1. Background / Motivation / Summary of the Project

Given recent push-backs and controversy surrounding vaccines, viruses, and health information, we want to understand and examine the specific factors, and in particular, education, and their effects on recent controversy and conflicts. This debate has only been exacerbated by the coronavirus pandemic of 2020, but has existed in the United States for many years, with anti-vaxx debates sparking major protests throughout the country. In fact, there have been many outbreaks in the US of major viruses such as measles in the past year (2019), particularly in Washington and Oregon.

2. Goals

Goals:

- Comparison of education by state to vaccination rates
    - Comparison of the "level" of education of state populations (e.g. percentage of high school graduates, college graduates, etc.) as well as the amount spent per capita on education by state (to try and reduce the effect of the difference in public funding).
- Exemptions/legislation by state compared to vaccination rates as well as number of cases
- Public perception of healthcare institutions - including their recommendations, requirements, and information provided

3. Datasets You Plan to Use/Collect

https://www.cdc.gov/vaccines/imz-managers/coverage/childvaxview/data-reports/mmr/reports/2017.html

https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwi_ve_RjP7rAhUQmHIEHfJgAvAQFjACegQIAhAB&url=https%3A%2F%2Fwww.census.gov%2Fnewsroom%2Freleases%2Fxls%2Fcb12-33table1states.xls&usg=AOvVaw1zrklipwQ-m4zRTFshv0nZ

Average standardized testing scores by state (SAT, ACT, etc.) as a measure of educational abilities by state

https://www.cdc.gov/phlp/publications/topic/vaccinations.html

https://ourworldindata.org/roush-and-murphy-2007-data (change in cases after vaccination)

https://ourworldindata.org/grapher/immunization-measles-vs-vaccine-safety?tab=chart&stackMode=absolute&country=&region=World (reference)

https://www.kff.org/state-category/health-costs-budgets/

https://www.cdc.gov/vaccines/programs/vfc/awardees/vaccine-management/price-list/index.html

https://news.gallup.com/poll/276929/fewer-continue-vaccines-important.aspx

https://www.cdc.gov/vaccines/pubs/pinkbook/downloads/appendices/G/coverage.pdf

4. Questions We May/Are Answering

- How does education correlate with the amount of cases (based on population, for example: per 100,000 people)?
  - If we control for several other key variables, is there a correlation between education and the amount of cases?
    - Key variables include population density, public opinion on vaccines and medical professionals, cost of vaccines (may be tough to get), state spending on healthcare, etc.
- Are people with a high school diploma more likely to be vaccinated? What about people holding a bachelor's degree? Advanced degrees?
- Comparison of virus cases in states with the highest vaccination rate versus lowest and a comparison of education measurements between those two specific states, taking into consideration factors such as location and city density. May also use states with "medium" or in-between vaccination rates for comparison or a baseline.
  - Different routes to take: can look at past virus transmission and education or the current coronavirus situation.
- What exemptions do states have? Do these affect vaccination rates?
- How does public opinion of medical professionals affect vaccination rates?
  - Examine the effects of the gallup poll, which shows that less Americans think it's "very important" for children to be vaccinated now than in previous years

(https://news.gallup.com/poll/276929/fewer-continue-vaccines-important.aspx)

- Compare this poll with the vaccination rates for the country as a whole in the past few years (corresponding to the dates in the poll) (vaccination rates: https://www.cdc.gov/vaccines/pubs/pinkbook/downloads/appendices/G/coverage.pdf)

5. Approach

1. Gather all datasets (including polls, surveys, and previous research/papers that may be referenced) that are necessary for analysis and evaluation
2. Clean all data down to relevant statistics
3. Explore whether adding a composite variable for education is a good idea or not (research needed on composite variables).
4. Determine what other factors
5. Create models, test, analyze and repeat (as we examine the results)
6. Generate visualizations
7. Write writeup