Forecasting Airbnb Pricing During COVID-19 Pandemic

Final Project Paper

Authors: Evelyn (Wanyi) Dong, Jenny Shang, Qiongqiong Lin, Yoki (Lingrou) Wang
DSO 522: Applied Time Series Analysis for Forecasting

Table of Contents:

1. Executive Summary

This report assesses the impact of COVID-19 on the tourism industry by summarizing the analysis and findings of forecasted average listing price for Airbnb properties in Oakland, California during the period from January 13, 2020 to July 16, 2020; and extends the forecast to predict average listing price between July 17, 2020 to Aug. 15, 2020. Various time series forecasting methods are used including the seasonal naive model, the Holt-Winters exponential smoothing model, the autoregressive integrated moving average model (ARIMA), neural network, as well as the linear regression model and the decision tree model with inclusion of external variables. The forecasts are used to compare the estimated average listing price under the assumption of no occurrence of COVID-19 versus the actual listing price to assess the pandemic's impact on Airbnb listing price. Results show that using the data up until January 13, 2020, the seasonal ARIMA (SARIMA) model achieves the best result with minimal error. Finally, using the most optimal SARIMA, the predicted average listing price for the 30 days from July 17, 2020 to Aug. 15, 2020 is between $122 and $125.

2. Introduction and Motivation

On December 31, 2019, the Wuhan Municipal Health Commision first reported a cluster of viral pneumonia cases of unknown cause. Nine days later, the World Health Organization (WHO) announced that authorities in China have determined that the pneumonia cases are caused by a novel coronavirus, thus commencing a series of teleconferences and genetic research among the world's clinic network. The surprise discovery of the novel coronavirus, named COVID-19 by the WHO, caught the global society off-guard, sunk the world into panic and inflicted a worldwide economic recession. Over the next six months, cases of COVID-19 escalated at unprecedented speed, reaching 1 million in April 2020 and over 10 million by the beginning of July; leading WHO to declare COVID-19 a global pandemic. From the explosion of cases onboard the Princess Diamond cruise ship, to widespread lockdowns across Europe, COVID-19 led governments to impose severe travel restrictions between countries. Evidently, the pandemic has devastated the global tourism industry.

Since its foundation in 2008, Airbnb, Inc has experienced strong growth in its online accommodations platform, Airbnb.com. The platform lists available housing for different cities and allows individuals to connect directly with the hosts. Over the years, Airbnb has become increasingly popular as a new form of travel accommodation. It offers travellers the flexibility in choosing from a variety of accommodation types, such as renting the whole apartment for larger groups at an affordable price. At the same time, it provides city locals with a platform to make additional income on spare housing space. For many of its hosts, most of the income is generated from travel seasons, typically during the summer as children have a break from school. However, this year is different. With travel restrictions in place, Airbnb hosts are struggling from booking declines due to decrease in domestic and international travellers. Now more than ever, Airbnb and its hosts need to estimate the impact of COVID-19 to more efficiently plan its resources and expenditures. Furthermore, a reliable pricing forecast can help various stakeholders in making plans for the future.

With such motivation in mind, the team seeks to use time-series forecasting to assess the impact of COVID-19 on Airbnb average listings price and generate a 30-day average price forecast first for

Oakland, California, with the additional possibility of extending it to other cities. Our goal is to develop valuable insights and recommendations that can help Airbnb and its stakeholders in resource planning, and in generating a solid pricing strategy to achieve customer retention and sustainable growth. The steps we take to achieve our goal is summarized as follows:

1. Understand the historical patterns of the Airbnb average listing price by studying potential trends and periodic components.
2. Explore a variety of time-series models and determine the most appropriate candidate model on the basis of accuracy and simplicity to analyze COVID-19 impact and forecast future prices.

Our analysis also explores external factors that play a role in determining the average listing price such that we may advise Airbnb stakeholders in non-Black-Swan scenarios. We are confident that with our models and insights, our project will be a great success.

3. Data: Source and Frequency

3.1 Source

Data was obtained from http://insideairbnb.com/get-the-data.html, and the data frequency is daily from 7/16/2018 to 7/16/2020 (722 data points total). According to the website, data was collected from publicly available data from the Airbnb website. Data were cleaned from raw data files (calendar.csv.gz), by aggregating all listings and computing the average listing price daily. Then the output cleaned_data.csv which contained data, and the average listing price of all listings available in the Oakland, CA area is used as the data input to build models.

3.2 External Variables

To improve prediction accuracy, monthly data that is publicly available are examined such as GDP (nationwide), unemployment rate (nationwide), and housing data in the Oakland area are used in modeling to see if there's a significant lift to the model.

3.3 Frequency

The given frequency was retained without any accumulation, as daily price prediction was identified to be the most beneficial interval for target users. Another consideration was a constraint of the length of data, which was only 2 years, and any accumulation would lose a granular level of information.

4. Methodology: Description of Procedure

4.1 Data Staging

The year of 2020 has been a disruptive one because of the outbreak of coronavirus, COVID-19. A test was performed to find the dropping point of the data in 2020, and the date was identified as 1/13/2020 by using the 'changepoint' function in R, which takes both trend and variability into consideration. It's believed that after this point, the data pattern showed a significant difference compared with historical data. To examine the impact of the pandemic, two stagings has been adopted:

1. Only using Pre-COVID Data
    a. Data is identified to be data from 8/16/2018 to 1/13/2020
    b. A 60 days validation window is reserved to slide and test model performance
    c. Forecast the next 30-day average listing prices as final deliverables
2. Using all data available
    a. Data is identified to be all data available
    b. A 60 days validation window is reserved to slide and test model performance
    c. Forecast the next 30-day average listing prices as final deliverables

4.2 Predictive Modeling

A total of 4 methods are developed as candidate models, where the best model is identified to have low training and testing errors (defined as RMSE and MAPE) and is robust enough which means the performance on training and testing data should be similar. A naive forecast has been set as the baseline model because of its simplicity and easy to explain, and thus the model that has the balance between outperforming a naive forecasting model and has good robustness will be finalized as the model for practical use and potential deployment to the audience.

4.3 Cross-Validation of Prediction Errors

A sliding window method was adopted, which means a forecast of upcoming 30 days will be made at a time, and one error datapoint will be calculated based on differences among the 30 forecasts with the actual data available in the testing set. Finally, an average method was applied to all 30 error points to get a cross-validated error estimate. In this way, each day will have multiple forecasts and errors are calculated accordingly, which enhances the credibility and reliability of the models.

4.4 Deployment of the Predictive Models

By comparing the two staging methods, predictions without pre-COVID data showed very different patterns compared with staging including COVID data. As expected, models with COVID-19 data in 2020 helped to reduce forecasting errors significantly, which suggested the occurrence of the emergency pandemic does not align with previous historic patterns and thus needs to be modeled and monitored carefully, which will be very valuable information to the target audience.

5. Presentation of Forecasts

5.1 Baseline Naive Forecasting

(1) Forecast 1 using pre-COVID data
A simple method was to adopt naive forecasting, which uses previous actual data, in this case, actual Airbnb prices in the past as the prediction of future prices. The rationale behind this method is that the future data will have a similar pattern as historical data, and it will be very easy to explain and interpret to users of the model. However, compared with more complicated methods as described above, naive forecasting is treated as a baseline and tradeoffs will be made with the best balance among model complexity and accuracy.

Similar to the methodology adopted previously, two models were built with different sets of training data, which function 1 uses all data available before COVID time and function 2 includes some data during COVID, and their prediction is visualized (Figure 1).

