The University of Melbourne
School of Computing and Information Systems
COMP20008 Elements of Data Processing
Assignment 2, Semester 2 2021

FINAL REPORT

Hieu Nguyen 940627, Ayesha Tabassum 1166531, Sammi Li 1271851, Telia Tung 993672
Group: 38

Since the beginning of the COVID-19 pandemic in 2019, people around the world are adapting to ensure safety and survival against the infectious disease. At the time of this report, Victoria and New South Wales combined had over 125,000 confirmed cases, which is over 95% of the total cases in Australia.[1]. Thus new problems, such as overwhelmed healthcare systems are brought to light.

The aim of this report is to forecast the data around elective surgery in Victorian and New South Wales hospitals for the year 2021-2022. Therefore, the research question is "How will COVID-19 affect the aspects of elective surgery in Victorian and New South Wales hospitals in the following year?". This research intends to help hospital workers and healthcare professionals better understand the impacts of COVID-19 on the healthcare system, as the regulations interfere with the citizens' access to urgent hospital treatment. Hospitals may also be able to determine what parts of their system are lacking, develop the resources to sustainably prepare for long-term impacts of the disease, and generate plans to increase the sustainability of the healthcare system and liveability of hospital workers and citizens in Australia.

1. Datasets and Data Wrangling

To do this, hospital datasets were collected from the Australian Government Australian Institute of Health and Welfare's (AIHW) website.[2] A total of 15 private and public hospitals from Victoria and New South Wales were chosen from Newsweek's list of best hospitals in Australia to conduct the report.[3] These datasets are in the form of excel files containing information about the number of elective surgeries, the percentage of people waiting more than 365 days, and the median waiting time for elective surgery between the years 2011-2020, which is divided by the types of elective surgeries. Note that the data is reported according to the financial year, meaning 2019-2020 would be June 2019 to June 2020. The dataset is then used to build a model to forecast the data for the next year.

The data from each hospital were preprocessed to collate data that corresponds to the desired information about elective surgeries and were parsed by merging the datasets as per their AIHW codes. Once collated, the data was cleaned to remove entries that have no value and alter entries that are not measurable to a quantifiable value. The Royal Melbourne Hospital was chosen to

undergo linear regression to forecast a predicted value and compare it to the actual value in 2019-2020 as it was ranked number 1 as the best hospital in Australia in 2019.[3] Therefore, may be a good representation of the impacts of COVID-19 on hospitals in Melbourne.

From the predicted values, a scatterplot was generated, showing the true values from 2011 to 2020 as blue dots and the predicted value as a red dot. This decision was made over other forms of representations of data as a scatter plot would best showcase the data differences and possible linear relationships. Using the data from all 15 hospitals, a linear regression was used to predict the median waiting time and number of elective surgeries for the following year with the predictor being the previous year's value. This data was visualized using a table as there are many numerical values that would be easier to read using a table. Links for the datasets collected from the Australian Institute of Health and Welfare's (AIHW) website are below:

Hospital	**Link to dataset**
Royal Melbourne Hospital [City Campus]	https://myhospitalsapi.aihw.gov.au/api/v1/reporting-units-downloads/H0371
Royal Hospital for Women:	https://myhospitalsapi.aihw.gov.au/api/v1/reporting-units-downloads/H0040
Campbelltown Hospital	https://myhospitalsapi.aihw.gov.au/api/v1/reporting-units-downloads/H0029
Royal Children's Hospital [Parkville]	https://myhospitalsapi.aihw.gov.au/api/v1/reporting-units-downloads/H0364
Austin Hospital [Heidelberg]	https://myhospitalsapi.aihw.gov.au/api/v1/reporting-units-downloads/H0359
The Alfred:	https://myhospitalsapi.aihw.gov.au/api/v1/reporting-units-downloads/H0358
Royal Women's Hospital [Parkville]:	https://myhospitalsapi.aihw.gov.au/api/v1/reporting-units-downloads/H0366
St Vincent's Hospital [Melbourne]:	https://myhospitalsapi.aihw.gov.au/api/v1/reporting-units-downloads/H0374
Western Hospital [Footscray]	https://myhospitalsapi.aihw.gov.au/api/v1/reporting-units-downloads/H0363
St George Hospital	https://myhospitalsapi.aihw.gov.au/api/v1/reporting-units-downloads/H0038
Maitland Hospital	https://myhospitalsapi.aihw.gov.au/api/v1/reporting-units-downloads/H0102
St Vincent's Hospital [Darlinghurst]	https://myhospitalsapi.aihw.gov.au/api/v1/reporting-units-downloads/H0017
Royal Prince Alfred Hospital	https://myhospitalsapi.aihw.gov.au/api/v1/reporting-units-downloads/H0021
Royal North Shore Hospital	https://myhospitalsapi.aihw.gov.au/api/v1/reporting-units-downloads/H0068
Nepean Hospital	https://myhospitalsapi.aihw.gov.au/api/v1/reporting-units-downloads/H0056

2. Data Analysis and Results

Using Royal Melbourne Hospital as a basis to forecast the value for the following year, a linear autoregressive model with one lag was used to model the impact of the first three months of COVID-19. The value under normal conditions before COVID-19 was forecasted for 2019-2020 and compared to the actual value.

