University ID : 2001274638
Name : Dilip Nikhil Francies

For each of the following scientific questions, state whether the question is better answered using a randomized experiment or an observational study, and briefly explain why.

(a) Does the flu vaccine prevent the flu?

Answer : In my opinion, I believe a randomized controlled experiment can answer the question of whether administering a flu vaccine plays a factor in preventing the flu. For instance, consider a randomized experiment conducted during a flu season where individuals are divided into a control group and treatment group where the latter is injected with a vaccine and the former group is not. The dependant variable here is to check the effectiveness of the flu vaccine. Because we are controlling the factor of vaccine here, and dividing the group without or very minimal bias, once the experiment is carried out for a particular period of time frame, one can deduce if in fact the vaccine had an impact in the reducing the number of flu infected patients in the treatment group compared to the control group. But, in case of an observational study, the factor of "control" does not exist which makes it difficult to conclude the causality between flu and flu vaccine.

b) *Has support for same-sex marriage increased over time?*

Answer : Because the question being asked here is whether the support for same sex marriage is increased "over time", I believe an observational study is better in answering the above question. A sample of people [carefully chosen to eliminated any selection bias etc.] in an area where the study is conducted, and asking the right set of questions and comparing the mindset, attitude, of what the situation was an year back, or 5 years back, or even decades through the the help of past surveys and historical data would be much more feasible to understand if the support has grown over years.

c) Does banning laptops in class improve exam scores?

Answer : The above scientific question can be better answered with a randomized controlled experiment. By dividing a class into two groups where one group (treatment) were prohibited to use laptops in their classes and the control group were free to use their laptops, and eventually comparing the overall grade once their exams are finished will provide us with a satisfied information if their was any causality in grades and laptop usage in class. Even though there are additional independent variables such as what the laptop was being used for, or was the treatment group prohibited to use laptop not only in class, but also at home to study, does play a factor, a randomized controlled experiment is still the best way to understand if using laptops in class would decrease or increase students scores.

d) Are registered voters in the U.S. who earn over $100,000 per year more likely to vote Democratic than registered voters in the U.S. who earn less than $100,000 per year?

Answer : I believe the above questioned can be answered through an observational study rather than a controlled experiment mainly because it is not ethical to divide a group of random people who earn more and less than 100k and then control the factor of who they vote for? Or for that matter, the vice-verse is not possible as well. Whereas in an observational study, data can be gathered through public without interfering or controlling any of the factors, and then using the right statistical techniques to deduce the relationship between two variables.

e)Does bacon cause colorectal cancer?

Answer: This definitely can be answered through a randomized controlled experiment. By providing the treatment group with the right amount of bacon and studying the patterns over a period of time, one can deduce the causal relationship of bacon to colorectal cancer by having other independent variables such as smoking, physical activity, age etc in check. Eventhough, it is not moral to provide huge quantities of bacon until someone develops colorectal cancer, I still believe for the sake of research, randomized controlled experiment is still the best way to understand the relationship with two factors.

Question 2 :

According to a study done at Kaiser Permanente in Walnut Creek, California, women who use oral contraceptives ("the pill") have a higher rate of cervical cancer than women who do not use the pill, even after adjusting for age, education, and marital status.

A : *Was this study a blind randomized experiment or an observational study?*

Answer : Based on the information provided in the passage, I believe the study was an observational study where in the data was obtained by conducting a survey of women who had and had not used the "pill" in the past and later a causal reference towards cervical cancer was made. If it was randomized blind experiment, the women might have been divided into control and treatment groups, and effect of "the pill" on the treatment group would then be studied closely by making sure other factors would be applicable equally to all of them.

B) Does the study prove that the pill causes cervical cancer? Explain why or why not?

Answer : No, the study does not prove that the pill causes a cervical cancer. To what extent the pill plays a factor in causing cervical cancer is a different question altogether. There are other confounding variables that needs to be considered if we are trying to deduce that "the pill" has direct correlation to cervical cancer. For instance, information such as usage of tobacco, the kind of diet the women in the study followed, the sample that was considered may or may not have been big enough to generalize on the entire population, selection bias, or for that matter it might have been a random chance. Even though the study was adjusted for age, education and marital status, more rigorous and thorough research and experimentation is required to prove the casual relationship between the pill and cervical cancer.