	RMSE	MAPE
Training Set	2.69	1.80%
Validation Set (Up to Recent Data)	4.40	3.03%

Based on RMSE and MAPE results, the naive forecasting failed to capture the COVID-19 impact and has a much larger error on the predictions compared to training errors. This is expected as naive forecasting focuses on historical data patterns, and the outbreak of a pandemic has never happened in the past, and thus the increased forecasting error aligns with expectations. The naive method will be used as a baseline to compare with more advanced models.

(2) Forecast 2 using all data
The second function included all data available and left 60 days as testing period. The training data is from May 7th and within the training data, dates after April 8th are used to evaluate the training error of the model. The testing data starts from May 8th to July 16th, with dates after June 6th are used for calculating the testing error. The evaluation length of both training and testing are consistent at 30 days.

The cross-validation of model errors are not adopted in naive forecasting given the nature of the model, and errors are calculated as the difference between actual data and the forecasted data, which is the last 7 data points in the data available. The choice of using the last 7 data points repeatedly is related to travel seasonality on a weekly basis instead of using a one datapoint forecast universally for all future data (Figure 2).

	RMSE	MAPE
Training Set	3.23	2.26%
Validation Set (Up to Recent Data)	2.77	1.77%

The result for function 2 outperforms function 1, as expected because of additional data in the pandemic. This model sets the baseline for future model developments.

5.2 Exponential Smoothing

The rationality behind choosing the Holt-Winters exponential smoothing model as a possible forecast method is that the Airbnb price data presents an overall gradual upward trend with a weekly seasonality.

(1) Forecast 1 using pre-COVID data

The Holt-Winters exponential smoothing model requires three parameter inputs: error, trend, and seasonality. For each of these parameters, there are two possible values to choose from, which are additive or multiplicative. Additive refers to a constant magnitude while multiplicative refers to an increasing or decreasing magnitude. For example, an additive seasonality generally means that the seasonal fluctuation is approximately the same magnitude for every period, whereas a multiplicative seasonality could be one where the difference between peak and trough in a season is growing overtime. In determining the possible parameter combinations of error, trend, and seasonality of a Holt-Winters exponential smoothing model, we looked at the plot of Airbnb prices over time prior to Jan. 13, 2020 which we have identified as the COVID-19 turning point.

The graph (Figure 3) shows the Airbnb price before and after the COVID-19 turning point, which is separated by the red dotted line. To the left of the line, there is a difference in the magnitude of seasonal fluctuations prior to January, 2019 compared to the fluctuations after this month. In addition, aside from the drastic drop in price in mid-2018 the average pricing looks to be trending steadily upwards.

After conducting a thorough analysis of various parameter combinations, the most optimal one with the multiplicative error, a damped-additive trend, and additive seasonality (MAdA). Using this parameter combination, we conducted 30-days rolling forward predictions setting 60 days prior to the COVID-19 turning point as our validation set all data prior to the validation set as training data. This model generates the smallest validation MAPE of 1.52%, with a training MAPE of 0.52%.

The graphical depiction of our forecast for the 30-day period post-COVID-19 turning point is shown in Figure 4.

	RMSE	MAPE
Training Set	1.25	0.52%
Validation Set (30 Iterations)	1.89	1.52%

Although this model gives relatively smaller errors, the downside is that it does not fully satisfy the statistical assumptions of the residuals. As shown in Figure 5figure, the residuals show non-stationarity as their variance greatly decreases in 2019. Recall that this pattern mirrors the variation seen in the Airbnb prices over time. Consequently, the distribution of the residuals does not follow a standard normal distribution as a great number of residuals cluster around 0. A potential solution to resolve this may be to separate the dataset further into two sections and use only the 2019 data to make predictions since the 2018 seasonal fluctuation is no longer relevant. Alternatively, we can consider other transformations.

In addition, the residuals also appear to exhibit autocorrelation at an eight days lag. As analyzed via an ACF and a PACF plot (Figure 6), the residuals have significant correlation around lag 5, 13, 21, etc. Given the potential autocorrelation, we further explore the SARIMA model, which is covered in the next section of the report. Other parameter combinations we have explored include MAM, AAA, MMM (Figure 9).

Parameters	Training RMSE	Training MAPE	Validation RMSE	Validation MAPE
MAM	1.25	0.52%	1.91	1.53%
AAA	1.25	0.52%	1.90	1.53%
MMM	1.25	0.52%	1.92	1.54%

As shown in the table, all parameter combinations produced very similar results but a Holt-Winters exponential smoothing model using multiplicative error, damped-additive trend and additive seasonality performed marginally better.

(2) Forecast 2 using all data

For this approach, instead of using the data prior to the COVID-19 turning point to forecast what Airbnb price would be under the scenario of no COVID-19, we instead take all data and forecast the Airbnb price under the current scenario for the next 30 days.

Using similar procedures as in (1) to determine possible parameter combinations, we have ultimately chosen to use a multiplicative error, damped-additive trend and additive seasonality (MAdA) enabled model. This model gives a training MAPE of 0.68% and a testing MAPE of 1.91% (Figure 7).

	RMSE	MAPE
Training Set	1.52	0.68%
Validation Set (30 Iterations)	2.32	1.91%

Similar to the residual behavior in (1), the residuals of predictions (Figure 8) using the full data is also subject to violation of assumptions, showing non-stationarity, non-normal distribution, and autocorrelation. Other parameters explored include AAA, MMM, and MAdM (Figure 9).

Parameters	Training RMSE	Training MAPE	Validation RMSE	Validation MAPE
AAA	1.51	0.68%	2.34	1.92%
MMM	1.51	0.68%	2.36	1.94%

MAdM	1.50	0.68%	2.40	1.97%

5.3 ARIMA

(1) Forecast 1 using pre-COVID data

For the ARIMA model, clearly, there is an upward trend in the original time series, which makes the time series non-stationary. As a result, we take a first-order differencing and check ACF and PACF graphs (Figure 10). From the ACF and PACF graphs of detrended data, we can see the PACF graph cuts off and the ACF graph slowly tails off.

However, since the ACF graph has spiked in seasonal lags and doesn't converge to 0, this suggests a weekly seasonality pattern in the data. These also appear to be non-stationary, so we take an additional weekly difference to track weekly seasonality (Figure 11).

We aim to find an appropriate ARIMA model based on the ACF and PACF graph shown above. The significant spike at lag 1 and no significant spikes immediately followed after lag 1 in the ACF suggests a non-seasonal MA(1) component, and the significant spike at lag 7 in the ACF suggests a seasonal MA(1) component. Consequently, we begin with a SARIMA(0,1,1)(0,1,1)[7] model, indicating a first and seasonal difference, and non-seasonal and seasonal MA(1) components.

By analogous logic applied to the PACF, there is a significant spike in lag 1, followed by an almost significant spike at lag 2, which doesn't indicate a non-seasonal AR term that we are interested in. The significant spike at lag 7 in the PACF suggests a seasonal AR(1) component. As a result, we could have another SARIMA(0,1,0)(1,1,0)[7] model.

By combining the AR and MA terms together, we could have a third SARIMA(0,1,1)(1,1,1)[7] model, and a fourth SARIMA(1,1,1)(1,1,1)[7] to validate our assumption about the AR term (Figure 12).

(1.1) SARIMA(0,1,1)(0,1,1)[7]

	RMSE	MAPE
Training Set	1.27	0.54%
Validation Set (30 Iterations)	1.88	1.51%

(1.2) SARIMA(0,1,0)(1,1,0)[7]

	RMSE	MAPE
Training Set	1.51	0.62%
Validation Set (30 Iterations)	1.82	1.46%

(1.3) SARIMA(0,1,0)(1,1,1)[7]

	RMSE	MAPE
Training Set	1.29	0.54%
Validation Set (30 Iterations)	1.90	1.52%

(1.4) SARIMA(1,1,1)(1,1,1)[7]

	RMSE	MAPE
Training Set	1.27	0.54%
Validation Set (30 Iterations)	1.88	1.51%

After comparing the results (Figure 13) from the four SARIMA models, we decide to forecast with (1) SARIMA(0,1,1)(0,1,1)[7] since (a) it gives a pretty good MAPE value on the training and validation sets (b) the model is robust (c) the graph makes more sense than model 2 when we comparing the actual COVID-19 period (d) it gives a similar result compared to the more complex models.

