Name: _____*Key*_____   **Math 127 – Test 1 – Fall 2016 Version α**

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.

Testing Center Staff Instructions

1.   One sheet of handwritten or typed notes is OK.

        Students may **not** use the "pink sheet" or any copied or scanned answer keys or Math 127 department documents.

2.   Collect the sheet of notes and staple it to the test when submitted.

3.   Testing Center issued TI calculator is OK.

4.   www.statcrunch.com is required. All other webpages are prohibited.

5.   Test must be completed in one sitting, but it is untimed. Very short bathroom breaks are permitted.

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.

Retired —

1. Load up the "**Calendar Year 2016 Food Bank**" dataset. Round all summary statistics to two decimals if necessary. Give all percentages rounded to two decimals, e.g. 13.58%.

1a. __Median__ = __100__ *Calories*   Best measure of center for "*Calories*" plus its value. **(3)**

1b. __IQR__ = __7__ *Fat Grams*   Best measure of spread for "*Fat Grams*" plus its value. **(3)**

1c. Product with the most "*Sugar*" is __Coca Cola Cherry__ with __70__ grams of sugar. **(3)**

1d. __117__ __**How many**__ products have under 10 grams of "*Sugar*"? **(3)**   (123 is -1)

1e. __2.23__ Convert the "*Calories*" for "*Wolf Brand Chili*" in row 80 to a z-score. Show calculation. **(3)**

$$z = \frac{400 - 142.48}{115.36} = 2.23$$

1f. $\frac{17}{145} = 0.1172 = 11.72\%$ __What **percentage**__ of products have at least 100 "*Fat Calories*"? **(3)**

1g. Describe the distribution of "*Protein Grams*" using the best summary statistics. Bullet points OK with me. **(5)**

Shape: Skewed Right ~~the best~~
Bimodal
Center: Median = 5 g
Spread: IQR = 5 g
Outliers: 6 high outliers (anything past 14, 14.5)
16 g, 17 g, 18 g, 19 g, 24 g

1h. __0.945__ The "*Chicken Noodle Soup*" in row 127 has a missing value for its "*Fat Grams*". Suppose the z-score is z = –0.50. Solve for its "*Fat Grams*" and show your calculation below. **(3)**

$$z = \frac{y - \bar{y}}{s}$$

$$-0.50 = \frac{y - 4.05}{6.21}$$

$$y = 4.05 - 0.5(6.21)$$

$$\boxed{y = 0.945}$$

2

2.   Short answer.

2a. (7) Suppose on the second day of class, we did this instead at the library. We start at the very first book on the very first shelve on the very first aisle and take that book. Then we just count and take every 70th book until we get to the back of the library. Those books will be our sample.

In StatCrunch, we record "*Book Title*", "*Book Author*", "*Bar Code Number*", and "*Number of Times Checked Out*".

