SARS-COV-2 Variant Analysis

Florida Polytechnic University

Ethan Bard

07 April 2022

Table of Contents

[Abstract 3](#_Toc100236712)

[Introduction 3](#_Toc100236713)

[Data 3](#_Toc100236714)

[Methods 3](#_Toc100236715)

[Results 3](#_Toc100236716)

[Conclusions 3](#_Toc100236717)

[References 4](#_Toc100236718)

# Abstract

* Analysis of worldwide covid data examining variant effect on various trends
* Major variables of interest:
  + Cases
  + Hospitalizations
  + Deaths
  + Vaccinations
  + Variant (% of sequences)
* Time Series Analysis
  + Clustering
  + Modeling
  + Forecasting
    - ARIMA
    - Multivariate Regression
* Goal: Identify countries with emerging variants and forecast cases, deaths, and variant (%)
* Goal: Forecast cases and identify countries still seeing increased cases
* Goal: Cluster trends and examine those that stand out

# Introduction

# Data

# Methods

# Results

# Conclusions

# References

<https://www.cdc.gov/coronavirus/2019-ncov/your-health/about-covid-19.html>

<https://www.lobdata.com.br/2020/09/15/how-to-perform-correlation-analysis-in-time-series-data-using-r/>

<https://machinelearningmastery.com/gentle-introduction-autocorrelation-partial-autocorrelation/>

<https://statisticsbyjim.com/time-series/autocorrelation-partial-autocorrelation/>

* Really good explanations about how to interpret the correlation plots

<https://towardsdatascience.com/setting-arima-model-parameters-in-r-grid-search-vs-auto-arima-19055aacafdf>

* R arima grid search method

<https://www.sciencedirect.com/science/article/pii/S0166093421003724>

* COVID forecasting with arima

<https://towardsdatascience.com/how-to-apply-k-means-clustering-to-time-series-data-28d04a8f7da3>

* K-means Clustering time series

<https://cran.r-project.org/web/packages/Rssa/Rssa.pdf>

* R SSA documentation

<https://www.researchgate.net/publication/228092069_Basic_Singular_Spectrum_Analysis_and_Forecasting_with_R>

* R SSA tutorial and explanation