(2) Forecast 2 using all data

Similar to procedures in (1), this time we use all data and set aside 60 days from all data as a validation set. Four models same as above are applied and compared. The forecasted August pricing values are shown in Figure 14.

(2.1) SARIMA(0,1,1)(0,1,1)[7]

	RMSE	MAPE
Training Set	1.51	0.64%
Validation Set (30 Iterations)	2.63	2.16%

(2.2) SARIMA(0,1,0)(1,1,0)[7]

	RMSE	MAPE
Training Set	1.64	0.67%
Validation Set (30 Iterations)	2.66	2.19%

(2.3) SARIMA(0,1,0)(1,1,1)[7]

	RMSE	MAPE
Training Set	1.54	0.61%
Validation Set (30 Iterations)	1.41	1.16%

(2.4) SARIMA(1,1,1)(1,1,1)[7]

	RMSE	MAPE
Training Set	1.50	0.62%
Validation Set (30 Iterations)	1.41	1.16%

After comparing the results from the four SARIMA models, we decide to forecast with (4) SARIMA(1,1,1)(1,1,1)[7] since (a) it gives the best RMSE and MAPE values on the training and validation sets among the four models (b) the model is robust (c) the graph shows the model is able to capture the most recent upward trend in the COVID period, which indicates a potential recovery.

5.4. Neural Network Autoregression

(1) Forecast 1 using pre-COVID data

NNAR is equivalent to an ARIMA model but without stationarity restrictions. At the first try, we simply built an automatic NNAR model which is NNAR(27,1,14)[365] without tuning and using any exogenous variable. We set aside 60 days from Pre-COVID data as a validation set and have result as below:

	RMSE	MAPE
Training Set	0.07	0.04%
Validation Set (30 Iterations)	1.84	1.47%

To further explore how some external variables have impacted on Airbnb prices before COVID-19, we set up exogenous variables for training and validation set. Since we only have monthly data for external variables: national GDP and national unemployment rate, we repeat those monthly data 30 times to form daily external variables.

In the second model, we use external variables national GDP to fit our model and have model NNAR(27,1,15)[365] with xreg equal to GDP:

	RMSE	MAPE
Training Set	0.05	0.03%
Validation Set (30 Iterations)	1.87	1.50%

To compare which external variables has a larger impact on Airbnb prices, in the third model we use national unemployment rate to fit our model and have model NNAR(27,1,15)[365] with xreg equal to unemployment rate:

	RMSE	MAPE
Training Set	0.05	0.03%
Validation Set (30 Iterations)	1.80	1.44%

Obviously, the unemployment rate can help us have more accurate prediction results than national GDP. Moreover, the model without any external variable has better performance than the model with GDP as external variables. But we still want to try having both of these two external variables fitted our model, and we have the model 4 NNAR(27,1,16)[365] with xreg equal to GDP and unemployment rate:

	RMSE	MAPE
Training Set	0.04	0.02%
Validation Set (30 Iterations)	1.84	1.47%

The model 4 has the best performance in the training set, however, it has worse performance in the testing set than previous models, hence, model 4 is kind of overfitting.

Except automatic models which are generated by system, we also check ACF and Pacf graphs to check if there is any other model we can build (Figure 15). The ACF graph tails off as a damped wave pattern and the Pacf graph totally cuts off at lag 36 but we will try p=8 here since Pacf becomes stable after lag 8 and p=36 will be too large for fitting the model. For model 5 NNAR(8,1,5)[365] we have:

	RMSE	MAPE
Training Set	0.04	0.1%
Validation Set (30 Iterations)	1.78	1.42%

Model 5 is the best performance model since it has the most stable performance and does not overfit the training set while having the highest accuracy on validation set. Even though the model 5 has the best performance in the testing set, we still want to use both model 5 and model 3 to make predictions and compare the accuracy of the prediction result. Further discussions are in Section 6.

(2) Forecast 2 using all data

In this part, we will use all of the data to train and test NNAR models, meanwhile we also use all of the external variables data to fit our models. Same as part 1, we begin our model with system-generated automatic model NNAR(28,1,15)[365]:

	RMSE	MAPE
Training Set	0.09	0.05%
Validation Set (30 Iterations)	6.35	5.04%

The model 2 NNAR(28,1,16)[365] with exogenous variable-GDP:

	RMSE	MAPE
Training Set	0.10	0.04%
Validation Set (30 Iterations)	6.39	5.07%

The model 3 NNAR(28,1,16)[365] with exogenous variable-unemployment rate:

	RMSE	MAPE
Training Set	0.09	0.04%
Validation Set (30 Iterations)	6.15	4.88%

The model 4 NNAR(28,1,16)[365] with both exogenous variables:

	RMSE	MAPE
Training Set	0.09	0.04%
Validation Set (30 Iterations)	6.82	5.42%

We still check ACF and Pacf graphs (Figure 16) to see if there is any other model we can build. The ACF graph tails off as a damped wave pattern, the Pacf graph totally cuts off at lag 59, p=59 will cause excessive weights, hence we use p=8 since the most of spikes in the Pcaf graph roughly fall into the confidence band after lag 8. Therefore, we have model 5 NNAR(8,1,5)[365]:

	RMSE	MAPE
Training Set	0.65	0.3%
Validation Set (30 Iterations)	6.01	4.77%

The model 6 NNAR(8,1,5)[365] with exogenous variable-GDP:

	RMSE	MAPE

Training Set	0.52	0.25%
Validation Set (30 Iterations)	6.05	4.81%

The model 7 NNAR(8,1,5)[365] with exogenous variable-unemployment rate:

	RMSE	**MAPE**
Training Set	0.54	0.26%
Validation Set (30 Iterations)	6.56	5.21%

The model 8 NNAR(28,1,16)[365] with both exogenous variables:

	RMSE	**MAPE**
Training Set	0.50	0.25%
Validation Set (30 Iterations)	6.93	5.50%

In terms of the accuracy and model stability, model 5 NNAR(8,1,5)[365] has the best performance since it has the smallest MAPE and RMSE for the testing set and it has the smallest gap between training set accuracy and testing set accuracy as well. Therefore, we will use model 5 to make August Airbnb price predictions in Oakland.

5.5 Regression

(1) Forecast 1 using pre-COVID data

For a regression model, we first create pairs of scatter plots (Figure 17) to observe the relationships between different variables. Because we use monthly variables as daily variables through repeating the monthly data 30 times, the scatter plots do not continuously spread but we still can tell there is a downward trend between price and unemployment rate and an upward trend between price and the national GDP.

We create four regression models to capture the general trend of the Airbnb prices in Oakland before COVID-19. In the model 1, besides adding trend as x, we also use both unemployment rate and national GDP as x to forecast price y. The formula is tslm(price~trend+ur+gdp). In the second model we move the trend to see if there are only external variables, how exogenous variables will influence the price and the formula for model 2 is tslm(price~ur+gdp). For model 3 and model 4, we are going to use a single external variable as only one x to make linear regression models and the formulas for them are tslm(price~ur) and tslm(price~gdp). The accuracy for those models are:

	RMSE	**MAPE**
Model 1	3.34	1.97%

Model 2	3.34	1.98%
Model 3	3.52	2.23%
Model 4	3.48	1.98%

Least squares using all data will perform reasonably well on the underlying data (in terms of prediction error) as the least square model is fitted with the use of exactly these data points. Therefore, we are not going to split data into a training set and testing set here, but use all of the data to make predictions. However, when we do forecasts we use predicted external variables to predict price. For example, after we have all of the accuracy of the models, the model 1 has the highest accuracy and we will use model 1 to make predictions. The formula for model 1 is tslm(price~trend+ur+gdp), therefore, we need generate a new data set which includes predicted values of unemployment rate and national GDP in January and February as input values to forecast the Airbnb price in January and February. Therefore, for Pre-COVID19 forecast, we will choose model 1 to predict the Airbnb price in January and February.

