Time Series Analysis of COVID-19 Vaccine Data

Sammy Stern, M.S. Data Science
Yash Kamat, M.S. Data Science

Introduction

Vaccines provide us with our best chance to fight the COVID-19 pandemic that has affected our world for nearly the past two years. The advent of successful vaccines gave governments around the world an opportunity to slow down rates of community spread. 12 months after the first inoculations started across North America and Europe, we aim to study the trends in vaccine roll-outs in 32 different European and North American countries, using time series methods.

Data Sources

1. Weekly new vaccine doses in EU per country/age group/vaccine manufacturer (*source*: *ECDC*)

2. Daily new vaccine doses for Canada and the United States (*source*: *Our World in Data*)

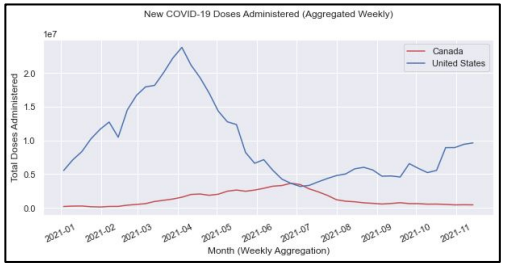

Fig 1: Weekly vaccine doses per United States & Canada

Data Collection/Preprocessing

1. **European Union Data**
   - Manually create 'Total Doses' column by adding up first doses, second doses, unknown doses and booster doses
   - Group data by 'Year-Week' (ex: '2021-W32')
   - Pivot from long to wide form by each EU country
   - Pivot from long to wide form by each manufacturer
   - Pivot from long to wide form by each age group

2. **North American Data**
   - Filter to only United States and Canada
   - Convert each date to first day of that calendar week
   - Group by week, convert to same format as EU

3. **EU and North America Aggregated**
   - Aggregate all EU and NA countries by week
   - Merge data on Year-Week column

Methods

1. **Seasonal Decomposition of Weekly Vaccine Data**
   a. EU & North America
2. **Time Series Distance Metrics**
   a. Euclidean Distance
   b. Cosine Similarity
3. **Pairwise Dynamic Time Warping**
4. **Time Series Modeling and Prediction: FBProphet**

Results/Findings

1. **Largest Euclidean Distance (EU)**
   - Countries: France & Liechtenstein: **3.290725e+07**
   - Age Groups: 0-4 & 25-49: **4.313324e+07**
   - Manufacturers: COM and SPU: **9.847537e+07**
2. **Largest Cosine Similarity (EU)**
   - Countries: France & Liechtenstein: **0.988743**
   - Age Groups: 10-14 & 15-17: **0.954548**
   - Manufacturers: MOD & UNK: **0.989050**
3. **Largest Pairwise DTW Cost (EU)**
   - Countries: France & Liechtenstein: **1.082887e+15**
   - Age Groups: 0-4 & 25-49: **1.860476e+15**
   - Manufacturers: COM & SPU: **9.697399e+15**

Time Series Predictions (FBProphet)
United States and Canada

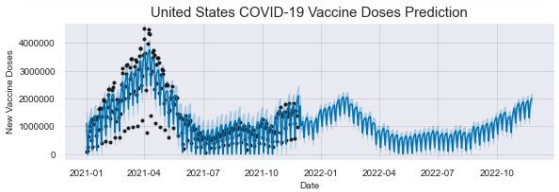

Fig 2: Vaccine predictions for United States (FBProphet)

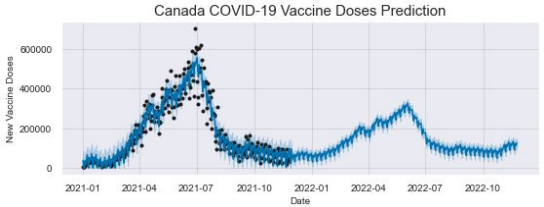

Fig 3: Vaccine predictions for Canada (FBProphet)