**Name: You Math 127 – Test 1A – Summer 2015**

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

**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**](http://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.** You can use a calculator, but you cannot use your phone. You can use the calculator on the computers if you wish.

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

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

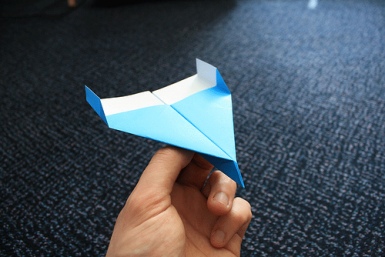
You may **not** use the pink sheet or copies of the pink sheet.

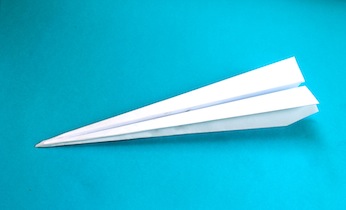
You must produce (handwritten or typed up) your own sheet of notes.

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

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

**1.** The “**Paper Planes**” dataset contains the results of a randomized experiment. The variable “***Distance***” is measured in feet from the throw point. The student running this experiment wanted to see if “***Design***” mattered and if adding “***Flaps***” mattered for how far a plane could be expected to fly.

 This is a glider with flaps. It could also be without flaps.



This is a dart without flaps. Flaps could be added.

**Identify the following:**

**1a. (2)** First factor with levels:

**1b. (2)** Second factor with levels:

**1c. (3)** Response variable: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**1d. (2)** How many different treatments are there? Circle: 1 2 3 4 64

**1e. (4)** Give the treatment means:

Dart with flaps:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Dart without flaps:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Glider with flaps:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Glider without flaps:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

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

**2a. (3)** At StatCrunch U, the 46,000 students were numbered, went home and put on their numbered T-shirts, came back to campus, and then using a random number generator, *n* = 32 students were selected.

**Sampling method:**

**2b. (3)** The registrar at Cecil College keeps grade records for every single student that has ever attended this institution. **Sampling method:**

**2c. (3)** Professor Kupe has his financial records for every purchase he has made in the last 8 years on [www.mint.com](http://www.mint.com). He takes every 13th purchase starting at the beginning of the list to determine his median purchase amount. **Sampling method:**

**2d. (3)** Suppose on the 2nd day of class, instead we did this at the library. With 20 aisles and 8 bookcases per aisle, there are 160 total bookcases in our library. We draw a random number between 1 and 160, go find that bookcase, and take every book found on that bookcase as our sample. Then we record the same variables as we did for our “**Calendar Year 2016 Library Data**” dataset.

**Sampling method:**

**3. (3)** Circle the only correct expression.







**4. (2)** Which of the following statistics **could** take on negative values? Circle all that are correct.

median mean standard deviation IQR minimum

range *Q*1 *z*-score *Q*3 maximum

**5. (3)** In words, what is a *z*-score? Give the common language definition. No formulas accepted for credit.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**6a. (10)** Describe the distribution of “***Recovery (in days)”*** in the “**Hip Surgery Outcomes**” dataset. Use the values of the best summary statistics in your write up. Bullet points OK.

Giving every single statistic that StatCrunch can produce will result in a penalty.

Finally, let’s agree that for all intents and purposes, “***Recovery (in days)***” is pretty darn symmetric.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**6b. (4)** Looks like there was a tie for the longest recovery time. Convert that value to a *z*-score. Show work.

**6c. (4)** A patient’s recovery time was lost. If his *z*-score was 0, what was his recovery time? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**7.** Load up the “**ZZZ Retired -** **Calendar Year 2015 Large Survey**” dataset. Show fraction, then decimal, then percentage rounded to two decimal places on all the categorical variable questions.

**7a. (3)** What percentage of the ***“Females”*** drink “***Very Often***”?

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**7b. (3)** What percentage of all respondents are “***Extremely Religious***”?

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**7c. (3)** What percentage of our “***Married***” students think “***Marriage is Obsolete***”?

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**7d. (3)** How often does our tallest student drink “***Alcohol***”? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**7e. (3)** Show calculation. Professor Kupe is in row 1. Using *z*-scores, is the “***Number of States***” he has visited unusually high?

**7f. (3)** Interpret with a sentence, the 90th percentile for “***Number of Tattoos***”:

**7g. (3)** What is the mean “***Height***” of the females? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**7h. (3)** What is the best measure of center for “***TV Time***”? Give its name and value.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**7i. (3)** What is the best measure of spread for “***TV Time***”? Give its name and value.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**7j. (3)** Are a majority of our respondents in the 20 to 29 age bracket? Circle: Yes No

**8a. (3)** Invent a dataset with 11 values with a mean of **exactly** 6.5 and a *Q*3 of **exactly** 8.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |  |

**8b. (3)** Invent a dataset with a standard deviation of exactly 1, as computed by StatCrunch:

**9. (4)** Fire up the “**NYPD January 2012**” dataset. Included are the police interrogation records from all 69,073 police interactions that month. Argue if getting frisked is independent of or dependent on gender. Support with the proper conditional and / or marginal percentages.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**10.** In the “**Honshu Japan Earthquake**” dataset, we have the 446 earthquakes that occurred in the region immediately before and after the magnitude 8.9 quake on March 11, 2011.

**10a.(3)** Give the “***Who***”: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**10b.(2)** The variable “***NST***” is the number of earthquake stations that reported earthquake activity for each particular earthquake. Is the variable categorical or quantitative?

Categorical Quantitative

**10c.(4)** Calculate the upper and lower fences for “***Magnitude***”. Show work: