

12. (33) Use the "**Calendar Year 2017 Large Survey**" on this one.

12a. 6 — **How many** students reported exactly $5000 in "***Credit Card Debt***"?

12b. 14 — Use Data → Compute Expression to create a new variable that sums "***Credit Card Debt***" with "***Student Loan Debt***". How many students have at least $30,000 in total debt?

12c. 22.92 — Mean "***Age***" of our "***Republicans***"

12d. 19, 48 — "***Ages***" of our two students with the highest "***Salaries***".

12e. 402 — **How many** students drink "***Alcohol***"?

12f. 82 — **How many** students "***Smoke***"?

12g. 404 — **How many** students drink "***Alcohol***" or "***Smoke***"?

12h. 2 — The best measure of center for "***Number of Siblings***" is the Median. Put its value on the first line.

12i. 10 — The best measure of spread for "***TV Time***" is the IQR. Put its value on the first line.

12j. For the variable "***Sleep Time***", create a range of values that would not be unusual. Base your interval on *z*-score ideas. Show the calculation: $\bar{y} = 6.704$, $s = 1.445$

$\bar{y} \pm 2(s) \rightarrow 6.704 \pm 2(1.445)$

6.704 ± 2.89

$(3.814, 9.594)$

12k. Does "***Facebook***" usage depend on "***Gender***"? Run the appropriate pie charts and write your response where you clearly state independent / dependent and you support your choice with the appropriate conditional percentages.

Dependent! 37.97% of Females use Facebook All The Time, compared to only 19.75% of Males.

More than 10% difference, so dependent.

12j. (2) Extra Credit – How many students in the sample are "***Females***" with no "***Children***" who don't "***Smoke***" and are "***Atheist /Agnostic***", and at least 29 but at most 42 years old?

Answer: 1