Sampling method: _Systematic (3)_

Type of Variables (Q, C, or I):   "*Book Title*" _I_ (1)   "*Book Author*" _C_ (1)
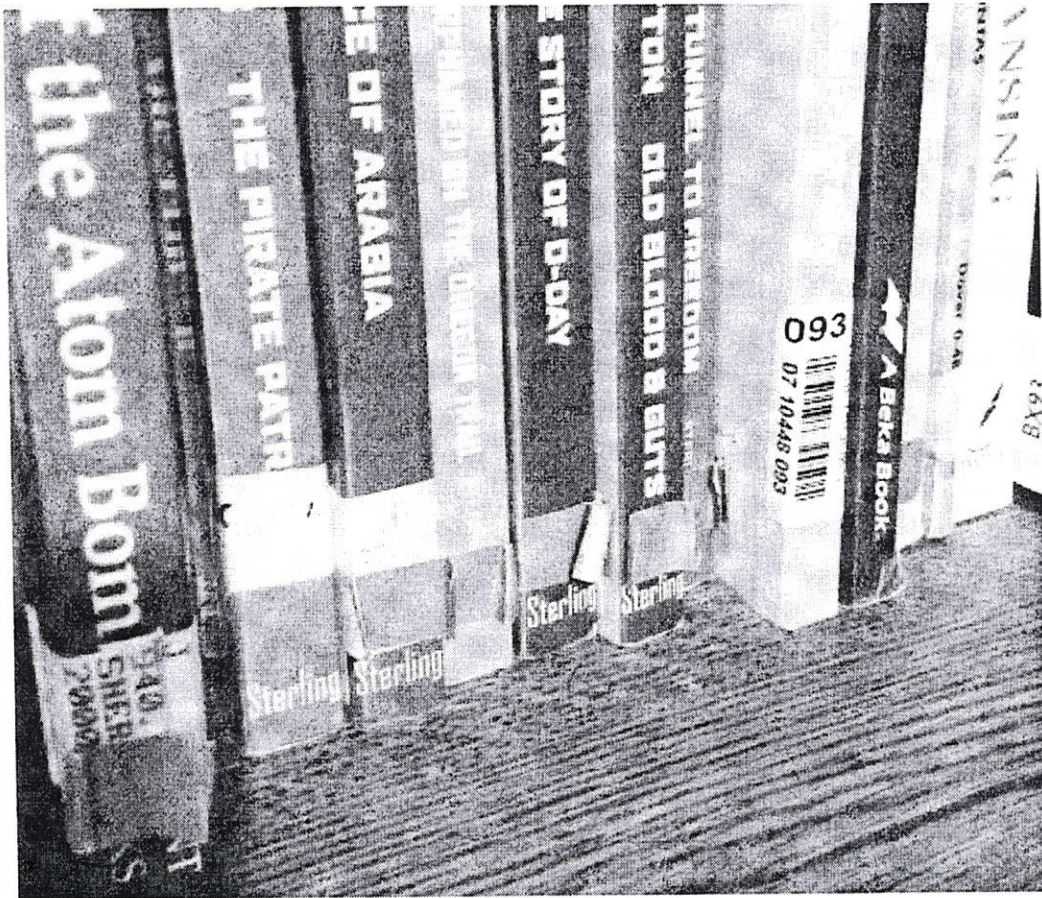
"*Bar Code Number*" _I_ (1) "*Number of Times Checked Out*" _Q_ (1)

2b. (3) In Towson, MD, Professor Kupe lives in a building with 18 floors. The apartment manager wants to survey the residents about the cleanliness of the refuse room on each floor. The manager will take a simple random sample of two residents from each and every floor in the building.

Sampling method: _Stratified (3)_

3

2c. (3) Invent a dataset that has a standard deviation of exactly 1. Doesn't matter how many data values you need.

1, 2, 3 works

2d. (3) Explain in common language the meaning of "***standard deviation***".

"like" the average / typical distance to the mean

2e. (3) On this Math 127 Exam, suppose your "***Exam Score***" exceeds the upper fence. You did

Pretty Crappy   (Pretty Awesome)   Can't Tell With the Given Information

2f. (3) Invent a dataset with nine values such that the IQR = 5 and the mean is negative.

e.g.

-1000	0	0	0	0	5	5	5	5

Retired

2g. (3) In the "**Calendar Year 2016 Personality Type**" dataset, give **The Who**.

Each Student

Retired

2h. (3) Was our data collection for the "**Calendar Year 2016 Food Bank**" a designed experiment or an observational study?

Observational Study

Retired

2i. (4) In the "**Calendar Year 2016 Large Survey**" dataset, look at the conditional distributions of "***Marital Status***" based on "***Military***" experience.

Does it appear the two variables are independent or dependent? Explain, and you must support with conditional percentages for full credit.

Dependent for sure.

No Military, 8.82% ~~have been~~ are "married"
Yes Military, 40.00% are "married"
Big Difference, so dependent.

4

Retired

2j. (3) Use our "**Calendar Year 2016 Large Survey**" to calculate the fences for "**Online Time**". Show your calculations.

$$\text{Lower Fence: } Q_1 - 1.5\,IQR = 4 - 1.5(11)$$
$$= -12.5$$

$$\text{Upper Fence: } Q_3 + 1.5\,IQR = 15 + 1.5(11)$$
$$= 31.5$$

Retired

2k. (2) Use the "**Calendar Year 2016 Large Survey**" dataset to determine how many official outliers we have for "**Number of Tattoos**".

