Mechanical & Industrial Engineering
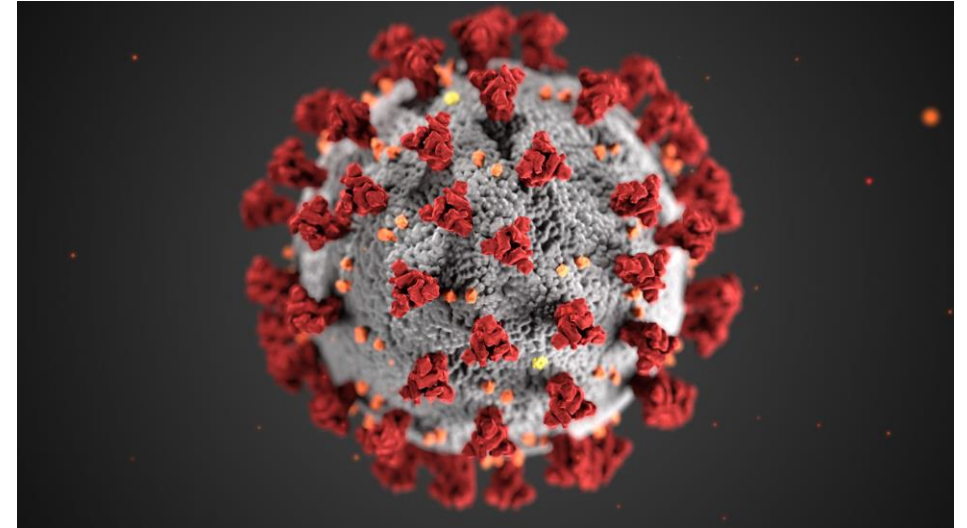UNIVERSITY OF TORONTO

MIE 1624: Final Project Presentation
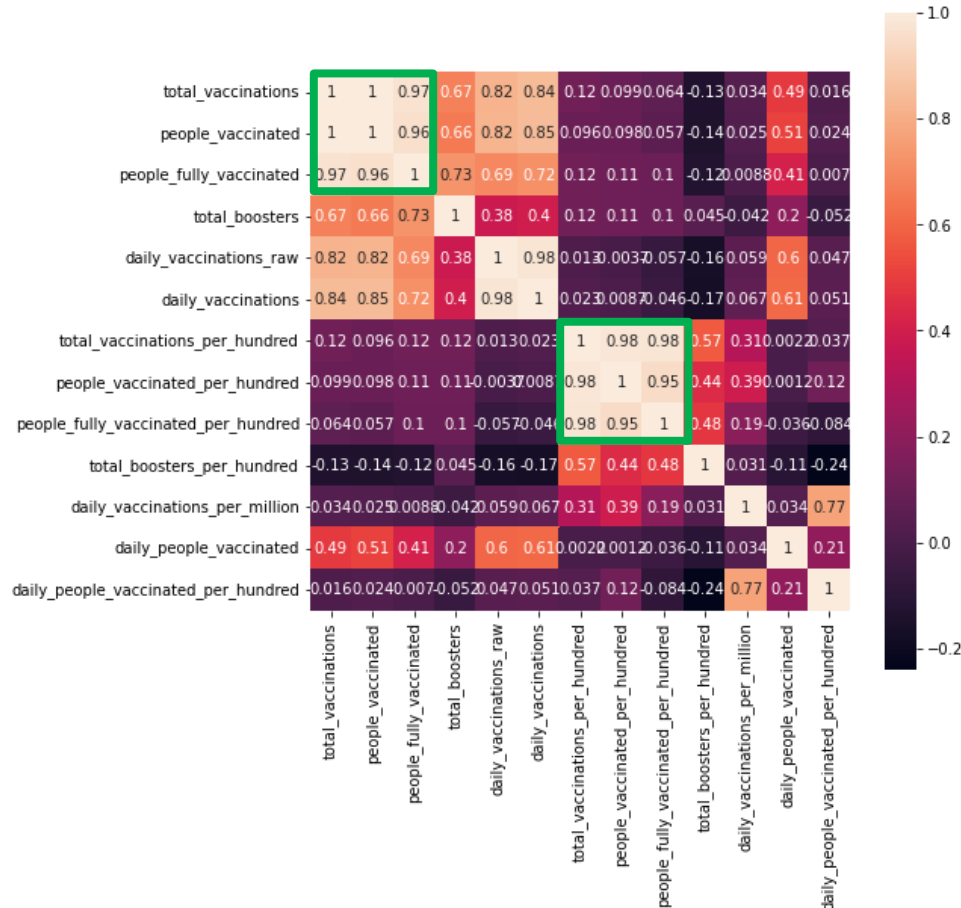
COVID-19 Vaccination Dataset

