**MLM model - ARIMA**

Originally we tried to do a linear regression model from 4 different data sets with data by year from 2010 or 2015. Each dataset (GDP, unemployment, employment, and population) we ran a linear regression model to predict 2024 values and then took the rate of change. Then we merged all the datasets into one large 2024 prediction data frame and ran a logistic regression model to rank the top cities to move too.

After discussing with James about how this was a time series model we pivoted and are now running 4 separate ARIMA models to predict the values. We ran ARIMA models on GDP, Unemployment, Population, and Employment. Once this model is run on each dataset then then we will run a simple classification model that will rank/score the cities to decide if these cities are “Emerging” or “Static”.

To run this ARIMA model we had to find and/or convert our data sets from yearly into monthly to have more datapoints to make the model fit the data.

After trying monthly data, we realized we had to convert this data back to yearly, as yearly data better fit the ARIMA model.