C) *Besides age, education, and marital status, what other factor(s) related to cervical cancer are different (on average) between women who use the pill and women who don't?*

Answer :

1. Tobacco consumption : One of the major causes of cancer in human beings irrespective of gender is tobacco usage of any form.
2. Diet, supplemets,Hydration : Poor diet, and water usage that consists of heavy elements such as Lead,Mercury,Arsenal and Nickel have been proved to cause cancer. So the kind of diet women followed, and access to good quality water and food, I believe is a major factor that should have been added along age, marriage and education.
3. Geography, Medical and Family history : Perhaps, the women in question were chosen from an area where cervical cancer was profound due to other medical ailments, or their family history. Choosing the right representative sample without any hidden bias is one of the crucial factors as well.
4. Hygiene : Sexual hygiene, number of sexual partners [and their hygiene] does play a factor as well in carrying viruses that may cause cancer.

Question 3:

The Center for American Progress to study public attitudes toward sports teams that ex-pressed opinions on issues that could be controversial, such as LGBT issues. One research question they were interested in was: "To what degree do people believe that professional sports teams should take public stances on social causes?"In general, respondents either "somewhat agreed" (33.2 percent) or "strongly agreed" (19.3 percent) that "Professional sports teams should utilize their platforms to advocate for causes they believe in." One in three respondents (30 percent) stated they were neutral on this issue or had "no opinion." Among respondents who identified as men, 46.5 percent either "somewhat agreed" or "strongly agreed" that sports teams should use their platforms to advocate for causes, while 33 percent had "no opinion." Among respondents who identified as women, 60.4 percent either"somewhat agreed" or "strongly agreed" with that statement, with another 25.6percent stating that they were neutral or had "no opinion."1The study also gave the following information about their survey:

The results presented above are from a convenience sample of 367 respondents recruited using Amazon Mechanical Turk, an online platform that allows for the purposeful sampling of respondents who meet relevant criteria. This survey-hosting website has been shown to be an efficient platform for gathering reliable data from diverse populations. . . 44.1 percent of respondents were between the ages of 18 and29, 38.3 percent were between the ages of 30 and 44, and 17.8 percent were above the age of 44.

(a) Does this survey give a statistically unbiased picture of how much of the general U.S. adult public would agree or disagree with the statement: "Professional sports teams should utilize their platforms to advocate for causes they believe in"? If not, describe and ex-plain the likely direction of the bias—is the survey likely overestimating the proportion of the public who agree, or the proportion who disagree?

Answer : The study does not give an accurate representation of the general opinion of the United States population regarding the question *"To what degree do people believe that professional sports teams should take public stances on social causes?".* It is more than likely to overestimating the ratio of population that agrees with the statement solely relying on sample of unrepresentative data filled with bias to draw conclusion over a general consensus. Some of the bias that I believe that exits in the data are as follows.

1. Sample chosen : According to Wikipedia[link], the percent of population that is above the age of 44 corresponds to a total of 50% of the entire population as of 2021. However, in the sample survey, only 17.8 % of the people who voted are above the age of 44, which in my opinion is a case of unrepresentative sample.

2. Random sampling : The fact that, people who had access to vote on Amazon Turk were considered a sample, and moreover the total number of voters being only 367 in total was considered to draw a conclusion clearly indicates that there was no big enough random sampling done to mimic the general population of the United States. Additionally, What about people who had no access to Amazon Turk Platform or even in which states and cities was this platform live to draw the conclusion based on only 367 votes?

3. What happened to the remaining percentage of men and women who did not respond with any of the options? Did they strongly disagree? There is lack of data here as well.

Hence I believe, a better representative data, and the medium of survey would have made a significant difference in conducting the study and to draw a large enough conclusion of the US population.

(b) Suppose an interested party gives you a reasonable budget to carry out a more rigorous study of the Center for American Progress' research question. Describe briefly the study you would perform:

Answer : I would mainly do the following changes to make the study more rigorous and representative:

1. Data Gathering : instead of solely relying on Amazon Turk, I would gather data from different mediums such as online survey, paper based survey, telephone calls, text messages, social media voting, through camps outside stadiums etc making sure I can reach out to all different types of people from different age to sex, from sports fans to non-fans, and include all the races.

2. Random Sampling : To make sure the sample is big enough and represents the entire population, I would carry out the data gathering in as many states, cities, suburbs as possible.

3. Have options other than strongly agree, strongly disagree, neutral so that I can obtain votes of wider range, and not just the first three like the study mentions.

4. Employ better statistical techniques to make sure any analysis done during and after data gathering adheres to proper statistical methods in conducting a study so that a well proven conclusion can be drawn, making sure the money spent does not go to waste.