(2) Forecast 2 using all data

For part2, we have a similar approach as in part 1. We first use all of the external variables data to update the predicted external variables values for August and September. Then we build another four models for forecasting the Airbnb prices in August and September. The first model with formula tslm(price~trend+ur+gdp); the second model with formula tslm(price~ur+gdp); the third model and the fourth one with formula tslm(price~ur) and tslm(price~gdp). All of methods are same as in part1, except we use all of data to fit our model and use August and September predicted external variables values to make forecasts:

	RMSE	MAPE
Model 1	3.42	2.15%
Model 2	3.47	2.22%
Model 3	3.88	2.59%
Model 4	3.72	2.45%

Since model 1 has the smallest errors we will model 1 to make predictions for August and September.

5.6 Decision Tree

Decision trees partition all possible values of the attributes into different subcategories or regions and each subcategory is assigned a value. By splitting variables we can predict the Airbnb price through a decision tree model.

(1) Forecast 1 using pre-COVID data

Preprocessing the variables is a significant step when we build the decision tree model. We first look at the unemployment rate variable before COVID-19 (Figure 18). The Acf graph tails off as a damped wave

pattern and Pacf graph totally cuts off at lag 1. Therefore, we use the AR(1) model to predict the unemployment rate in January and February.
Next, we look at the national GDP variable (Figure 19). Similar to variable unemployment rate, national GDP's Acf graph tails off as a damped wave pattern and Pacf graph totally cuts off at lag 1. Hence, we continue to use the AR(1) model to predict the GDP in January and February. Another reason we choose the AR(1) model is that both of these two variables are stationary before COVID-19 and using an Arima model will be appropriate.

Next step is to generate a training set and testing set. We create a data set which includes three variables: prices, unemployment rate and national GDP. We randomly select 363 records into the training set and 183 records into the testing set. Before we use variables to fit into a decision model, we need to create a linear regression model first and our linear regression model is using GDP and unemployment as X using price as Y; the formula is price~ur+gdp. We use this linear regression model and training set as input values to train our decision model and we have model visualizations as shown in Figure 20. We further use testing set to make predictions and have accuracy:

	R^2	**MAPE**
Training Set	0.69	1.82%
Testing Set	0.64	2.07%

This model is robust because the R square of the testing set is slightly lower than the training set while the MAPE of the testing set is slightly higher than the training set. Finally, we use the predicted external variable to make predictions for January and February Airbnb price.

(2) Forecast 2 using all data
We first need to make predictions for external variables in August and September (Figure 21). The unemployment rate sharply increased and the GDP substantially dropped at early 2020 due to COVID-19. Therefore, it is more appropriate to use the moving average method to make forecasts for external variables in August and September. The following steps are similar to part 1, we use all of the data to randomly generate a training set and testing set. We have 487 records as a training set and 245 records as a testing set. By using a linear regression formula price~ur+gdp as input, we have a decision model for all data as shown in Figure 22. We further use testing set to make predictions and have accuracy as below:

	R^2	**MAPE**
Training Set	0.63	1.69%
Testing Set	0.63	1.78%

6. Evaluation of Forecasts

6.1 Naive Forecasting

Serving as the baseline model, naive forecasting has reasonable results for both functions. Without showing COVID-19 to the model in function 1, the testing RMSE is 4.40 and MAPE is 3.03%. After showing additional current COVID-19 data to the model in function 2, the model performance had significant improvement with RMSE at 1.77 and MAPE at 1.77%.

By examining two different lengths of data to develop a naive forecast, it suggests that to adopt this model, it's most beneficial to include the most recent data to achieve the best model performance results. The final prediction with all data available using naive forecasting shows an average in the upcoming days at $123.1, which showed the price is bouncing back from the below $120 average in March (Figure 23).
However, one potential concern of naive forecasting is the critical role of frequency selection. Due to constraints on historical data, the frequency option is very limited. For example, a one year frequency in seasonality of 365 days will not capture any 2020 COVID-19 pattern because we only have half a year worth of data in 2020. In addition, the nature of naive forecasting only takes in data patterns that has happened in the past, under the uncertainty of 2020, is not believed to be the best model to adopt.

Therefore, although naive forecasting has reasonable performance, it should only be used as a baseline to compare with other models. If chosen to be the final model to deploy will require some re-evaluation of model parameters before actual use.

6.2 Exponential Smoothing

(1) Forecast 1 using pre-COVID data
Recall from the previous section that the top performing Holt-Winters exponential smoothing model based on pre-COVID data is one with multiplicative error, damped-additive trend, and additive seasonality. Using this model and training on all available pre-COVID data up until January 13, 2020, we achieved a training RMSE of 1.20 and a training MAPE of 0.50%. Applying this model to the COVID period up to July 16, 2020 yields a testing RMSE of 4.31 and a testing MAPE of 2.94%. The jump in testing MAPE suggests that this model is not a robust model and further shows that the average listing price during the COVID period deviates from historical pattern. In fact, the predicted price versus the actual price differs by an average of $3.48 dollars each day. Therefore, under the assumption that all other environmental factors are unchanged, COVID-19 is associated with an average price drop of about 3%. For Airbnb hosts that rely on collecting accommodation payments as a main source of income, they need to factor in the price decrease in addition to the loss in booking in their effort to cut cost. For Airbnb, who has a greater influence, they can strategically assist the hosts in pitching for additional subsidies from state and federal governments.

(2) Forecast 2 using all data
Applying the most optimal Holt-Winters model with multiplicative error, damped-additive trend, and additive seasonality on the entire dataset gives a training RMSE of 1.48 and a training MAPE of 0.67%.

From the graph below, it can be seen that the model fitted data traces the actual values fairly well. Although the model failed to capture the bigger fluctuations between February and June of this year, it was able to replicate the general direction of the change. Using this mode, the forecasted price for 30 days post July 16, 2020 is between $122 and $125, which is around the same level as June of last year (Figure 24).

For Airbnb hosts, this is definitely a good sign of recovery. In order to capitalize on the opportunity, hosts should ensure that their properties are properly sanitized after each stay. For Airbnb, it can use this forecast as a benchmark in estimating bookings and revenue for the next month.

6.3 ARIMA

(1) Forecast 1 using pre-COVID data

After training the best model SARIMA(0,1,1)(0,1,1)[7] on all pre-COVID data and testing on the post-COVID period, we found the RMSE on the training set is 1.23, on the test set is 4.00; MAPE on the training set is 0.51%, on the test set is 2.74%. The model is not robust enough because when the COVID-19 started, the pricing data fluctuated and showed a higher variability across time. Thus, the model, which performs well for the 2019 data as shown in the residuals plot, doesn't perform well enough on the COVID data. This indicates the different behaviors between the pre-COVID and during-COVID Airbnb housing prices (Figure 25).

(2) Forecast 2 using all data

After training the best model SARIMA(1,1,1)(1,1,1)[7] on all data and forecasting the next 30 days, we found the RMSE on the training set is 1.50; the MAPE on the training set is 0.62%. From the residuals graph, we can see that the model was able to capture the trend and seasonality very well in 2019; however, at the beginning of 2020 when the COVID pandemic started, the data showed a higher variability and the model didn't perform well. As time goes by, the range of residuals is narrowing down, which indicates the fact that the model works better in the middle of 2020. Thus, we are more confident that applying this SARIMA model to all data will give us an accurate result during the COVID period (Figure 26).