Fall 2021

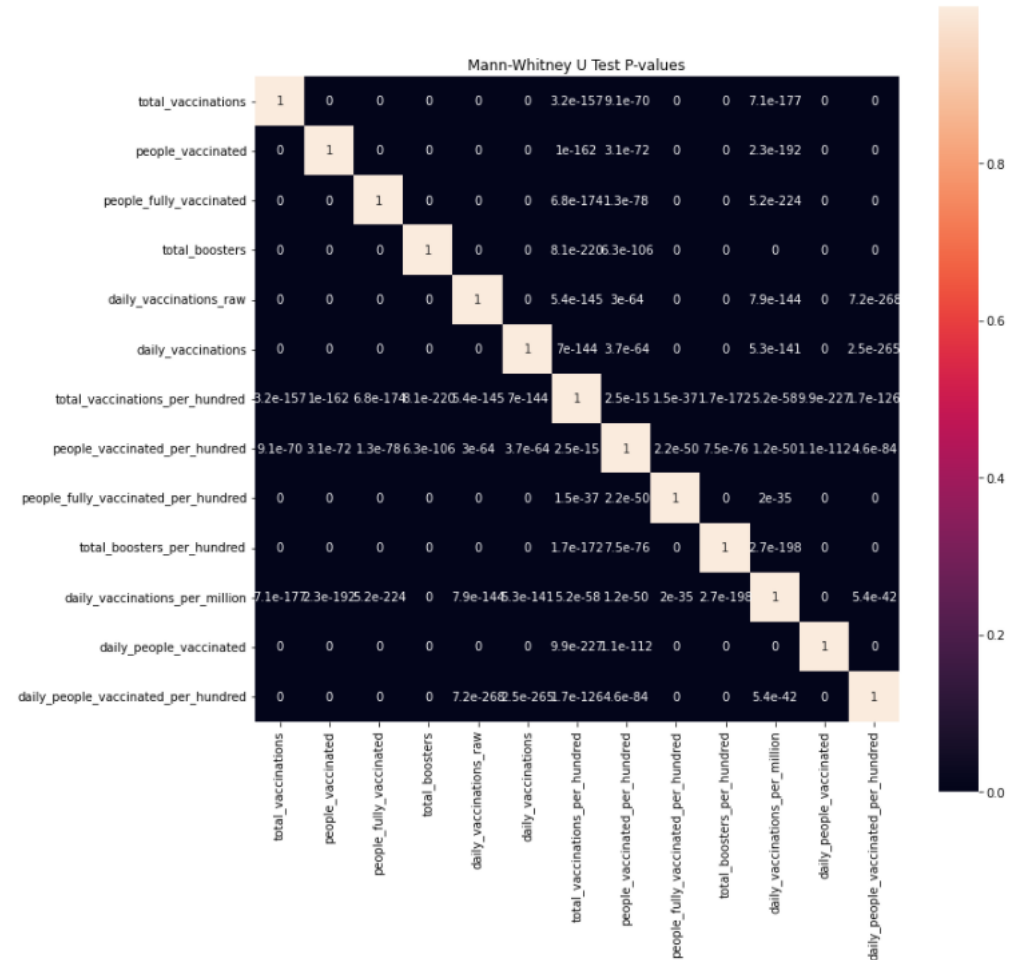Presenters: Alexander Cheng (1001634298)
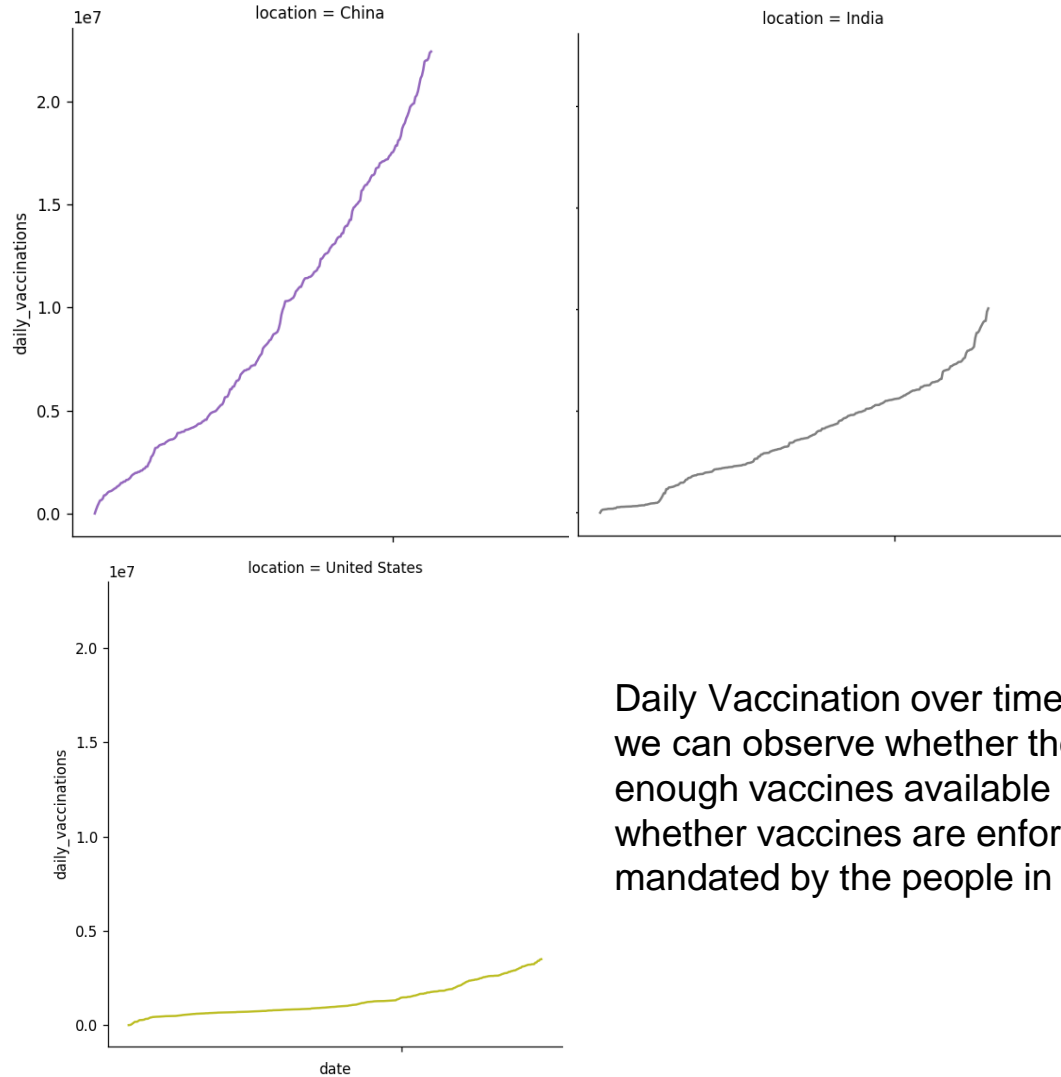Date: Dec 16, 2021

Question 1 – Data Cleaning

