Question 01

Great work so far on solving the Flu-X outbreak! But now, it is time to experiment.

After intense laboratory experiments it appears a cure may be at hand—pizza. To confirm your scientific breakthrough you decide to run an experimental study that enrolls students infected with Flu-X. With a probability of 1/2 you assign students to eat pizza versus not eat pizza.

(A)

What is it called when you assign observations to different groups (eating versus not eating pizza)?

(B)

Are there any potential advantages to enrolling and assigning students infected with Flu-X to eat or not pizza versus an observational study that records student symptoms at a local pizzeria?

Question 02

Friends of yours, other biostatisticians not yet infected with Flu-X, and yourself meet to discuss data that was collected on campus. When everyone pools together their data you find we have several variables.

Please classify the following variables as numerical continuous, numerical discrete, categorical ordinal, and categorical nominal.

1. Student's age
2. Whether they have do or do not have Flu-X symptoms
3. Study's body temperature (elevated temperature is a sign of Flu-X)

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

4. Whether the student is a freshman, sophomore, junior, or senior.

Question 03

You want to see as soon as possible, is pizza a cure? After collecting data on 5 infected students to the pizza group and 5 infected students to the non eating pizza group you observe the following data
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 variance 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 04

(A)

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

$$S = \sigma_{i=-3}^{3} \frac{x_i}{i} \tag{1}$$

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?

After rigorous experimentation you find the mean hours of studying statistics from 100 students in the pizza group equals 1.2 hours. Students in the non-pizza group averaged 3.2 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 s number of hours of study.
You do, however, know that in the pizza group the total number of hours was 224 and in the non-pizza group the total number of hours was 445.

B

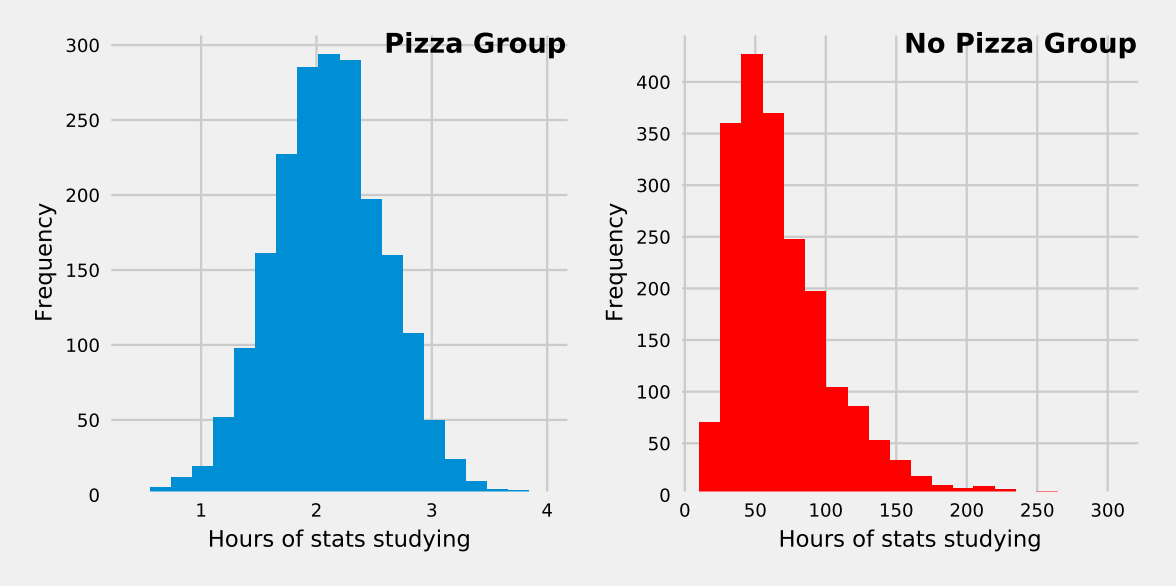
How many hours did student 72 in the pizza group study?

C

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

Question 04

Based on the positive results of your experiment, you and your biostat friends plan to run a larger experimental study that compares 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 unimodel? Multimodel? Why or why not?