Print Name: _Key_

Math 127 – Exam 1 – Spring 2017

Version Patriots

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

Sign Name: _Key_

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

Dedication

This exam is dedicated to the college students who say "please" and "thank you" when asking for something. To those who respect the classroom, the syllabus, others sharing the space, and the instructor. It is dedicated to the students who I see working before class, after class, asking questions in office hours and emailing / texting honest-to-goodness questions at all hours of the night. It is dedicated to those who have jobs, who have kids, who take care of their parents, who are dealing with illnesses or who are battling problems outside of their control and still make time for statistics. It is especially dedicated to those who call me by my name and not the word "hey". To the students who are mature enough to sit in a college classroom without being rude by talking during class or distracting others by playing on the computers, Professor Kupe salutes you and wishes you the very best of luck on this examination.

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 <u>only</u> 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 <u>not</u> use the pink sheet or copies of the pink sheet.

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

    You may <u>not</u> 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. (4) Here is the product listing for the *"Blanc Noir Texture Sweatshirt"* on www.zappos.com.

Blanc Noir Texture Sweatshirt

SKU 8699027

BLANC NOIR.

$99.00

Ships Free!

White/Black ▼

Choose Women's Siz ▼

Don't see your size?

Add To Cart

Item Information

- Redefine comfort with this stylish look!
- Textured fabrication with solid contrasts for a bold look.
- Mesh ventilation panels throughout help heat escape.
- Crew neckline.
- Long sleeves with thumb holes at cuffs to keep hands warm.
- Banded hem.
- 97% polyester, 3% spandex.
  Contrast: 90% polyester, 10% spandex.
  Contrast 2: 96% cotton, 4% spandex.
- Machine wash, tumble dry.
- Imported.
- Measurements:
  - Length: 28 in
- Product measurements were taken using size SM. Please note tha

View Zappos.com Glossary of Terms

Find something wrong in this description? Help us fix it! Report An Er

1a. What kind of variable is the *"SKU"*? ___I___ Q, C, or I (for all of these questions)

1b. What kind of variable is the *"Product Name"*? ___I___

1c. What kind of variable is the *"Price"*? ___Q___

1d. What kind of variable is *"Length"*? ___Q___

1e. What kind of variable is *"Polyester %"* ___Q___

1f. What kind of variable is *"Imported"*? ___C___

1g. What kind of variable is *"Color"*? ___C___

1h. What kind of variable is *"Women's Size"* (was listed as XS, S, M, etc…) ? ___C___

2. (5) True or False. Write clearly, especially if you are a "T" or "F" person!

2a. _____F_____ A negative standard deviation indicates that most of the data values fall below the mean.

2b. _____F_____ Simple random samples are relatively easy to do in the real world.

2c. _____F_____ The types of conclusions you can make using a convenience sample and a designed experiment are essentially the same.

2d. _____F_____ A *z*-score of 0 would be really unusual.

2e. _____F_____ If the IQR is 0, then the upper and lower fences must also be 0.

ZZZ Retired –

3. (3) On StatCrunch, we have the "**Calendar Year 2017 Library Data**" dataset. Be extra crystal clear on this one! The three answers are different from each other.

3a. Give *The Who*: Each book (collected by students)

3b. Give the **population**: All books in our library

3c. Give the **sample**: The n=~~202~~ books we collected
451

4. On StatCrunch, use the "**Fatal Encounters Updated…**" dataset for all of problem 4. Probably best to report percentage answers like this:

333 / 998 = 0.3337 = 33.37%

4a. (3) What percentage of the reported cases had a "*Location of death (state)*" from "*MD*" = Maryland?

208/14881 = 0.0140 = 1.40%

4b. (3) What percentage of the "*Location of death (state)*" = "*MD*" cases had "*Subject's Gender*" = "*Male*"?

192/208 = 0.9231 = 92.31%

4c. (3) What percentage of the "*Males*" died by "*Gunshot*"?

10676/13593 = 0.7854 = 78.54%

4d. (3) There is one case from our school's zip code of "*21901*".

Give the "*Subject's Name*": Charles Hall

4e. (3) Most common "*Age*" in the dataset is: 22

4f. (3) What was "*Nickolos Cyrus's*" "*Cause of Death*"? Taser

4g. (2) How many "*Age's*" are official outliers?   Low Outliers: _0_   High Outliers: _210_

4

5. (5) Use the **"Calendar Year 2017 Large Survey"**. Discuss the independence or dependence of the variables **"Astrological Sign"** and **"Tattoo"**.

You must include the appropriate conditional proportions / percentages to support your side.

You must say the correct words that go with the proportions / percentages you have chosen.

One or two good sentences will cover it.

