

Covid Data Analysis

Prediction Model of stock market and covid data

Covid_p11

2016/12/12 (updated: 2022-05-22)

Background & aim

Analyse the relationship & impact of the Covid-19 on public equity markets and respective country economy.

1. Exploratory analysis of data
2. Correlation modelling & determining relationships
3. Model development and machine learning analysis
 - Focusing on key countries (respective indices)
 - China
 - India
 - Japan
 - Australia
 - United States

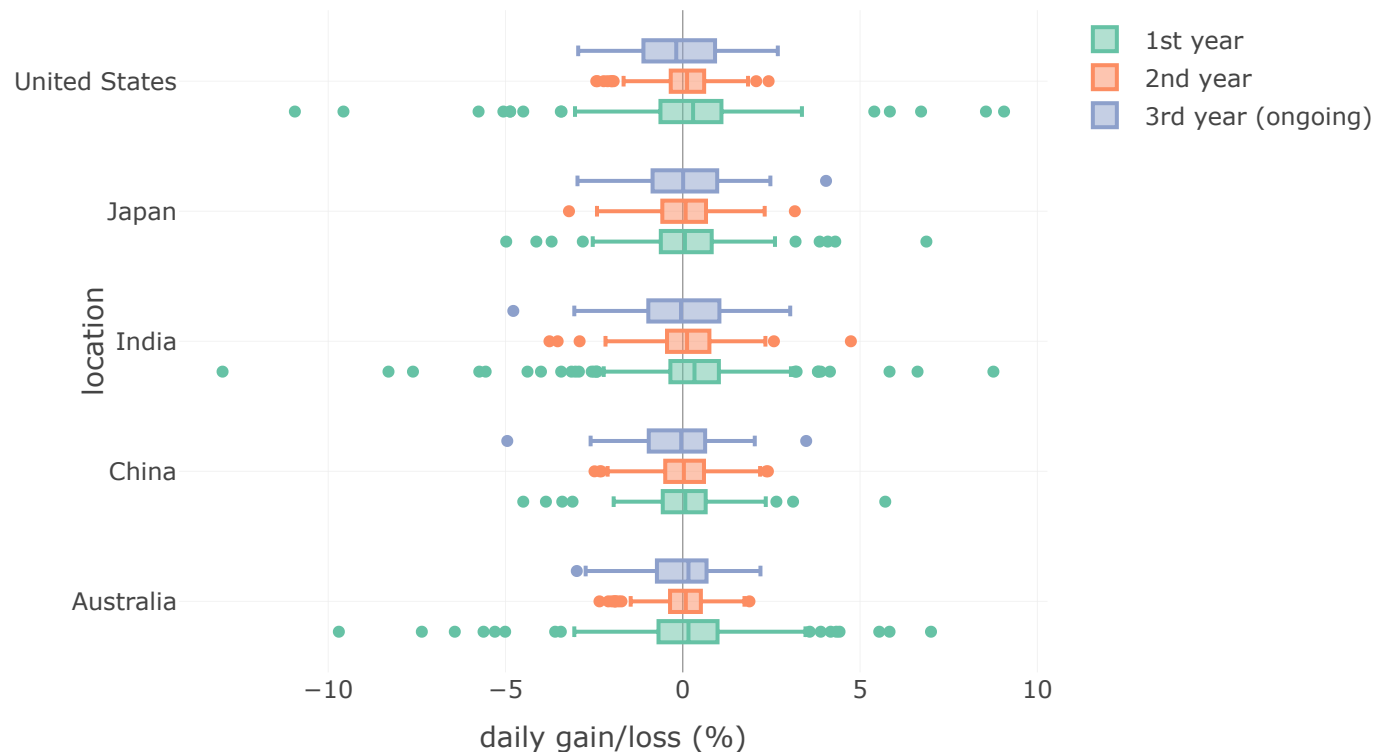
Approach

- Data collection
 - Various methods of web scrapping and csv's joined with covid data
 - Using data from respective country index
- Data analysis
 - Adopting rolling window statistics
- Models
 - Adopting other models (PLS, RDA) to predict price levels

Section 1: Exploratory Analysis

Analysing distribution of gains and losses over pandemic for respective country stock market

- **Excess kurtosis** in 1st year is consistent with the covid crisis
- **Stablisation** over 2nd and 2nd year



correlation modelling

Correlation Matrix

```
cor(df_q2_cor)
```

```
##               new_vaccinations new_tests  new_cases      Price
## new_vaccinations      1.00000000 0.4663069  0.03867642  0.5685671
## new_tests             0.46630689 1.0000000  0.60799915  0.1228084
## new_cases             0.03867642 0.6079992  1.00000000 -0.1504513
## Price                 0.56856712 0.1228084 -0.15045129  1.0000000
```