Correlation Heatmap between features – this helps us on filling out missing values
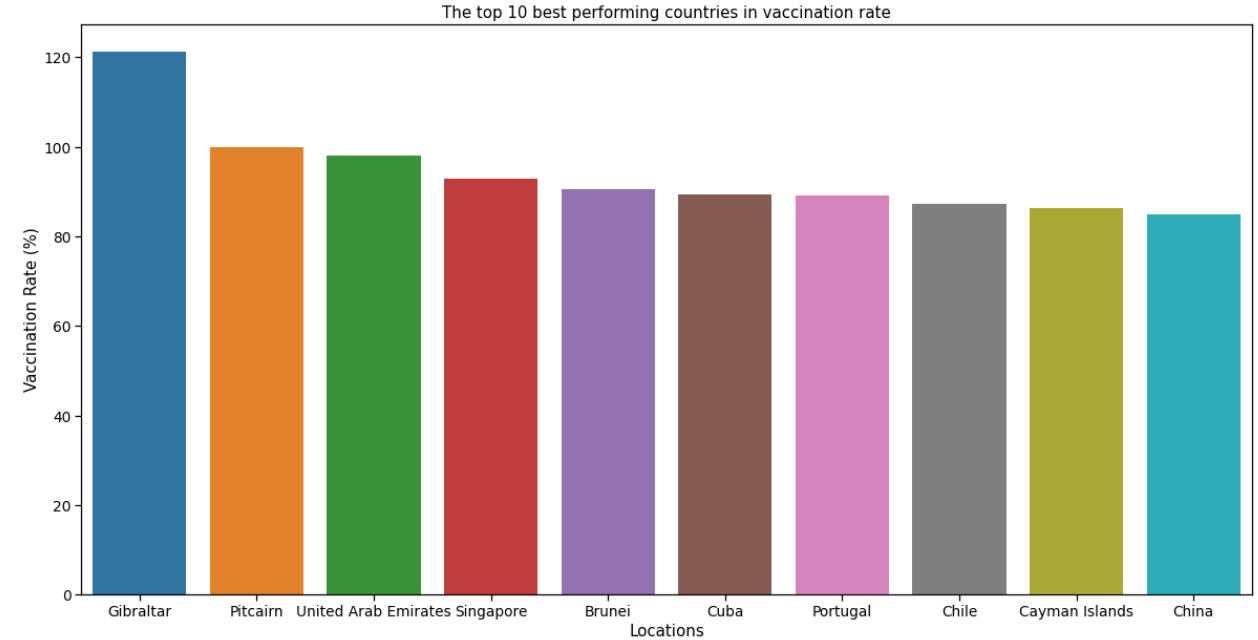

Mann-Whitney U-test P-values - compare whether there is a difference in the dependent variable for two independent groups