33.67%

~~People with tattoos:~~ ~~40.4%~~ believe in Astrology

~~People w/out tattoos:~~ 26.37% ~~20.7%~~ believe the Same.

We have a 10%+ difference, so the variables are ~~dependent~~.

ZZZ Retired

6. (5) Use the "**Calendar Year 2017 Large Survey**". Describe the distribution of the variable "***Number of Tattoos***".

You must hit all the points we've hit in class when we've done this exact type of problem.

Bullet points OK, use only the best summary statistics for your description.

Shape: Unimodal + Skewed Right

Center: Median = 0 tattoos (mean = 1.32)

Spread: IQR = 2 tattoos (SD = 2.68)

Outliers: ~~1~~ 30 high outliers, Anyone with 6+ tats.

7. (3) Invent a dataset with 11 values that has a $Q_3 = 10$ and a mean = 100. If you don't get it within 5 minutes, skip ahead and come back later.

e.g. (infinite possibilities)

| 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 1000 |

8. (3) The typical human adult emits about 40 radioactive atoms per second (on average) with a standard deviation of about 10 atoms. Professor Kupe has a Geiger counter and at 11:00 a.m. on Saturday, February 11, 2017, his count was 4,827 per second. Convert that value to a z-score. Show calculation. Is Kupe's count unusually large?

z-score: 478.7   Unusually Large? Yes!

$$z = \frac{4827 - 40}{10} = 478.7$$

"ZZZ Retired" *(handwritten, red)*

9. Use the "**Calendar Year 2017 Large Survey**" dataset for all questions on this page.

9a. (4) Mean ideal "*Marriage Age*" for "*Females*" = ~~26.59~~ 26.29 and then for "*Males*" = ~~26.66~~ 26.90.

9b. (4) ☐ Circle the best measure of spread for "*Number of Siblings*": Standard Deviation or (IQR.) Value = ~~2.106~~ ~~3~~ 2

9c. (4) How many students are "*Divorced*" "*Democrats*"? ~~3~~ 8

9d. (4) Professor Sheppard is in Row 4. Convert her "*Commute*" to a *z*-score, show the calculation:

$$z = \frac{13 - 19.74}{15.05} = -0.448$$

$$z = \frac{13 - 19.46}{13.49} = -0.479$$ *(red)*

9e. (4) Using the idea of *z*-scores, give a range of "*Sleep*" hours that would **not** be considered unusual. Show calculation:

$$\bar{y} \pm 2S$$

$$6.71 \pm 2(1.45) \qquad 6.704 \pm 2(1.445)$$

$$(3.81 \text{ hours}, \ 9.61 \text{ hours})$$

3.814   9.594 *(red)*

9f. (4) Calculate by hand the fences for "*Student Loan Debt*":

$$LF = 0 - 1.5(3000) = -4500$$

$$UF = 3000 + 1.5(3000) = 7500$$

9g. (4) Interpret, with a sentence or two, the 99th percentile for "*Credit Card Debt*": 99% of our respondents have at most ~~$10,000~~ $11000 in CC debt. 1% have ~~$10,000~~ or more in CC debt. $11,000

9h. (4) Professor Kupe did not report his "*Salary*", but because he is an incredibly wealthy man, his *z*-score is 3.774. Show the calculation and solve for his "*Salary*".

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

~~$y = \$74,000$~~

$$3.774 = \frac{y - \overset{15,981.69}{\overbrace{14,642.16}}}{\underset{21845.47}{15728.12}}$$

$98,426.49 *(red)*

6

10. (6) For the problem, the population of interest is the entire student body of Towson University. At Towson, there are 22,285 students, of which 18,807 are undergraduate students and 3,478 are graduate students. We will aim to take a sample of 1,000 students.

Identify the sampling method for each scenario: census, simple random, systematic, cluster, stratified, or convenience. Some answers are used more than once and some answers are not used at all.

10a. _Stratified_ Take a random sample of 840 undergrads and a random sample of 160 graduate students to arrive at a total sample of 1000 students. Towson U. is 84% undergraduate, 16% graduate, by the way.

10b. _Stratified_ Towson students represent 101 countries of origin. Divide the population into 101 groups, one for each country of origin, and take a random sample from each group. The sample sizes will be dictated by how large each group is (in other words, if 70% of Towson enrollment is Americans, then we will randomly select 700 Americans). Total sample size is still 1,000.

10c. _Convenience_ (-½ cluster) The largest class at Towson might be Psychology 101, with total enrollment of 1,388 when you add up all the sections across one school year. To take our sample, we decide to visit all the PSYC 101 courses over the course of year 2017 and survey those students to get our number of 1,000.

10d. _Systematic_ Using an email list that is sorted by "*Student ID*" number, we program a computer to email every 20th student.