6.4. NN

(1) Forecast 1 using pre-COVID data

We have two best models for predicting pre-COVID data. One is model 3 NNAR(28,1,16)[365] with exogenous variable-unemployment rate and another one is model 5 NNAR(8,1,5)[365]. We used predicted external variables to predict the Airbnb price in January and February and then compare with the actual prices to check which model is more accurate:

	RMSE	MAPE
Model 3	2.22	1.58%
Model 5	3.93	3.16%

Model 3 has better performance even though model 5 has better performance in the training set and testing set. But the model 3 is still underfitting because it overestimated the price fluctuations during January and February (Figure 27).

(2) Forecast 2 using all data

Model 5 has 0.65 RMSE and 0.3% MAPE in the training set; 6.01 RMSE and 4.77% MAPE in the testing set. This model is not robust since it is overfitting, the errors in the training set are quite small but the errors in testing are around ten times the errors in the training set. High variability means we can not guarantee the model is robust (Figure 28).

6.5 Regression

(1) Forecast 1 using pre-COVID data

The best model in part 1 is model 1; the RMSE is 3.34 and MAPE is 1.97%. Linear regression could be a good model to reflect the general trend of Airbnb prices but could not have accurate predictions for daily price since it is a least square model. The residuals graph (Figure 29) shows there is a sharp drop in 2019, it means the linear regression can not reflect a sudden change in the price since the actual price in 2019 is in a trough but the linear regression model may keep forecasting it as average or increasing price. Therefore, the linear regression may still be good for predicting pre-COVID price but not after COVID price. The ACF graph also indicates there are some correlations existing. Using the Arima model will be appropriate to solve this problem.

(2) Forecast 2 using all data

The best model in the second part is model1 with RMSE 3.42 and MAPE 2.15%. Both RMSE and MAPE increase when using all of the data to make predictions. The reason is the price becomes much fluctuated during COVID19 and the linear regression becomes harder to detect the variability. As the residuals graph (Figure 30) shows after 2020, the linear regression model has more unconstant residuals and it becomes harder for the linear regression model to have an accurate prediction. Linear regression could a good model check the general trend but not a robust model for a fluctuated period, such as COVID19.

6.6 Decision Tree

(1) Forecast 1 using pre-COVID data

The decision tree model has 1.82% MAPE for the training set and 2.07% MAPE for the testing set. The R square for the training set is 0.69 and for the testing set is 0.64 which means the model captures around 70% of the variance for the training set and captures 64% of the variance for testing. Even though the residuals are constant spread and normal, this decision tree is not robust enough since we expect R square over 90% to have a significant R square (Figure 31).

(2) Forecast 2 using all data

The decision tree model has 1.69% MAPE for the training set and 1.78% MAPE for the testing set. It reflects 63% of variance for both training and testing sets. This decision tree model is worse than part 1's decision tree model since it has lower R square in both training and testing sets. Moreover, the R square in the training set is the same as the testing set which means the training set is underfitting. The reason could be the model can not capture the steep change in all of the variables due to COVID-19 (Figure 32).

7. Conclusion

This project assesses the impact of the COVID-19 pandemic on the average listing price for Oakland Airbnb properties and recommends the most appropriate model for forecasting average listing price in a 30-day time horizon. By applying various time series models including the seasonal naive model, Holt-Winters Exponential Smoothing model, ARIMA model, neural network, linear regression model, and decision tree model with external variables GDP and unemployment rate, the results suggest the best model to be SARIMA(0,1,0)(1,1,1)[7], which gives a robust and accurate result with a training MAPE of 0.61% and a validation MAPE of 1.61%. We expect the actual prices to be around 1.61% off from the predicted values under the 30-day forecasting horizon from July 17, 2020 to Aug. 15, 2020.

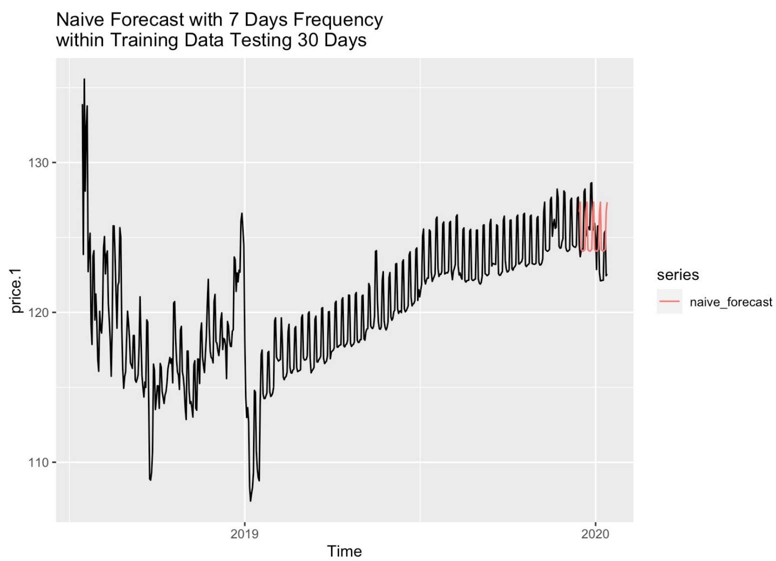
This project and the final model is scalable, valuable and multipurpose both externally and internally. Externally, with the insights from this project, Airbnb could work closely with the Oakland, CA government to help map out upcoming incentives or additional travel resources for the short-term rental market recovery. By maintaining and enhancing the relationship with the government, this win-win strategy between Airbnb and the government will be critical for Airbnb's long-term growth in the local rental business. In addition, this model could be incorporated into Airbnb websites as a resource. By recommending listing prices and forecasting travel recovery for property owners, Airbnb can foster a better relationship burned due to the COVID-19 pandemic. By providing forecasted housing prices for travellers, Airbnb could build a credible brand image and assist users in navigating through the uncertainty in this pandemic.

Internally, this project helps executives and managers to understand how exactly the COVID-19 pandemic influenced the tourism industry, especially Oakland Airbnb properties, by providing a comparison between forecasted average prices without COVID-19 and actual prices during COVID-19. In addition, this model is beneficial for business development and marketing teams to monitor the industry recovery and set up plans ahead according to the forecasted price values.

The next step is deploying the forecasting model and architecturing a pipeline. There are three key areas the team needs to consider before embarking on the project: data storage and retrieval, frameworks and tools, feedback and iteration. First, since the average listing price is on a daily basis, an offline learning system is sufficient for making predictions. Second, the frequency of providing predictions should be based on the business needs. Generally speaking, retrieving and retraining the model per week is suitable and cost-effective. Third, the feedback and performance of the model should be continuously monitored to track the recovery point. If the model performs worse and a large difference between predicted and actual values is discovered, we recommend repeating the model selection process and deciding on the most appropriate model to capture the potential future changes in trend or variability. By having a reliable model deployment pipeline, this project is scalable to smaller segmented locations in Oakland, California or more cities in the United States.

