Rhys and Rihaan MAP - Spring 2026

Github shared with Professor Kuiper

Defining the parameters for our study:

1) We begin by breaking our study down into several parts, including but not limited to:

    - Overarching metric we need for everything: population of US + county-wise
    - Economic indicators: Fed interest rates, Unemployment rates, Inflation
    - Health: AQI, BMI, blood pressure, lifestyle factors (alcohol, smoking, etc)
    - Personal: consumer sentiment

2) Standard range of years for most datasets we will be using → 1990 to 2022, maybe beyond that too. This would give us 32 years of data, which we feel is sufficient to make analyses and theories off. This also gives us a range across various shocks (dotcom crash, subprime mortgage crisis, covid pandemic, etc).

Confused about: should we have one overarching research question, or multiple questions for each subsection of data we are identifying?

See the data first

Potential research topic ideas:

1. Predicting high risk areas by county in the united states using several economic indicators and trends

Datasets found:

1. U.S. population by year (1990-2022): https://fred.stlouisfed.org/series/POPTOTUSA647NWDB
2. County Health 2021 dataset uploaded on Github
3. US Population by Race and County (Needs Discussion): https://seer.cancer.gov/popdata/download.html
4. US Freight Analysis Data: https://www.bts.gov/faf/county (This requires conversation)
5. SOI tax stats - Migration data https://www.irs.gov/statistics/soi-tax-stats-migration-data?
6. Personal Income by County (2023): https://www.bea.gov/data/income-saving/personal-income-county-metro-and-other-areas
7. GDP by County, Metro, and Other Areas (2023): https://www.bea.gov/data/gdp/gdp-county-metro-and-other-areas
8. Longitudinal Employer Household Dynamics (needs further research): https://lehd.ces.census.gov/data/

Start by looking at just one year (2020)

Migration data to see how people are moving (republican vs democratic) increase or decrease post-elections. Where are they moving from, and where are they moving to?

Relates to taxes. State income tax rates, property tax

Transit freight data

Interactive data viz

World factbook data

Debunk overgeneralizations

What does the leadership look like

Social indicators

Then expand to more years

Part 2 of our study will look at, and try to find solutions to the extensive data cleaning processes students have to carry out when working with real-world and messy datasets. We will be looking at solutions using Python and R Studio.