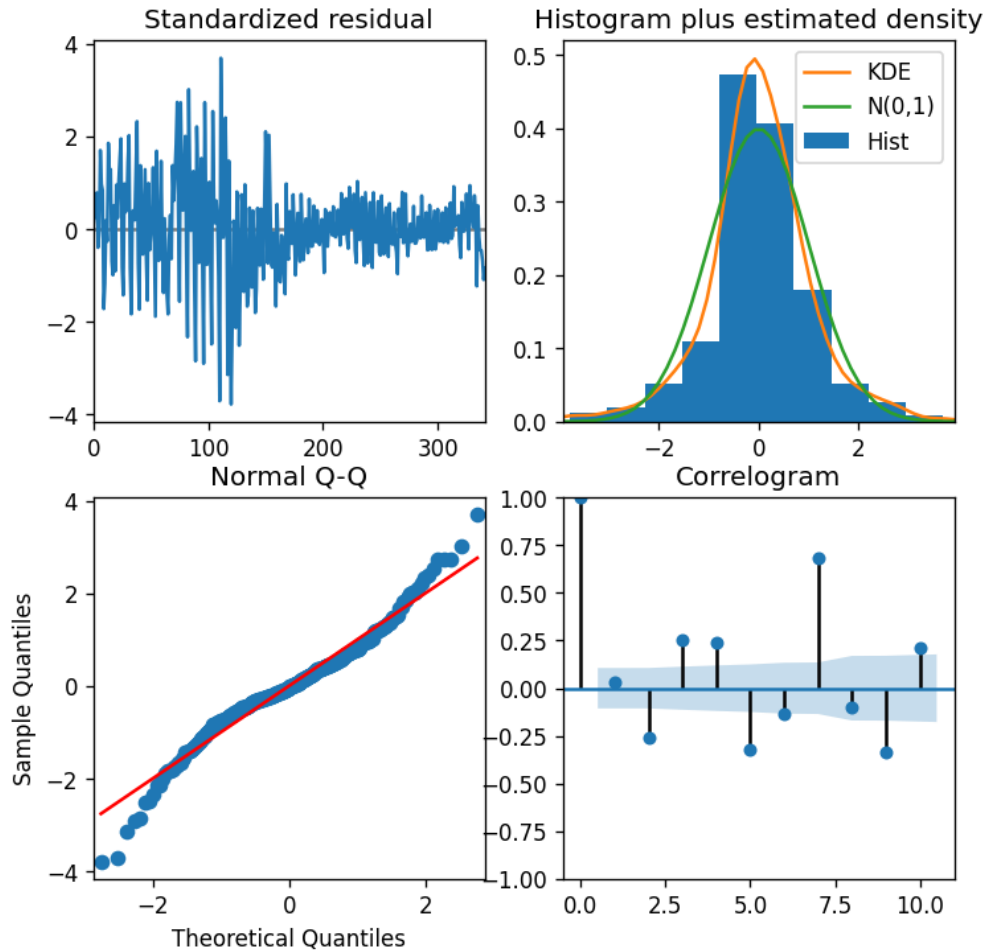
Daily Vaccination over time – here we can observe whether there's enough vaccines available or whether vaccines are enforced and mandated by the people in power