10e. _Simple Random_ Using an email list that is sorted by "*Student ID*" number, we ask StatCrunch to randomly select 1,000 of those numbers and then we email those students.

10f. _Convenience_ (-½ census) Using that email list, we just email everyone and take what we can get, because we know that students rarely check their emails. We expect about a 5% reply rate.

11. (8) Short Answer.

11a. A z-score of –2.73 means our data value was _2.73 SD below the mean. Or unusually small compared to the mean._

11b. Give two statistics that cannot be negative: _IQR, SD, Range_

11c. For data that is symmetric, unimodal, and clean, the best measure of the center is the _mean_ and the best measure of spread is the _std. dev._.

11d. Your "*Zip Code*" is certainly a number, but it is not Quantitative. Why not? _Doing math on Zip Code doesn't make sense. Zip Code is Categorical._

7

**12.** Here is a designed experiment. The treatments are assigned randomly to 40 white clothes.

First, we apply a "*Stain*", which can be "*Ketchup*" or "*Mustard*".
Then, the cloths are either washed in "*Temperature*" = "*Hot*" or "*Cold*"
Also, the washing machine is set to either "*Cycle*" = "*Heavy*" or "*Delicates*".

At the end, a Colorimeter will be used to score the amount of stain that remains on the cloth. The goal is to see which of the above combinations leads to the best Colorimeter "*Score*" (best being least stain remaining).

12a. (4) Draw a diagram (we did in class and on the videos, some call this a tree diagram, not sure if I did, but it looked like one when we were done) that starts with the 40 clothes up top and works its way down to the different treatments. You can assume equal / balanced group sizes for simplicity. Work top to bottom or left to right, either way is OK. You must include the number of clothes in each treatment as you work your way down.
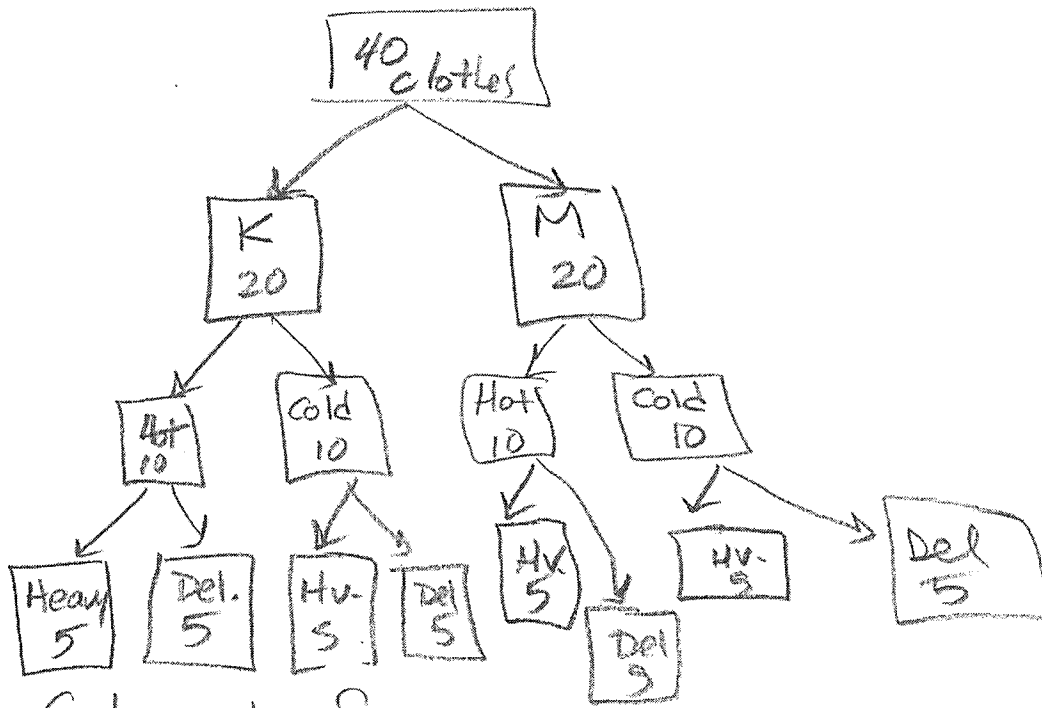
12b. (1) <u>Colorimeter Score</u> Identify the response variable.

12c. (1) ___8___ How many different treatments do we have for this experiment?

Extra Credit: One point each, based on the Hans Rosling video I told you about.

E1. ___T___ True or False. Everything is generally getting better.

E2. ___T___ True or False. The income distribution worldwide now has one hump.

E3. ___F___ True or False. Hans used an incredibly tiny pointer stick to make a point.

E4. ___F___ True or False. The humans are beating the chimps on the quizzes about worldwide trends.