IDS 701 -- Spring 2022 Causal Inference Project Rough Draft

Satvik Kishore, Tego Chang, Jaya Khan, Fides Regina Schwartz

TOTAL POINTS

0.85 / 1

QUESTION 1

1 Rough Draft **0.85 / 1**

✓ - **0 pts** Even in a report to a data scientist (e.g., me), the only stuff about your data manipulations that you want to include are manipulations where where **your judgment** came into play, and where the decisions you made might impact the results. For example if you decided to drop certain observations because they were missing entries for certain variable, that is a judgment call you made that impacts the substance of your analysis and that should be shared. But something like reshaping the data is not something you need to include.

✓ - **0.15 pts** A reasonable start, but underdeveloped in a number of important ways. A lot of work still to do.

💬 Your writing is very thin / outline like, and lots of aspects of your project aren't spelled out. Please take a look at some of the other reports I send out later this weekend to see what a richer writing style might look like.

1 This comes filled back to the question of whether your analysis is tautological—why are you asserting that you think that the people making the most money are superhosts, then going on to actually study it is a worthwhile question?

2 It's not really clear how these two questions relate to one another. They both seem like they're kind of asking "what is the causal effect of being a super host." Provided that is what you are interested in, and you are just working with a cross-sectional regression, then you certainly want to use all those predictors!

3

What exactly are you scraping, and what do you need locations that look similar?

4 See note about what you need to include and what isn't relevant

5 Why?

6 Why wouldn't you just keep each listing as a separate observation?

7 ? I think you need to flesh this out—why is the minimum night value equal o the average minimum number of nights a renter stays?

8 wait you can't match most of your data? Surely you can do better than that, no? Some binning of variables or something? Is there really that little overlap? Can you provide some diagnostics on the matching?

9 Remember you aren't done when the algorithm stops; that's just the point when it is gotten stuck. It's up to you to evaluate whether that's the place you want to stop or whether you would rather use fewer iterations.

10 Why each of these specifications? What are they doing differently? Can you be more clear about your motivation?

11 Um... how do you have so many hosts at sea?

12 seeing anything off about this plot? Like... the hundreds of AirBnBs at sea?

ıl gradescope

1 Rough Draft 0.85 / 1

✓ **- 0 pts** Even in a report to a data scientist (e.g., me), the only stuff about your data manipulations that you want to include are manipulations where where **your judgment** came into play, and where the decisions you made might impact the results. For example if you decided to drop certain observations because they were missing entries for certain variable, that is a judgment call you made that impacts the substance of your analysis and that should be shared. But something like reshaping the data is not something you need to include.

✓ **- 0.15 pts** A reasonable start, but underdeveloped in a number of important ways. A lot of work still to do.

💬 Your writing is very thin / outline like, and lots of aspects of your project aren't spelled out. Please take a look at some of the other reports I send out later this weekend to see what a richer writing style might look like.

1 This comes filled back to the question of whether your analysis is tautological—why are you asserting that you think that the people making the most money are superhosts, then going on to actually study it is a worthwhile question?

2 It's not really clear how these two questions relate to one another. They both seem like they're kind of asking "what is the causal effect of being a super host." Provided that is what you are interested in, and you are just working with a cross-sectional regression, then you certainly want to use all those predictors!

3 What exactly are you scraping, and what do you need locations that look similar?

4 See note about what you need to include and what isn't relevant

5 Why?

6 Why wouldn't you just keep each listing as a separate observation?

7 ? I think you need to flesh this out—why is the minimum night value equal o the average minimum number of nights a renter stays?

8 wait you can't match most of your data? Surely you can do better than that, no? Some binning of variables or something? Is there really that little overlap? Can you provide some diagnostics on the matching?

9 Remember you aren't done when the algorithm stops; that's just the point when it is gotten stuck. It's up to you to evaluate whether that's the place you want to stop or whether you would rather use fewer iterations.

10 Why each of these specifications? What are they doing differently? Can you be more clear about your motivation?

11 Um... how do you have so many hosts at sea?

12 seeing anything off about this plot? Like... the hundreds of AirBnBs at sea?

ıll gradescope