Public Health Policy Intervention

What is the problem you want to solve?

- The focus of the project is to forecast metrics such as correlation and covariance between GDP and various diseases present in the United States at the county level with the intent to push for some public health policy.

Who is your client and why do they care about this problem? In other words, what will your client do or decide based on your analysis that they wouldn't have done otherwise?

- The issue will be relevant to any US government agency as well as any health organization. Public Health is relevant at all levels of government and business. If a disease is on the rise and some relationship is able to capture its effect on economic factors such as GDP and employment, everyone will want to know, especially policy makers who need to know what to do with public funds to prevent both epidemic and/or economic crisis.

Briefly outline how you'll solve this problem. Your approach may change later, but this is a good first step to get you thinking about a method and solution.

- The goal of the analysis is to build a predictive model that can take more than two factors and accurately forecast their location on a 2 dimensional plane. Factors of interest will be diseases, counties, sex, year, population, GDP, GDP growth, unemployment, and all other related economic factors.
- I would like to use random forests for feature selection and neural networks for forecasting. If time permits, I will attempt to forecast with a simple regression model and compare with the results of a neural network model.

What data are you using? How will you acquire the data?
- Data: unemployment, GDP, GDP growth, and diseases, all by county
- Bureau of Economic Analysis API, California Health and Human Services Data Portal.
- https://data.chhs.ca.gov/dataset/infectious-disease-cases-by-county-year-and-sex/resource/741f25e1-db50-436a-a6a9-7b840176edbf
- https://apps.bea.gov/iTable/iTable.cfm?acrdn=7&isuri=1&reqid=70&step=1#acrdn=7&isuri=1&reqid=70&step=1

What are your deliverables? Typically, this includes code, a paper, or a slide deck.
- Jupyter notebook with data visualization and stats.