Research Question and personal codebook assignment

Purpose

Learn more about your chosen research data and formulate a Research Question. This will guide your focus for the rest of the semester. In this assignment you will

1. Decide on a specific research topic
2. Choose variables that relate to your topic
3. Formulate a Research Question

Instructions

- You will conduct the work directly in a Google document in the `02 Research Quesetion and Codebook` folder.
- Name this file `rq_userid`
- Drop a note in Blackboard Learn when your document is ready to grade.
- Use header formats to clearly denote steps 1-3.

Step 1: Topic area statement:

Thoroughly review the codebook for your dataset of choice and make note of topics that interest you. Use this as a brainstorming session and identify as much as you can. Write a statement similar to the one below regarding what topic(s) you are interested in. You should also explain WHY you are interested in your topic.

> Example: After looking through the codebook for the U.S. Longitudinal Study of Adolescent to Adult Health study database, I have decided that I am particularly interested in nicotine dependence. I am not sure which variables I will use regarding nicotine dependence (e.g. symptoms or diagnosis) so for now I will include all of the relevant variables in my personal codebook. While nicotine dependence is a good starting point, I need to determine what it is about nicotine dependence that I am interested in. It strikes me that friends and acquaintances that I have known through the years that became hooked on cigarettes did so across very different periods of time. Some seemed to be dependent soon after their first few experiences with smoking and others after many years of generally irregular smoking behavior. I decided that I am most interested in exploring the association between level of smoking and nicotine dependence. Below I have added to my codebook variables reflecting smoking levels (e.g. smoking quantity and frequency). If my topic is too vague, I would like to bring suicide in. I would be interested to see if there is any relationship between level of smoking with nicotine dependence and suicide.

Step 2: Personal Codebook:

Create your personal codebook. You will create a trimmed down version of the codebook to help you not get overwhelmed with the larger pool of available questions.

- In your personal codebook, include the questions/items/variables from the full data codebook that measure your selected topics.
- Choose between 10 and 20 variables. This is just a starting point, you can modify anytime later.
- **Do not go overboard here. Excessive variables being chosen indicate lack of thought and will not be reviewed highly**

- You can do this by taking little screenshots with the snipping tool. In the pdf file, go to Edit, Take a Snap Shot, then click and drag over what you want. If you are on a MAC, press Shift+Command+Ctrl+4 to get the snip tool.

Step 3: Research Questions:

Now that you have a few topics and picked out variables for each topic, what questions are you interested in trying to answer about your topics?

- Start by asking whether two constructs are associated. This is referred to as a **testable hypothesis**. A testable hypothsis is one that can be answered with a "yes" or "no".
- Good RQ's are made up by asking how does 1 explanatory variable affect 1 response variable. Stick with two variables only. You have to start somewhere. You can always build up but you have to start with 2.
- Be explicit with your measures. If you want to measure smoking quantity, you should have a variable in your personal coebook (step 2) that measures smoking quantity.
- Avoid being vague. This is easy to do.

  Example:
  a) Does smoking quantity relate to nicotine dependence?
  b) Is there a relationship between smoking frequency and suicide?
  c) Does a smoker have more suicidal behaviors than a non-smoker?

Now: Re-read your question out loud. Does it make sense? Can you clearly identify what your IV and DV (see below) are?

A note about the `explanatory` vs `response` variable terminology:

Depending on your field of study, these terms may be called different things. Statisticians tend to say *covariate* and *outcome*. Public health and the Social Sciences tend to use Independent variable (IV) and dependent variable (DV). Regardless of the word you use... the explanatory variable tries to **explain** the response variable.

Explanatory	Response
X	Y
independent variable	dependent variable
covariate	outcome