Homework #1a: Preliminaries

Math 4334: Mathematical Modeling
Dr. Scott Norris

August 20, 2023

As we have not yet learned very many sophisticated modeling tools, we will use this assignment to practice some of the broader skills associated with mathematical modeling, such as

- forming and justifying hypotheses
- obtaining, organizing, and visualizing data
- evaluating hypotheses in light of new data

Please respond to each prompt below with a moderately-sized paragraph, accompanied by a graph as directed in each prompt. We'll have a little bit of fun by exploring some moderately "hot" topics. My interest is not to push a specific *opinion* on these topics. Rather, it is to force ourselves to practice using a data-oriented approach to "real-world" topics, even ones where we might have pre-existing opinions.

Problem #1. Hot Topic: Immigration. A commonly-encountered claim is that "Immigrants depress wages and take jobs away from native-born Americans." As a hypothesis, this seems plausible, as immigrants are often stereotyped as being willing to work for lower wages than people born here, and the claim would then seem to follow from the laws of supply and demand. However, before accepting this claim, let's look for some data on the matter.

- Download data from https://www.latticepublishing.com/blog/states-with-the-most-immigrants on the percentage of each state's population that are immigrants (i.e. foreign-born).
- Find data on the *median* wages in each state (don't use *mean* wages, which are unrepresenative).
- Merge by state, and display as a well-formatted scatter plot with immigrant pct. as (x), and median wage as (y). [NOTE: It is typical to remove Washington, DC from such plots – why?]
- Discuss what you found. Address the following questions.
  - Where did you obtain your data? How reliable do you think this source is?
  - Does the hypothesis stated above appear to be consistent with the data?
  - If not, then formulate a hypothesis that you believe *could* explain the data (could include null hypothesis).

Problem #2. Hot Topic: Tax Rates. A commonly-encountered claim is that "High taxes strangle the economy. Low taxes increase economic growth, which benefits everyone." As a hypothesis, this seems plausible, as taxes take money away from people that they might otherwise spend to improve their standard of living, which in turn contributes to the economy. However, before accepting this claim, let's look for some data on the matter.

- Download data from https://wallethub.com/edu/states-with-highest-lowest-tax-burden/20494 on the state-level "tax burden" [combining property, income, and sales taxes].
- Find data on the per-capita GDP of each state (not necessarily the same as mean wages).
- Merge by state, and display as a well-formatted scatter plot with tax burden as (x), and per-capita GDP as (y). [NOTE: It is typical to remove Washington, DC from such plots – why?]
- Discuss what you found. Address the following questions.
  - Where did you obtain your data? How reliable do you think this source is?
  - Does the hypothesis stated above appear to be consistent with the data?
  - If not, then formulate a hypothesis that you believe *could* explain the data (could include null hypothesis).

Problem #3. Hot Topic: Gun Violence. A commonly-encountered claim is that "If guns are outlawed, then only outlaws will have guns. An armed society is a polite society" – i.e. widespread gun ownership keeps us safe. As a hypothesis, this seems plausible, because if someone is breaking into my home, I would like to have a way to defend myself – especially if that person is also armed! However, before accepting this claim, let's look for some data on the matter.

- Extract data from Kalesan et al, Injury Prevention **22** (2016) pp. 216-220, on the the rate of gun ownership in each state, measured as a percentage (yes, you'll have to type it out by hand into a spreadsheet!).
- Find data from the US on the per-capita rate of gun-related deaths in each state, (i.e. annual deaths / population).
- Merge by state, and display as a well-formatted scatter plot with ownership rate as (x), and death rate as (y).
- Discuss what you found. Address the following questions.
  - Where did you obtain your data? How reliable do you think this source is?
  - Does the hypothesis stated above appear to be consistent with the data?
  - If not, then formulate a hypothesis that you believe *could* explain the data (could include null hypothesis).

Problem #4. Equal opportunity. The previous three questions could – from a certain perspective – be seen as picking on conservatives. Here let's turn the tables. Describe a belief associated with U.S. liberals that you believe is (a) commonly held, and (b) contradicted by available data. Provide evidence (such as a link to a newspaper article) that this belief is, in fact, commonly held. Then find and cite data that you believe contradicts this belief, and present it as a graph.

Problem #5. Topic of your choice.

- Find data on *any two variables that interest you*, that you think might be correlated.
- Plot the data in a well-labeled scatter plot. Include axis labels and a title.
- Formulate a hypothesis that you believe could explain the observations.
- What additional data might you gather to test this hypothesis?