IVAR - Interactive Visual Analysis with R for Exploring COVID-19 Data

Lin Yongyan
Masters of IT in Business (MITB), Singapore
Management University (SMU)
yongyan.lin.2020@mitb.smu.edu.sg

Siow Mun Chai
Masters of IT in Business (MITB), Singapore
Management University (SMU)
mcsiow.2020@mitb.smu.edu.sg

Tan Wei Li Amanda
Masters of IT in Business (MITB), Singapore
Management University (SMU)
amandatan.2020@mitb.smu.edu.sg

ABSTRACT
This is the abstract.

It consists of two paragraphs.

1. INTRODUCTION
The Coronavirus (COVID-19) has caught the world's attention with the first COVID-19 cases reported in Wuhan, Hubei, China, in December 2019. In the global battle against the virus, countries seek to understand the virus, its spread, impact and more recently, receptivity towards the COVID-19 vaccination. In today's Data Age, there are many COVID-19 related data available across various platforms. We aim to leverage the richness of the available data to provide an interactive experience in generating insights and analyses from the key COVID-19 trends: (1) new cases; (2) deaths; and (3) vaccination receptivity.

2. MOTIVATION OF THE APPLICATION
There are several one-stop applications that allow interactive visualisation of COVID-19 related data across time. These applications typically report number of events i.e. number of new cases/deaths/tests conducted, number of people vaccinated. Deeper exploration and analysis on COVID-19 trends and relationships with other factors or indicators are done in silos and majority of such studies report their findings based on pre-defined variables and specific analysis models.

With this application, we hope to combine and provide an interactive experience in in-depth exploration and analysis of the COVID-19 data. As the scope of the COVID-19 data is too large, we will focus on three key aspects for the application:

- Predictive analysis of new cases
- Bivariate and multivariate analysis of deaths and death rates with health, economic and population structure indicators
- Exploratory and bivariate analysis of vaccination receptivity with virus perception and demographics

3. REVIEW AND CRITIC ON PAST WORKS
Review of the current and past works is done separately for the three aspects mentioned earlier.

3.1 New cases
Most of the sites that forecast the near-future case load uses time series charts and/or putting values in tabular form in white papers (See Fig 1). The site/papers will give detailed explanation of the mathematical models, without letting the users to explore the dataset.

These visualisations focus on showing the result of the chosen models, without allowing users to change any of the parameters of the models. Most of the time, users are also not given the choice on which forecasting models to use. If possible, to have a few models so that users can make a comparison and to allow the users to choose the models he/she wants.

As with all forecasting/prediction models, most of the forecast models do show the Confidence Interval of the outcome. This will show the users the range in which the forecast will fall in. We will continue to maintain this in our application.

Most forecasting site do not allow users to explore the dataset. Will be good if users are able to observe the "characteristics" of the dataset before proceeding to the forecast. Also, should explore allowing the users to select the date range to be used for the forecast.

Most site only show the Time Series plot of the forecast, without the values of the result (RMSE etc). Will be good to show the users the results so that users can gauge the accuracy of the models.

Most of the visualisations for the forecast of Covid-19 confirmed cases are time series charts, without much interactiv-

ity with the users. Most of the time, only 1 model result is shown.

In this assignment, we will attempt to create an interactive visualisation. Firstly, the application will have a page to allow users to explore the time-series chart (see the trend, seasonality, anomaly etc). Subsequently, at the forecasting page, will allow users to choose the date range to be used for the forecasting. Countries are seeing spikes in the Covid-19 confirmed cases and had different measures to control the spread. As such, allowing users to choose the range of date will result in different outcome that can best suit the users' requirement. Different Models will also be shown so that users can compare the outcome.

3.2 Deaths and death rates

3.3 Vaccination receptivity

4. DESIGN FRAMEWORK

5. DEMONSTRATION

6. DISCUSSION

7. FUTURE WORK

8. CONCLUSION

References

[1] Fenner, M. 2012. One-click science marketing. *Nature Materials*. 11, 4 (Mar. 2012), 261–263.

[2] Meier, R. 2012. *Professinal Android 4 Application Development*. John Wiley & Sons, Inc.