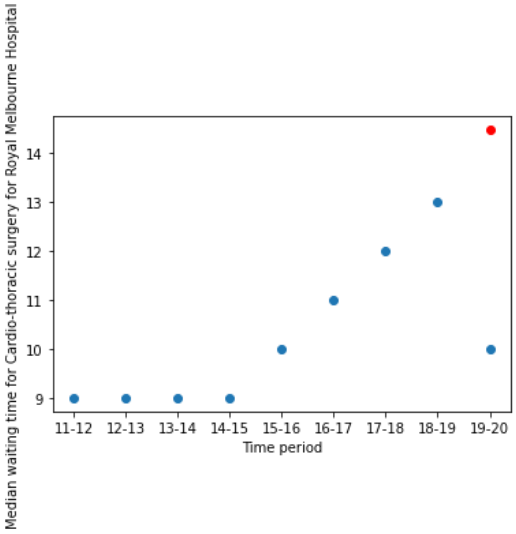

Figure 1: Median waiting time for Cardio-thoracic surgery for Royal Melbourne Hospital (days) vs Time period (2011 - 2020)

Figure 1 demonstrates that from 2014 up until 2019, the median surgery waiting time was increasing by 1 day after a continuous stable value of 9 days before 2014. However, in 2019-2020, the value had dropped from 13 days in 2018-2019 to 10 days. The predicted value, as indicated by the red dot, was determined to be 14.48 days. This is evaluated to be a 4.48 difference between the forecasted value and actual value.

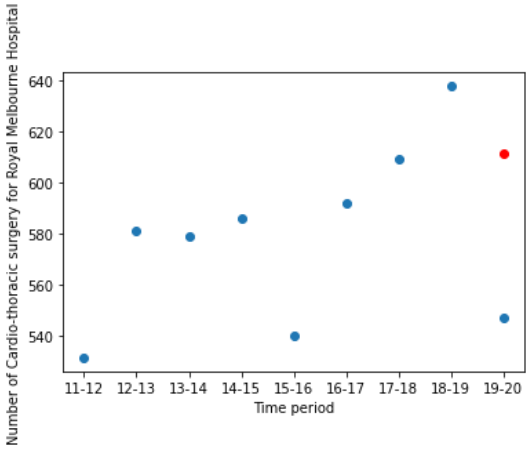

Figure 2: Number of cardio-thoracic surgeries at the Royal Melbourne Hospital (days) vs Time period (2011-2020)

Figure 2 shows the total number of cardiothoracic surgeries done at the Royal Melbourne Hospital from 2011 to 2020. It can be seen that there is a noticeable rise from 2011-12 to 2012-13. It was approximately stable around 580 numbers until 2014-15. However, the number of surgeries had a significant fall between 2014-15 and 2015-16. For the following three years period (2016-2019), the number kept increasing but had a huge decline in 2019-20. The forecast for 2019-2020 was around 611 which is 64 more than the actual number of surgeries that was 547.

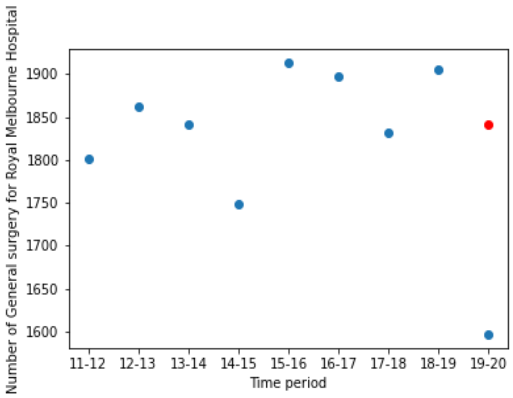

Figure 3: Number of general surgeries at the Royal Melbourne Hospital (days) vs Time period (2011-2020)

The number of general surgeries as shown in Figure 3 is predicted to be at around 1840 (shown by the red dot) as the number of surgeries performed fluctuates year to year but in a slight upward trend. The number has drastically decreased in the 2019-2020 time period, where only 1597 general surgeries have been performed. The drastic difference of 243 surgeries is easily noticeable, even when compared with the second-lowest amount at 1748 in the 2014-2015 time period.

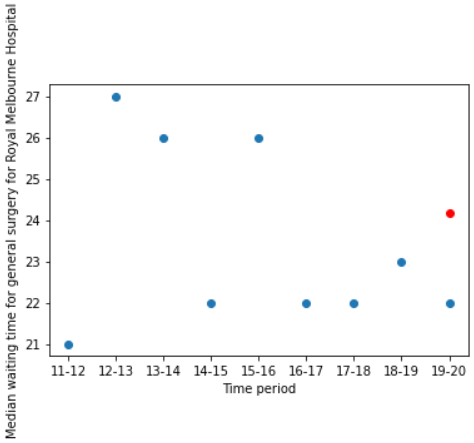

Figure 4: Median waiting time for general surgery for Royal Melbourne Hospital (days) vs Time period (2011 - 2020)

The median waiting time is predicted to be at 24.17 days as indicated by the red dot in Figure 4. The median waiting time fluctuates between 21 to 27 days throughout the decade, with a slight downward trend starting from 2012-2013. With that in mind, the median waiting time is around 22 days for a majority of different time periods, even though it is not continuously consistent. The actual median waiting time for the 2019-2020 period is lower than the forecasted waiting time by 2 days, where patients waited for a median of 22 days before their general surgery, which is similar to the median waiting time of other time periods.