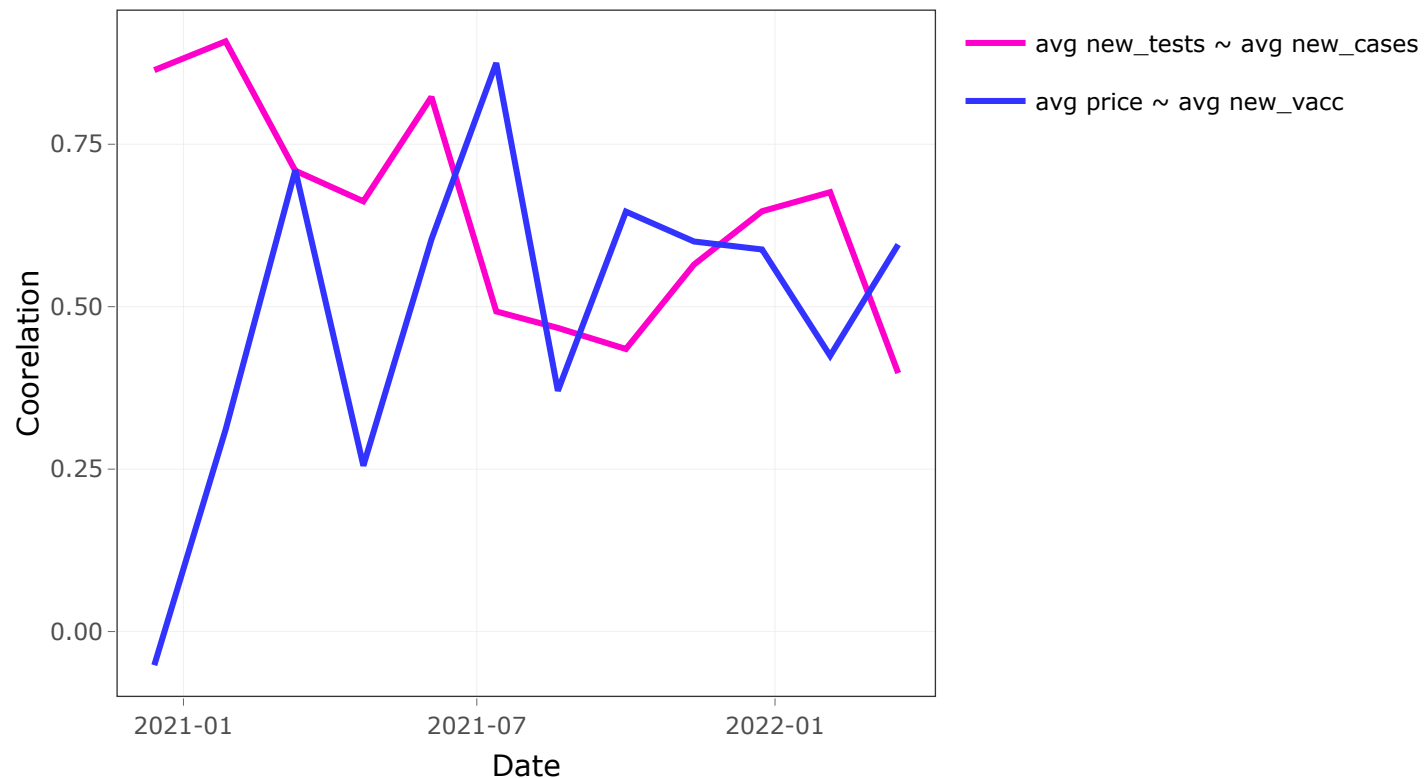
```
qtlcharts::iplotCorr(df_q2_cor)
```

```
## Set screen size to height=700 x width=1000
```


