Flu-X, the contagious virus that causes student to uncontrollaby study statistics rages on. It is up to you to find a cure.

Question 01

Pizza, pizza, pizza. There is something about how pizza interacts with Flu-X that makes you think—is this a cure? From Homework one we enrolled 5 infected students into a "pizza eating" group and another 5 infected students into a "non eating pizza" group. After collecting data on these 10 students you observe the following data

Pizza	Hours studying statistics within 24 hours of pizza
1	2.1
1	3.1
1	1
1	0.3
1	0.4
0	3.5
0	3.6
0	7.8
0	5.4
0	9.0

The variable Pizza equals 1 for observations (students) who were assigned to the pizza eating group and 0 for students assigned to the not-eating pizza group.

(A)

Compute the mean number of hours studying statistics in the pizza eating and the non-pizza eating groups.

(B)

Compute the variance of the number of hours studying statistics in the pizza eating and the non-pizza eating groups.

(C)

Compute the standard deviation of the number of hours studying statistics in the pizza eating and the non-pizza eating groups.

(D)

What does your preliminary evidence suggest about the effect of pizza (the exposure) on the number of hours studying statistics (the outcome)?

Question 02

(A)

During experimentation, one of the infected students says "Ah ha! I've found a new formula!!". They give you the following formula

$$S = \sum_{i=-3}^{3} \frac{x_i}{(i+4)}$$

Given the following table of data,

X	Value
x_{-3}	3
x_{-2}	2
x_{-1}	1
x_0	0
x_1	4
x_2	5
x_3	2

what is S?

You decide to expand your study and enroll 200 students. After rigorous experimentation you find the mean hours of studying statistics from 100 students in the pizza group equals 2.24 hours. You also find 100 students in the non-pizza group averaged 4.45 hours. It looks like pizza is working! You plan to compile all your data to present but DISASTER. In the pizza group, you lost student number 72's number of hours of study. Second DISASTER. In the non-pizza group you lost student number 34's number of hours of study.

You do, however, know that in the pizza group the total number of hours for the remaining 99 students who you do have data for was 189 hours and in the non-pizza group the total number of hours for the remaining 99 students was 410.

B

How many hours did student 72 in the pizza group study?

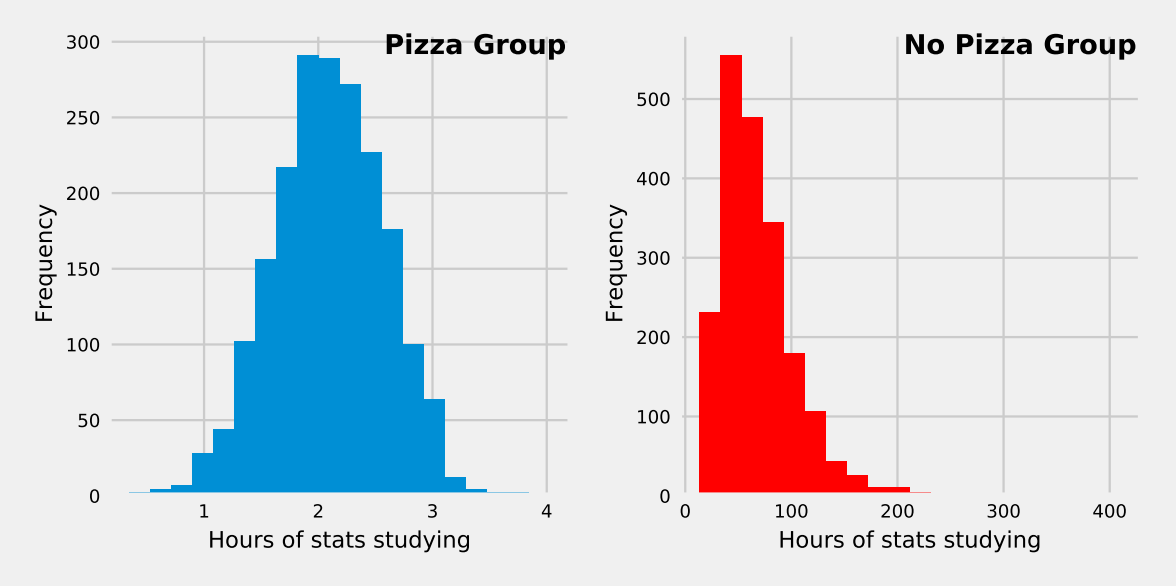
C

How many hours did student 34 in the non-pizza group study?

Question 03

Based on the positive results of this 200 student study you decide to run the largest pizza experimental study, comparing students who are assigned to a pizza vs non pizza group. You randomly assign 4,000 students to eating pizza versus not eating pizza and monitor how many hours they study statistics over a 2 week period.

The following histogram summarizes the data.



Please describe both of these histograms

A

Are any skew or symmetric?

B

What is (approximately) the mode of stats studying hours for each group?

C

Are these histograms unimodal? Multimodal? Why or why not?

Question 04

Please open up a jupyter notebook using Lehigh's Jupyter Cloud. Nagivate to the folder **BSTA001-2021S-CUL/h2** and open the file **Homework02Question04.ipynb**. Please complete the question in the notebook you opened, save them as a pdf, and include that report with the rest of your homework.