8. Appendix

Figure 1



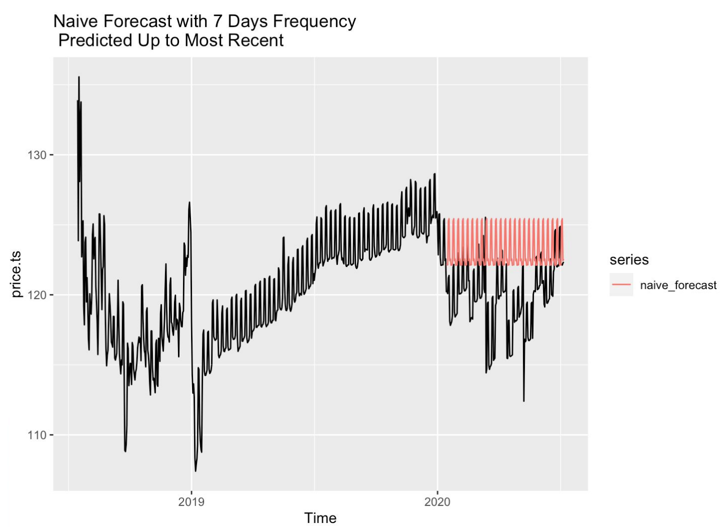
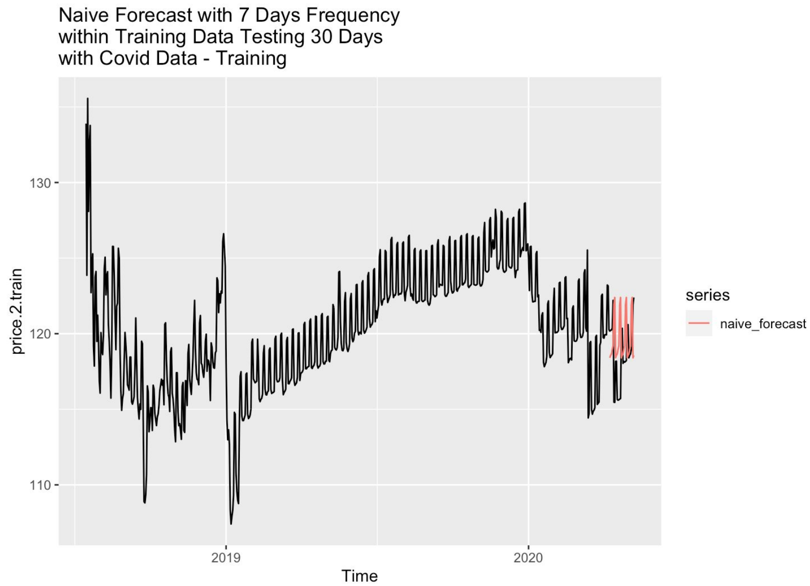


Figure 2


Naive Forecast with 7 Days Frequency
within Training Data Testing 30 Days
with Covid Data - Training
price.2.train
130
120
110
2019
2020
Time
series
naive_forecast

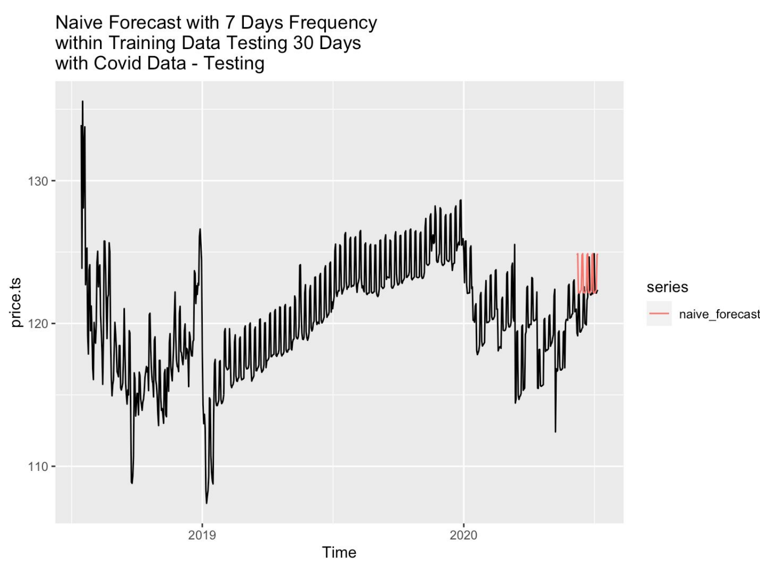

Naive Forecast with 7 Days Frequency
within Training Data Testing 30 Days
with Covid Data - Testing
price.ts
130
120
110
2019
2020
Time
series
naive_forecast