Correlation of variables over the time

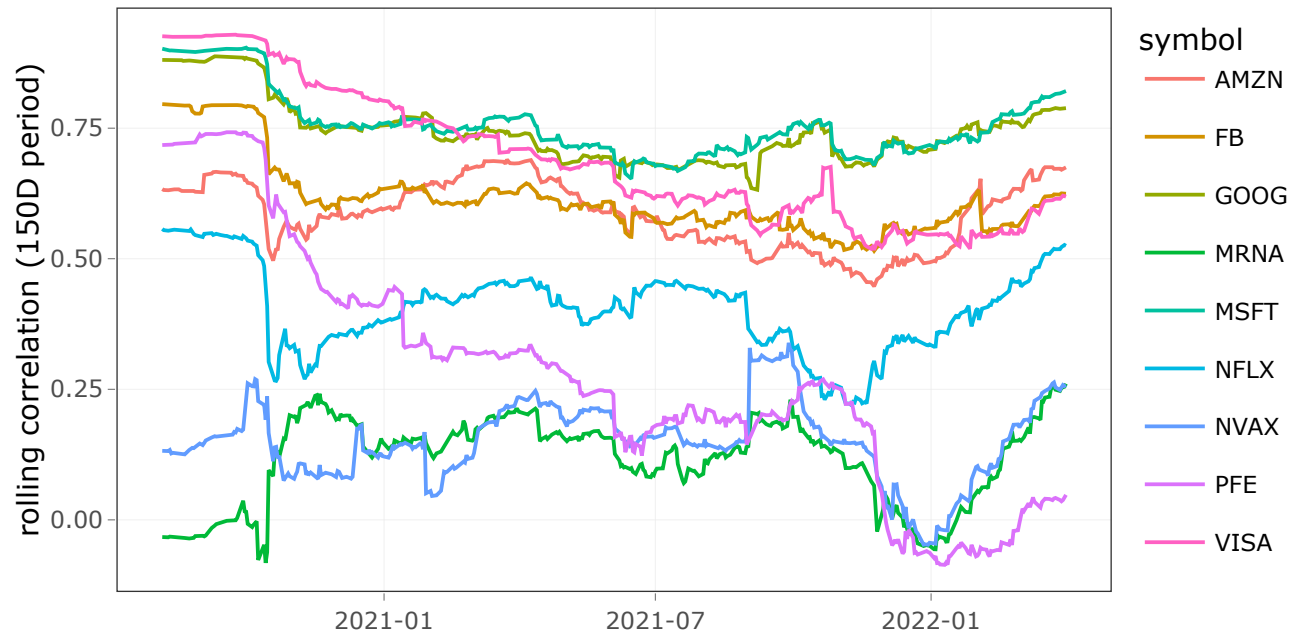
```
ggplotly(p1)
```

Price~New_vaccinations, New_tests~New_cases



Focused US (s&p500) analysis (150D rolling correlation)

- Vaccine manufacturers exhibited **anomalous** behavior
- they reached **negative** correlation
- Pfizer significant **decreasing** correlation as pandemic progressed



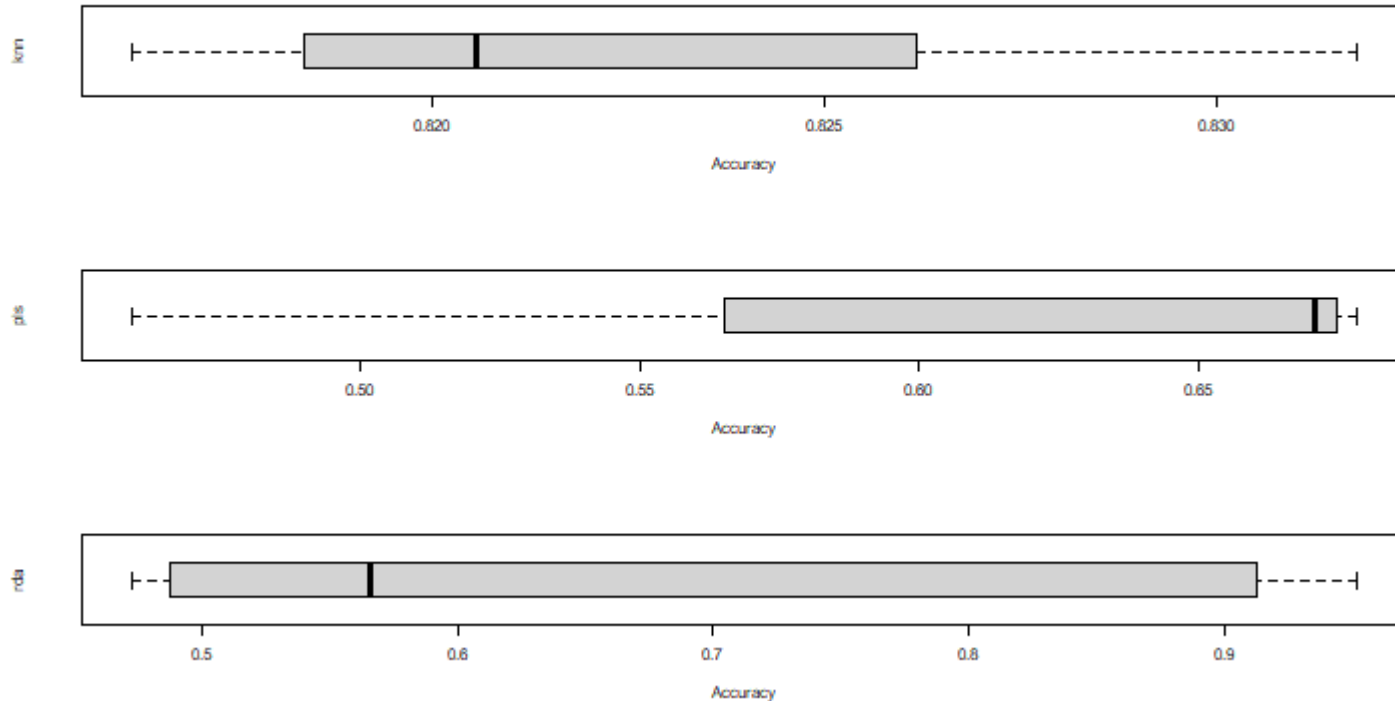
Modeling

Models developing and Machine learning

Using the historical data to classify and create models

- Classifying the Price data into 3 groups(High, Median, Low)
- building the model(KNN,PLS,RDA)
- calculate the accuracy of the model
- analysis and evaluate the model
- shiny app
- evaluation focused on accuracy comparison

Analysis and accuracy evaluation



group contribution

- Presentation: Maxim, Jasmine and Jia
- Q&A: all members
- Others:
- Section 1 code: Max
- Section 2 code: Jasmie and Christin
- Section 3 code: Paul and Marin
- Report: Jia
- Shiny app: Jinting

Thanks