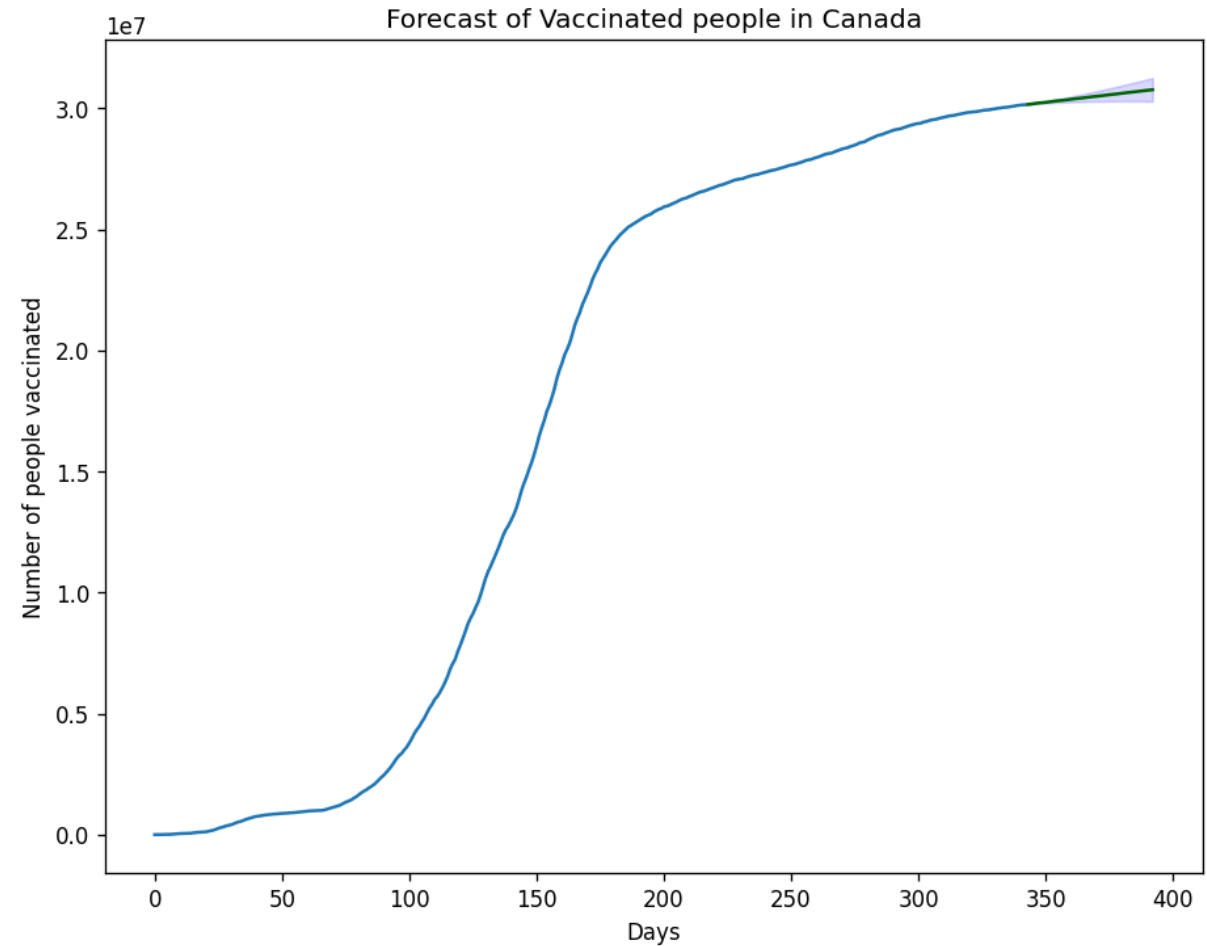
Here are the countries with the highest vaccination rate – small countries and wealthy nations such as UAE and Singapore have very high vaccination rate.