Figure 3

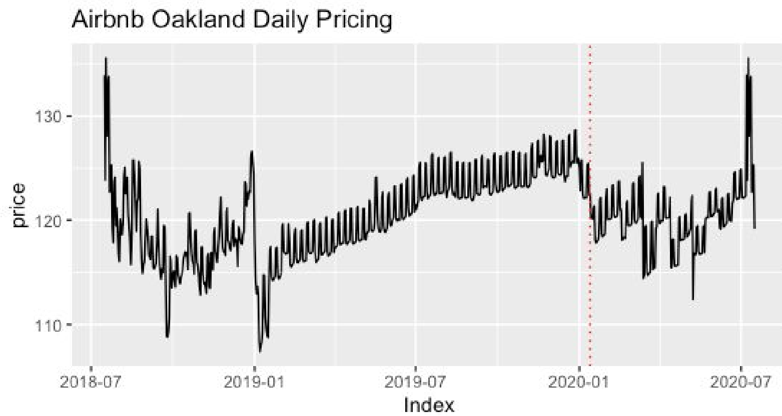


Figure 4

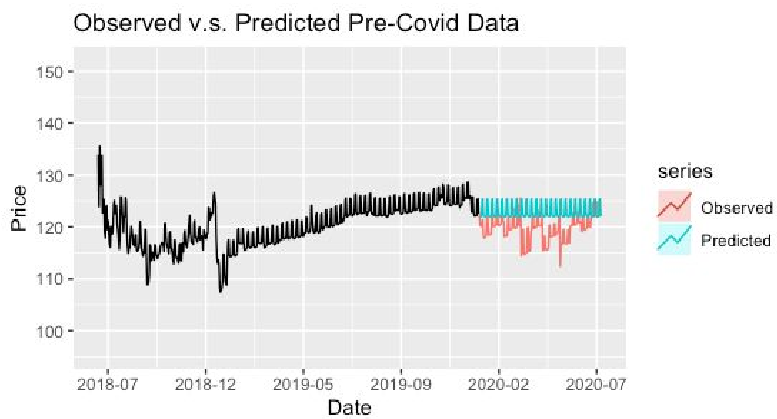


Figure 5

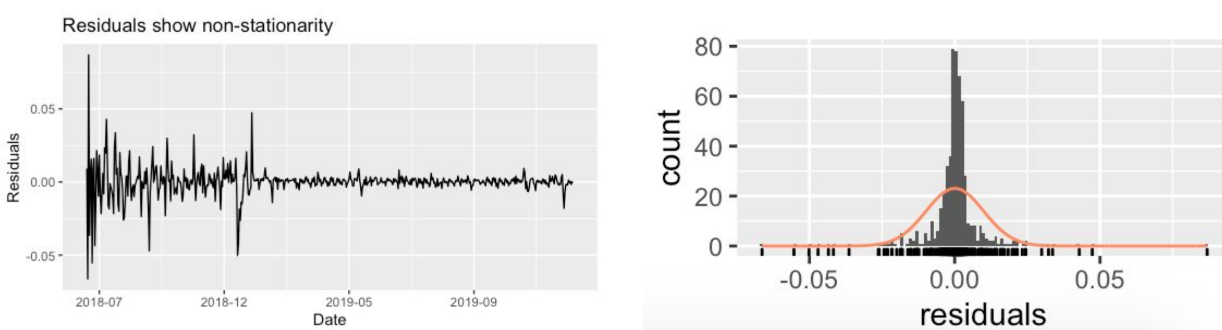


Figure 6

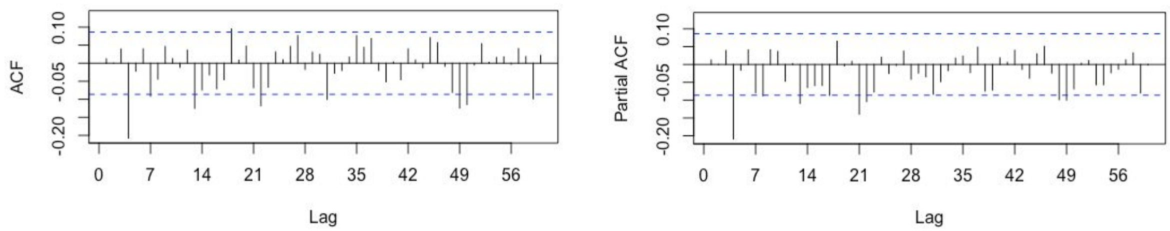


Figure 7

Figure 8

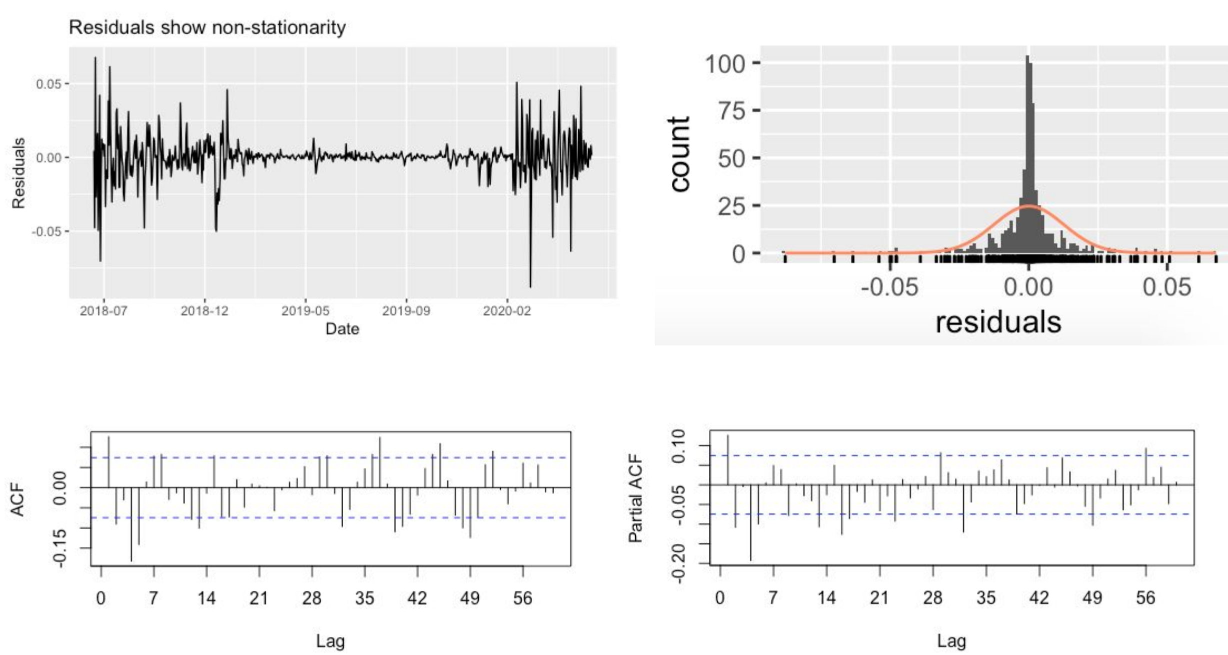
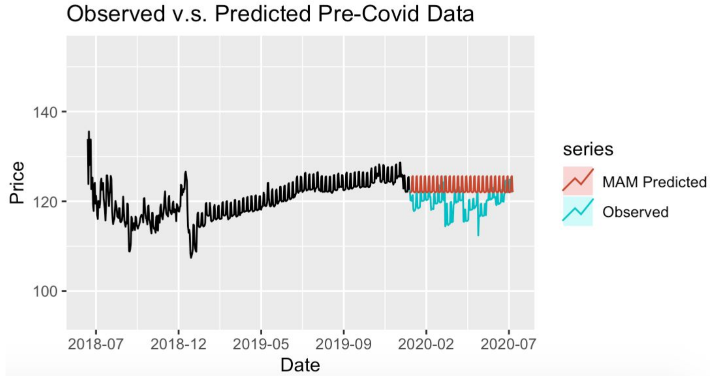
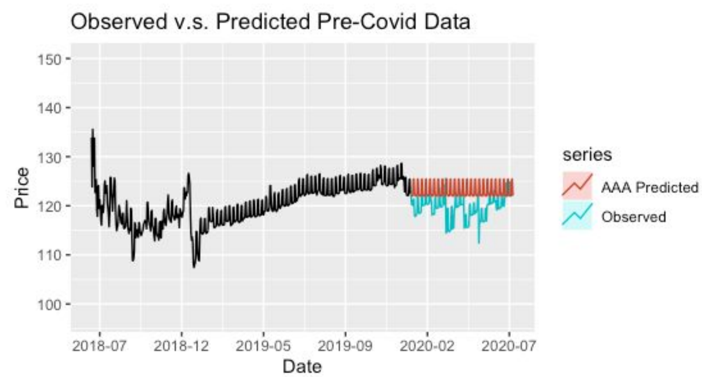
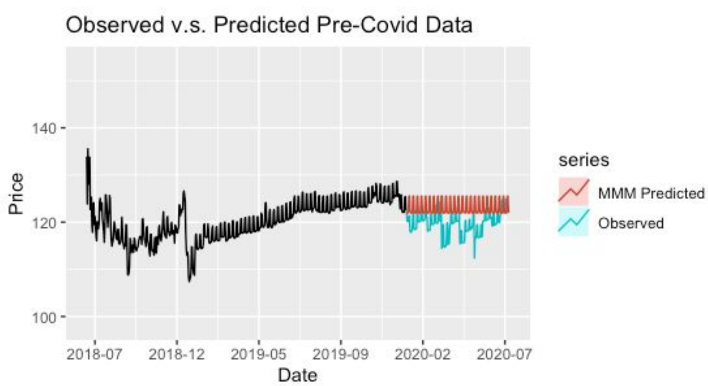


