**Quasi-Experiments: Activity**

1. Data scientist David Zhang wrote on his blog about the potential impact of Queen’s Gambit on interest in chess playing. He was able to analyze traffic on the website, chess.com, over the past several months. He analyzed the number of new users who started new accounts on the site. Chess.com is the most popular chess community on the Internet.

Zhang graphed the number of games played by United Kingdom (UK) players by day from July 2020 (three months before the Queen's Gambit was released) to March 2021 (five months after). The Netflix show appeared in October, 2020. If you want you can find the story here: [**this story link.**](https://towardsdatascience.com/how-has-the-queens-gambit-impacted-the-popularity-of-online-chess-43594efe5a98)

Chart, histogram

Description automatically generated

a) This graph, by itself, should remind you of one of the four quasi-experimental. Which one?

b) What is the quasi-independent variable in this study? What are its levels?

c) What is the dependent variable in this study?

d) When we analyze how well a quasi experiment can support a causal claim (such as "Queens Gambit caused people to join chess.com"), we need to consider the design and the results. What do you think--do the design and results show here help us rule out any internal validity threats?

e) Many folks would nominate *history threat* as a likely issue for this situation. That is, they might suggest that some other event, other than Queen's Gambit, occurred in October that could have caused chess to become more popular. What do you think--is that a reasonable critique?  If so, what historical event would you nominate that would have occurred at the same time (October 2020), and would have affected most people in this UK sample?   
(notice what the author of this post writes about a potential history threat:

....other chess events have occurred around the time frame analysed. For example, the first [PogChamps](https://en.wikipedia.org/wiki/PogChamps) tournament involving popular streamers was hosted in May 2020 by chess.com, which is likely to have spread awareness and garnered appeal for online chess.)

f) Other folks might say that maturation is a threat here. They might say that "perhaps people just naturally and spontaneously became more interested in chess over time." What helps rule this out, or make it implausible?

2. Since the COVID-19 pandemic began, several governments, including the U.S. federal government, have distributed cash payments to citizens in an attempt to alleviate the  financial hardship caused by the global shutdown.

As the [**New York Times reports here,**](https://www.nytimes.com/2021/06/02/us/politics/stimulus-checks-economic-hardship.html)policymakers wonder if such payments are useful. Now we have data to help address that research question.

In offering most Americans two more rounds of [stimulus checks](https://www.nytimes.com/2021/03/17/your-money/stimulus-payments-checks.html) in the past six months, totaling $2,000 a person, the federal government effectively conducted a huge experiment in safety net policy. Supporters said a quick, broad outpouring of cash would ease the economic hardships caused by the coronavirus pandemic. Skeptics called the policy wasteful and expensive. [...]

A new [analysis of Census Bureau surveys](http://sites.fordschool.umich.edu/poverty2021/files/2021/05/PovertySolutions-Hardship-After-COVID-19-Relief-Bill-PolicyBrief-r1.pdf) argues that the two latest rounds of aid significantly improved Americans’ ability to buy food and pay household bills and reduced anxiety and depression, with the largest benefits going to the poorest households and those with children. The analysis offers the fullest look at hardship reduction under the stimulus aid.

The full report is [**available here,**](http://sites.fordschool.umich.edu/poverty2021/files/2021/05/PovertySolutions-Hardship-After-COVID-19-Relief-Bill-PolicyBrief-r1.pdf) and provides some excellent graphed examples of quasi-experimental data. We'll focus on Figure 1 from the report, reprinted above with permission from the authors.

The y-axis on the figure represents the percentage of Americans who "sometimes or often had not enough food to eat in the last seven days." The question is asked by U.S. Census' Household Pulse survey about once per month to a random sample of Americans. The dates of the two U.S. stimulus checks are indicated in the laddered, light-blue vertical lines. Payments occurred in December, 2020 and March, 2021.

Chart, line chart

Description automatically generated

a) The journalist writes that "the federal government effectively conducted a huge experiment in safety net policy".  What is the independent variable in this "experiment?"  What is the dependent variable (i.e., the one depicted in Figure 1)?

b) The independent variable here is not a true IV--it is a quasi-independent variable. What makes this variable *quasi*-independent?

c) The authors claim that the results from Figure 1 are consistent with the argument that the stimulus checks helped people. The authors wrote in their report: "Based on the speed with which we see hardship fall, we suspect much of this drop was the result of EIP checks, which the federal government was able to quickly deliver to bank accounts for most U.S. households following passage of both bills." Locate the areas on the figure that the authors are referring to.

d) This quasi-experimental study is probably best described as an interrupted time-series design. What is the "interruption" here? What is the time series?

e) A history threat is one potential internal validity problem in an interrupted time-series design. Explain what a history threat is and also why a history threat *might* be a problem here. What do you think--to what extent can we be sure it was the stimulus checks, and not some other pair of events, that is responsible for the increased food security depicted here?

3. Examine the situations below and identify potential, plausible threats to internal validity.

a) A researcher was interested in the effects of embarrassment on people's willingness to comply with another person's request. (Are people more willing to do what other people ask after they've made a fool out of themselves?) The researcher randomly assigned 30 participants to one of two conditions. Participants in the no embarrassment condition were told that they would simply listen to 5 minutes of disco music. Participants in the high embarrassment condition were told that they would dance to 5 minutes of disco music in front of another person. After having the study described to them, three of the participants in the high embarrassment condition declined to participate in the study, and the researcher let them leave. After receiving the experimental manipulation (no vs. high embarrassment), participants were asked if they would volunteer for another hour‑long study. The proportion of participants in each condition who complied with the request was recorded and analyzed.

b) An educational psychologist was interested in determining whether a new teaching technique results in higher academic achievement than traditional teaching methods. To compare the two techniques, she identified a school system that taught in the traditional manner and another school system that used the new technique. She administered an achievement test to the children in the two schools and found that the children in the school that used the new technique significantly outperformed the children in the traditional school.

c) A study investigated the hypothesis that excessive sugar increases children's activity level. Fifty kindergarten children were randomly assigned to drink either a sugar‑sweetened drink or an artificially‑sweetened drink. Afterwards, the children engaged in free play for 30 minutes as their behavior was observed. Two observers rated the activity of level of each child's behavior every five minutes for an hour. To avoid the possibility of fatigue setting in, two observers watched the children for the first 30 minutes and two other observers watched the children for the second 30 minutes.

d) People who take the Scholastic Aptitude Test and get an extremely low score, sometimes take the test twice. When they take the test a second time, they are likely to get a score slightly closer to the mean SAT score for the overall population of SAT test takers. What term describes what is occurring here?