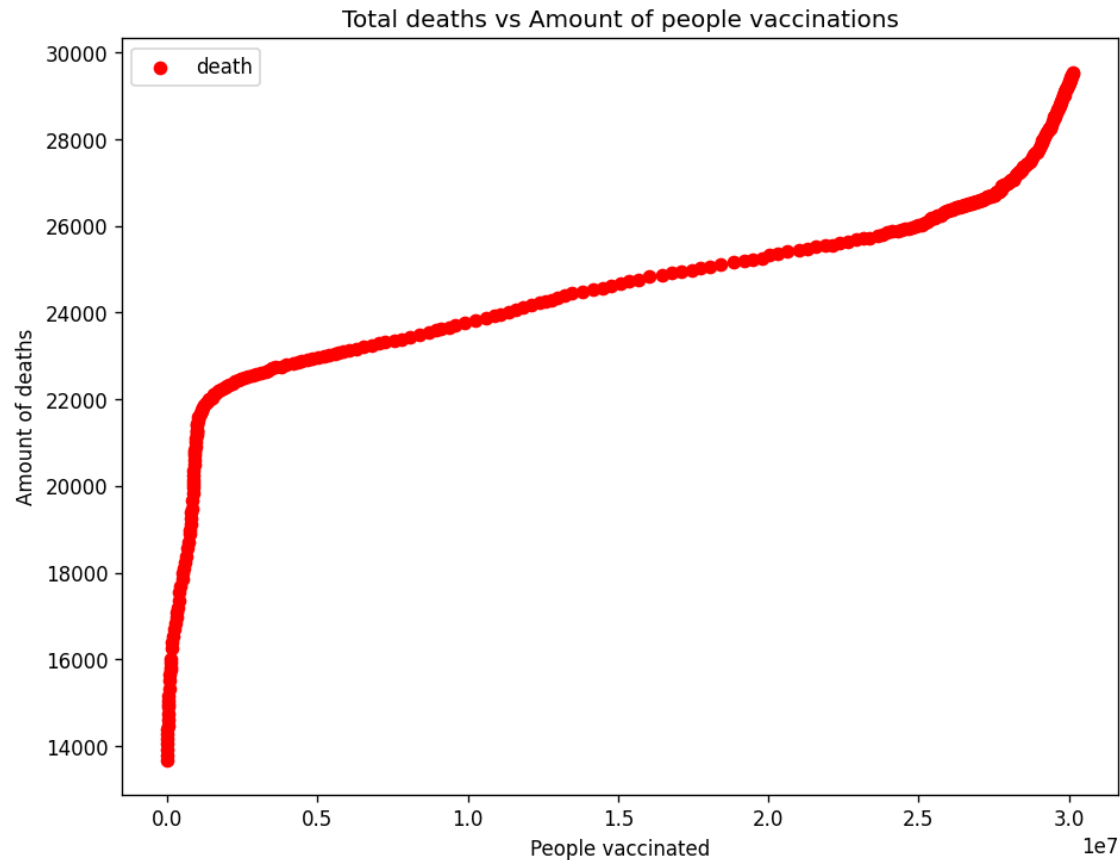
Using Auto ARIMA Model to find the best model based on Akaike's Information Criterion (AIC), we also want our time series to be stationary, we have to perform second order differencing
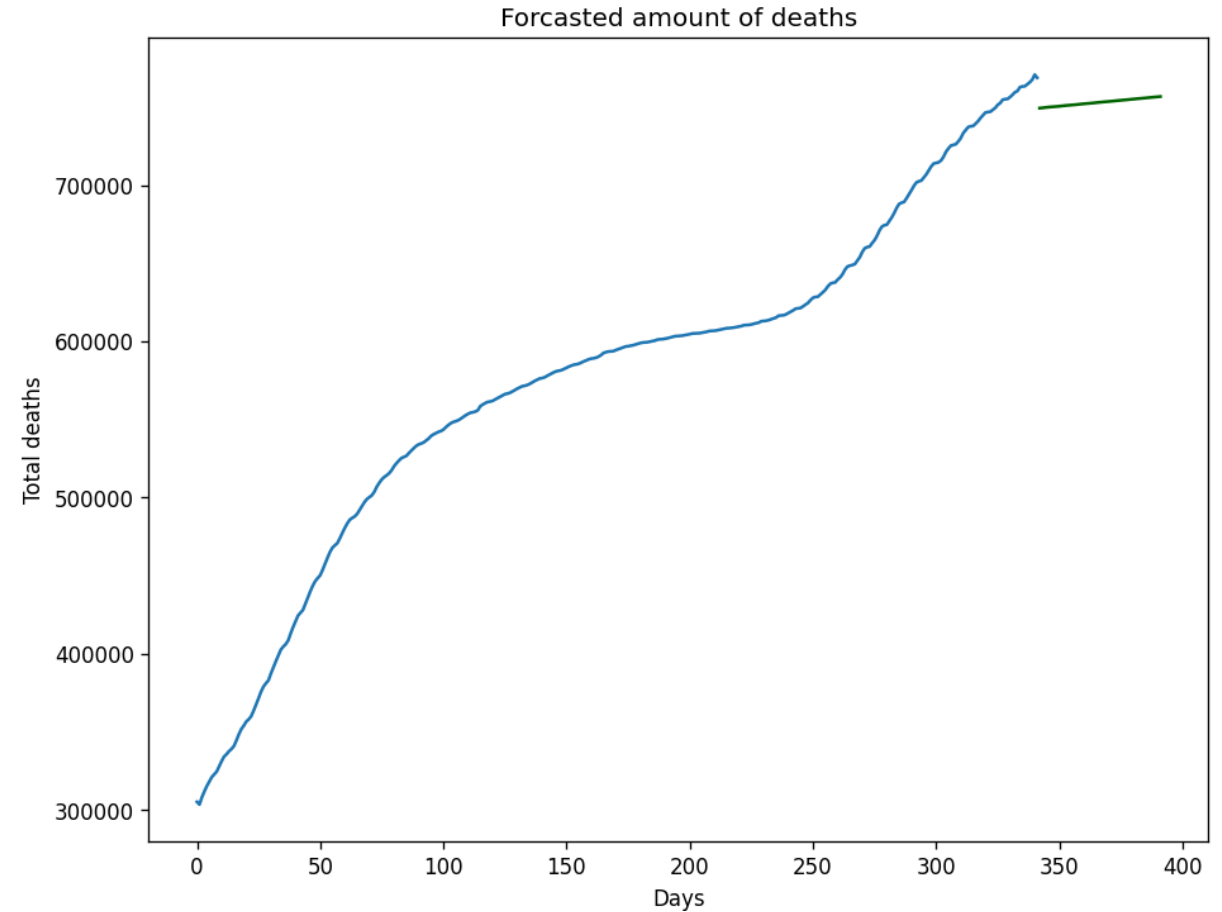
Using ARIMA Model to generate our time series for the forecasted vaccination rate and amount of vaccinated people in the next 50 days – has best, worst, and base scenarios

