

Backwards Design Template

November 17, 2019

1 Topic:

What is your project about?

2 Project Question

What specific question are you seeking to answer with this project?

3 Project Hypothesis

What is your hypothesized answer to your question?

4 Model Results

One of the hardest parts of developing a good data science project is developing a question that is actually answerable. Perhaps the best way to figure out if your question is answerable is to see if you can imagine what an answer to your question would look like. Below, draw the graph, regression table, etc. that you would consider to be an answer to your question. Then draw it

again, so you have a model result for if your hypothesized answer is true, and a model result for if your hypothesized answer is false. (If the answer to your question is continuous, not discrete (like: what is the level of inequality in the United States), draw it for high values (high inequality) and low values (low inequality)).

Result if you hypothesis is true

Result if you hypothesis is false

5 Final Variables Required

Now that you've specified what an answer to your question looks like, what data do you need to generate that answer.

For each variable, define both the variable you need **and** the population for which you need the variables to be defined.

You don't have to be too specific ("I need variable 7 from the NHGIS 2019 census 1% sample release") – just define it in the most general terms that are still specific enough to meet your needs (e.g. I need income data for a nationally representative sample of US citizens).

6 Data Sources

Finally, given the variables you need for your analysis, what actual data sources do you think will have the data you need?

*In specifying the datasets you need, if you list more than one **also** indicate how you think you can relate these datasets (i.e. if you're gonna merge them, what variables do you think those datasets will provide that will allow you merge them? There's no use saying "I'll merge this political survey with medical records of who has received bad care" if the political survey doesn't provide identifying information you can use to link survey respondents to medical records, even if you have both the survey and medical records!)*