Low-valued outliers: ___0___   High-valued outliers: ___63___

$$(9 \text{ is } -1)$$

2l. (3) Show the algebra or explain with a diagram why the lower fence and upper fence are 4 IQRs apart.

$$UF - LF = (Q_3 + 1.5\,IQR) - (Q_1 - 1.5\,IQR)$$
$$= (Q_3 - Q_1) + 1.5\,IQR + 1.5\,IQR$$
$$= IQR + 1.5\,IQR + 1.5\,IQR = 4\,IQR$$

2m. (3) A Cecil county resident has a "**Net Worth**" of \$43,500 and that translates to a z-score of 1.39. If the standard deviation for the whole county is \$50,000, solve for the mean "**Net Worth**" and show the calculation:

$$z = \frac{y - \bar{y}}{s} \qquad 1.39 = \frac{43,500 - \bar{y}}{50,000}$$

$$\bar{y} = 43,500 - 1.39(50,000)$$
$$\bar{y} = -26,000$$

3. (10) Use the "**NYPD January 2012**" dataset. In each row is the result of a police interaction. Report your answers as 54 / 67 = 0.8060 = 80.60%.

3a. Percentage of all interactions that were "**Male**".

$$64208/69073 = 0.9296 = 92.96\%$$

3b. Percentage of all interactions in which an "**Arrest**" was made.

$$3816/69073 = 0.0552 = 5.52\%$$

3c. Percentage of the "**Males**" that were "**Frisked**".

$$37144/64208 = 0.5785 = 57.85\%$$

3d. Percentage of those "**Frisked**" in which the officer "**Found Something**".

$$1684/38628 = 0.0436 = 4.36\%$$

1838

3e. Number of interactions in which "**Race**" was not reported by the officer: ___1838___

5

4. A designed experiment was run to grow a mold in a Petri dish.

Two different sizes of dish were used, "*Large*" and "*Small*".
Three different sweetener solutions were used, "*Sugar in the Raw*", "*Sweet 'n Low*", and "*Splenda*".
Two different breeding grounds were used, "*Bread*" and "*Sponge*".

At the end of two weeks, "*Amount of Mold*" was measured for each of the 144 Petri dishes.
The StatCrunch dataset is set up like this:

Row	Amount of Mold	Petri Dish	Sweetener	Breeding Grounds	Run
1	1.793852	Large	Splenda	Sponge	1
2	2.3163805	Large	Sugar in the Raw	Bread	2
3	2.7529563	Large	Sweet 'n Low	Sponge	3
4	2.5161157	Small	Splenda	Bread	4
5	2.3219717	Small	Sugar in the Raw	Sponge	5
6	2.655789	Small	Sweet 'n Low	Bread	6

4a. (3) What was the response variable for this experiment? ___Amount of Mold___

4b. (3) If every combination of the factors were used by the experimenter, how many different treatments were there in this experiment? Show calculation.

$$2 \times 3 \times 2 = 12 \text{ treatments}$$

5. (3) Use the "**Hospital Payments**" dataset in StatCrunch. Using the idea of z-scores, give a range of values that would be considered not unusual for the variable "*Average Covered Charges*". Show calculations.

$$\bar{y} \pm 2S$$
$$37593 \pm 2(43256)$$

Inside here, not unusual. +124,105
$(-48919, \cancel{0})$

OR $(\$0, \$124,105)$

6. Use the "**Bachelor's Degree Institutions**" dataset on this one.

6a. (3) ___8206.82___ Give the mean for "*Undergrads*" for the "*Public Schools*".

6b. (3) ___8825.12___ Give the mean for "*Undergrads*" for the "*Public Schools*" that offer "*Campus Housing*".

6c. (3) Interpret with a sentence the first quartile for "*Student-to-Faculty Ratio*". 25% of schools have a S-t-F ratio of 12 or below. 75% of schools have a S-t-F ratio of 12 or above

6d. (3) Create a new variable that takes the ratio of "*Undergrads*" to "*Students*". (Data button). If you did this right, the Minimum = 0.025799215. Give the mean and explain what this number means.

Mean: ___0.84___. Interpretation: On average, 84% of the student populations are undergrads. (16% are grad stud)

6