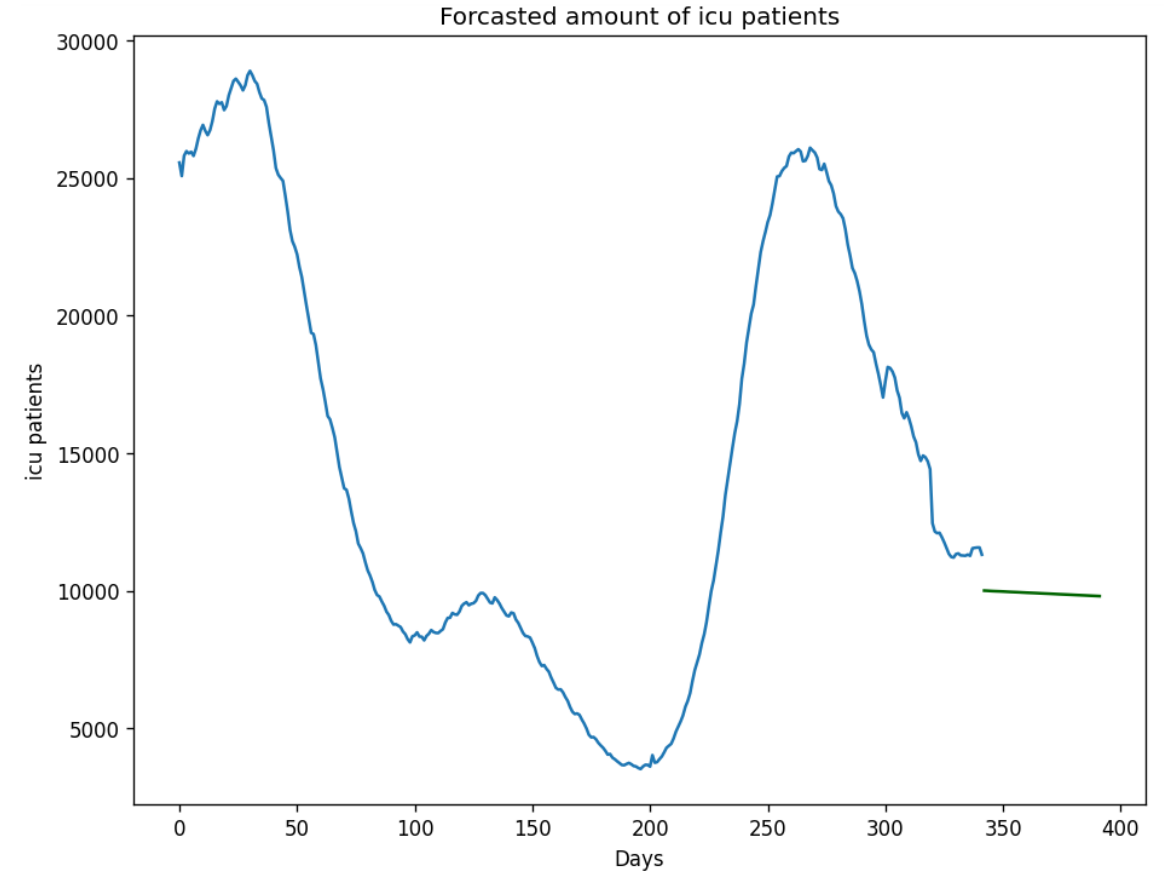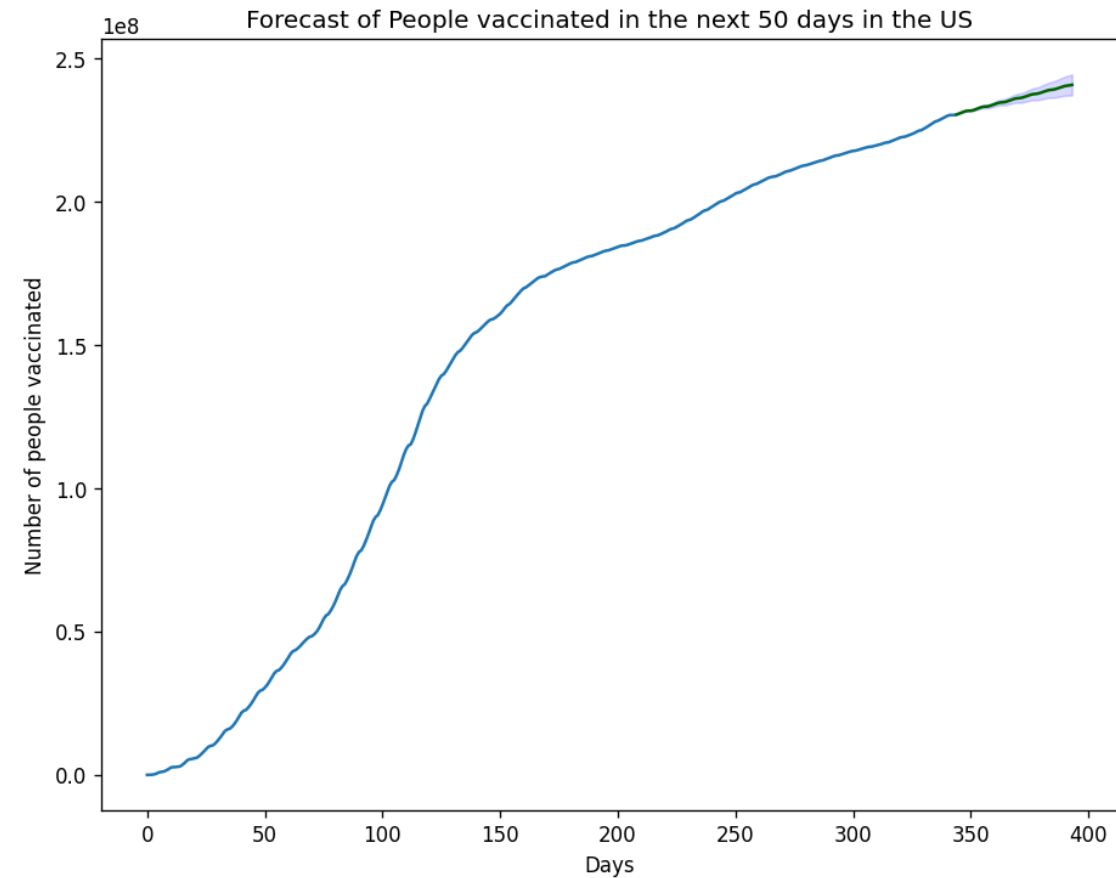# Question 4 – Second Feature Relation and Forecast – Canada only

Plotting Canada's number of deaths against number of people vaccinations – we see that it helps flatten the curve – starting to increase perhaps due to new covid variants

Using Linear regression with forecasted vaccination rates to predict the future amount of deaths due to covid in the next 50 days

Question 5 – Insights and Discussion

Vaccination Rate: Canada's 79.21% will increase to 80.79% after 50 days. US's 69.5% will increase to 71.57% after 50 days, which is double the increased rate of Canada.
Comparing the number of deaths due to COVID and icu patients, US has a population size of 8.74x greater than Canada's yet, 26.09x the deaths and 24.17x the number of icu patients.