Lisa’s Role Project 2, Murican Labs

Data cleaning: Data from both sources, UNICEF and World Bank, were joined with an outer merge using pandas to determine which countries from World Bank needed to be renamed to match the naming used by UNICEF. Only countries that appear in both data sets were kept. The vaccination data was restructured using a loop to take each year column and stack the information so that year is a single column and the vaccination coverage percent was another column. Missing values from the stacked data were removed. The infant mortality and life expectancy data were merged into a single data set. Primary and merge key values were added.

Python Flask coding: Created the framework of the app.py, template/index.html, static/js/app.js files to connect to the ProgresSQL database and create a hello world template to build on.

Time Series Visualization: Created drop downs to allow a user to choose a vaccine and a country. Pulled and filtered, creating a path to store jsonified data. Generated Plotly time series line graphs of filtered data.

Analysis: As with the time series visualization, drops downs allowed the user to choose a vaccine and country, and the filtered data was store in a path defined in app.py. Conducted regression analysis on the filter data in python using SciPy. Defined a path in app.py to store jsonified regression analysis results. Regression fit plot and related plots to check assumptions were created using Plotly. Provided description of regression analysis to ad to webpage.