Now, instead of a simple autoregressive model, the forecast table below (Figure 5) used similar data from all the hospitals in the datasets. For example, take St's George (NSW) 2011-2020 data on a number of general elective surgeries, the data would be split into halves, predictor (2011, 2012, 2013, 2014) on the response (2015) as training data, and predictor (2015, 2016, 2017, 2018) on the response (2019) as another training data. The same process is repeated on the other 14 hospitals gathering their data on general elective surgeries, to build a model to forecast RMH numbers of general surgery. Resulting in a more realistic model, as the first half can be considered the normal condition while the second half factors the initial impact of COVID-19.

Type of elective surgery	Number of elective surgeries	Median waiting time (days)
Cardio-thoracic surgery	489.190289	9.774964
Ear, nose and throat surgery	152.190509	43.332234
General surgery	1524.732022	20.308678
Neurosurgery	715.290678	42.151712
Ophthalmology	386.873005	113.103654
Orthopaedic surgery	1029.763940	44.068200
Plastic surgery	788.152901	23.457514
Urology	2010.754093	17.887952
Vascular surgery	195.351212	27.115538

Figure 5: Table of forecasted values for 2020-2021. Type of elective surgery vs Number of elective surgeries vs Median waiting time (days)

Figure 5 demonstrates the forecasted number of elective surgeries and median waiting time in days for the different types of elective surgery conducted in hospitals. Based on observation, it can be interpreted that clinics with a higher number of surgeries have less waiting time, as shown by general surgery and urology, with around 1,030 surgeries and 20 waiting days, and around 2,011 surgeries with 18 days of waiting respectively. Whilst the clinics with fewer surgeries a year have more waiting days than surgeries with a high number of surgeries, as shown by ophthalmology, which has around 387 surgeries and a waiting time of 113 days.

3. Significance of Results

Since the beginning of COVID-19 (2019-2020), it is shown that the forecasted median waiting times and the number of surgeries for the Royal Melbourne Hospital were higher than the actual value. This suggests that the effects of COVID-19 are prominent and should be taken into consideration for the long-term planning of hospital resources, indicating that demands for healthcare services may be impacted.

The forecasted number and median waiting time for different types of elective surgeries for 2020-2021 were also determined based on the hospitals' datasets. These forecasted values are useful for establishing what types and amount of resources the hospital may need to prepare for the following year. It is also brought to light that there are many elective surgeries with high median waiting times. Hospitals may benefit from this information by improving protocols and allocating staff and resources to the corresponding surgery types to lower waiting times and accommodate more patients.

4. Limitations

The research focused on the Royal Melbourne Hospital, hence the results might not be generalisable for other hospitals. This can be improved on by altering the approach to the research question by analysing more data from more hospitals and producing a more generalisable report for all hospitals. The prediction may be more accurate if monthly datasets were used; however, the datasets were limited to annual figures, possibly due to privacy concerns from within the hospitals, reducing accessibility to more accurate data. By obtaining data from more hospitals and data in monthly intervals, the precision and reliability of the results can be increased, and be more indicative of the activity of Victorian and New South Wales' hospitals.

Additionally, the evaluation of forecasted data on elective surgery overall used data from top hospitals of Victoria and New South Wales, thus the results are unlikely to be generalisable to private, specialised or rural hospitals. Further research on such hospitals is needed for a more accurate result that represents other hospitals. Thus, reliability and generalisability can be improved by assessing datasets separated by metropolis, rural and specialised hospitals, then analysing the data based on the type of hospital.

In addition to further research on rural hospitals in Victoria and New South Wales, it is also valuable to conduct research on the entirety of Australia and other cities that have been heavily affected by COVID-19 on a more global scale. This can provide further comparisons to gain insight on hospital systems that are most effectively preparing and approaching the issues raised due to COVID-19, and thus develop a better understanding of efficient long-term planning.

References:

1. Australian Government Department of Health. (2021, October 14). Coronavirus (COVID-19) case numbers and statistics. Australian Government Department of Health. Retrieved October 14, 2021, from https://www.health.gov.au/news/health-alerts/novel-coronavirus-2019-ncov-health-alert/coronavirus-covid-19-case-numbers-and-statistics
2. Hospital downloads. Australian Institute of Health and Welfare. (n.d.). Retrieved October 14, 2021, from https://www.aihw.gov.au/reports-data/myhospitals/content/data-downloads/hospital-downloads?page=1
3. Posimsetty, G. (2019, June 20). Best Hospitals - Australia. Newsweek. Retrieved October 14, 2021, from https://www.newsweek.com/best-hospitals-2019/australia