Figure 9

(1) Forecast using pre-COVID data







(2) Forecast using all data

Figure 10

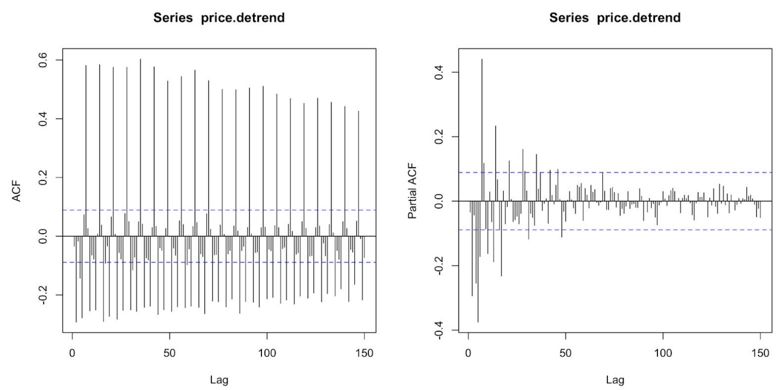


Figure 11

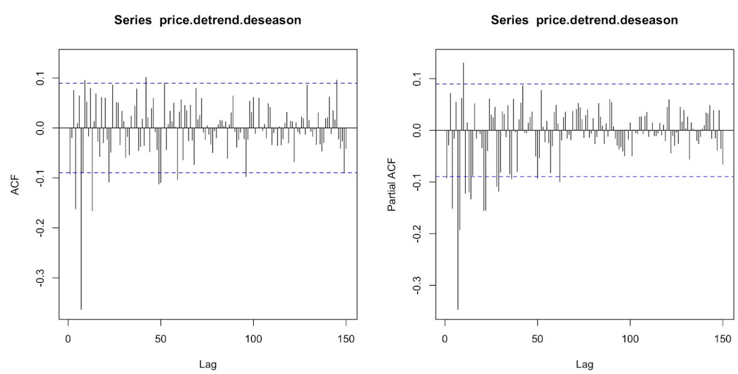


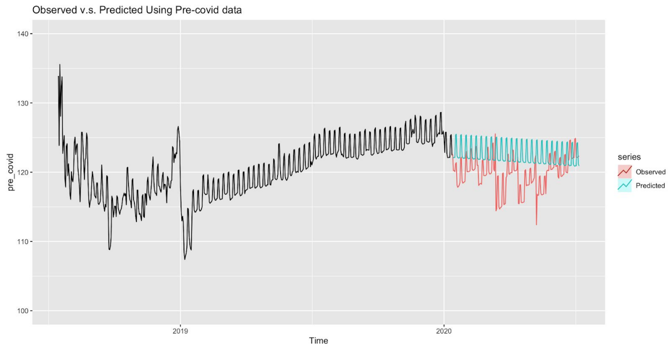
Figure 12

TABLE 6.2 ACF AND PACF PATTERNS FOR GENERAL AR, MA, AND ARMA SCHEMES

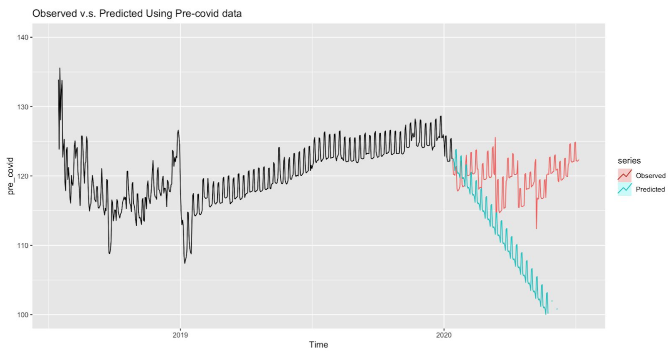
Function	AR(p) Scheme	MA(q) Scheme	Mixed ARMA Scheme
ACF	Tails off as a damped wave pattern or damped exponential	Finite, q spikes	Tails off as a damped wave pattern or damped exponential
PACF	Finite, p spikes	Tails off as a damped wave pattern or damped exponential	Tails off as a damped wave pattern or damped exponential

Figure 13

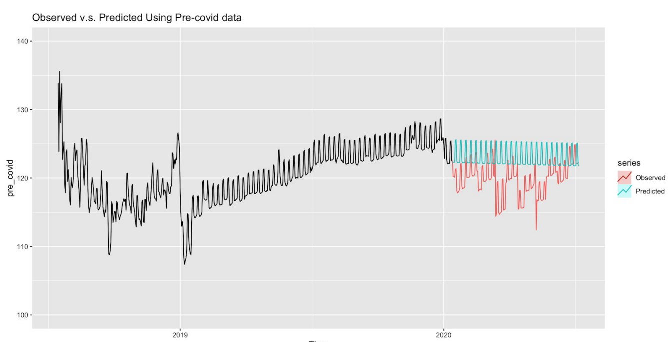
(1.1) SARIMA(0,1,1)(0,1,1)[7]



(1.2) SARIMA(0,1,0)(1,1,0)[7]



(1.3) SARIMA(0,1,0)(1,1,1)[7]



(1.4) SARIMA(1,1,1)(1,1,1)[7]

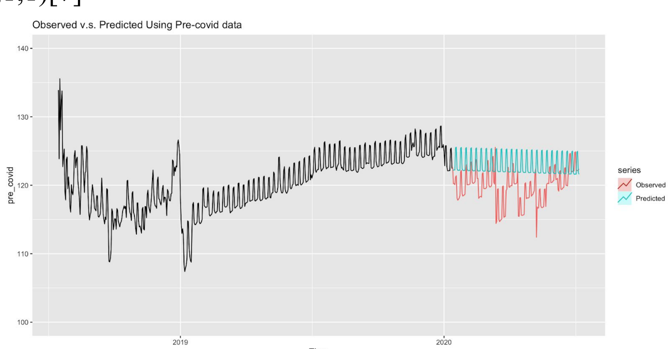
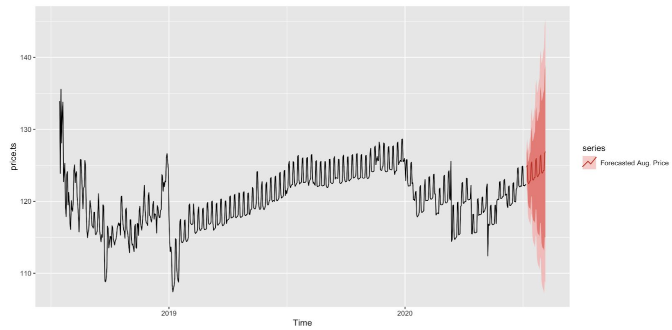
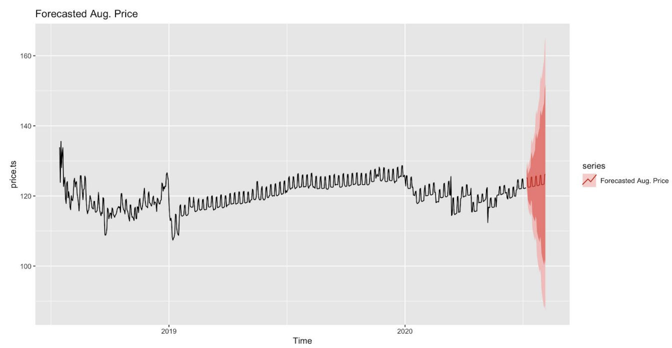


Figure 14

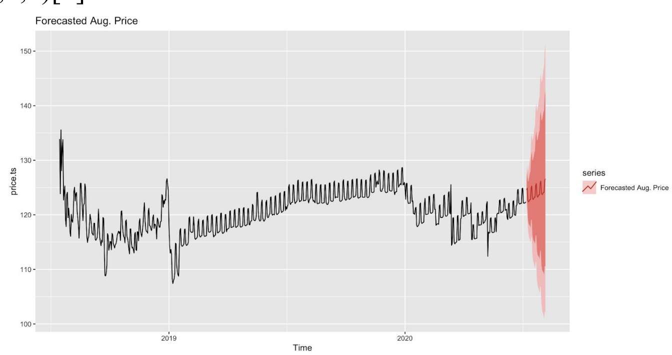
(2.1) SARIMA(0,1,1)(0,1,1)[7]



(2.2) SARIMA(0,1,0)(1,1,0)[7]



(2.3) SARIMA(0,1,0)(1,1,1)[7]



(2.4) SARIMA(1,1,1)(1,1,1)[7]

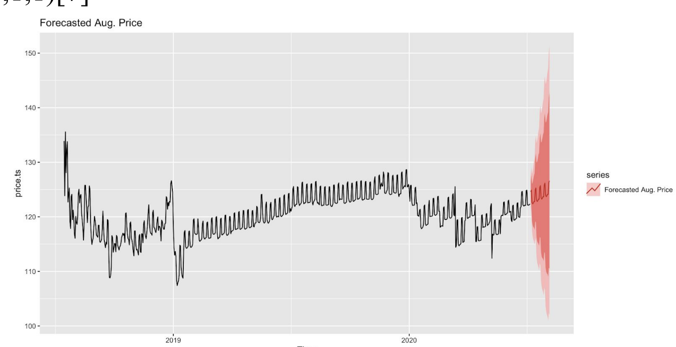


Figure 15

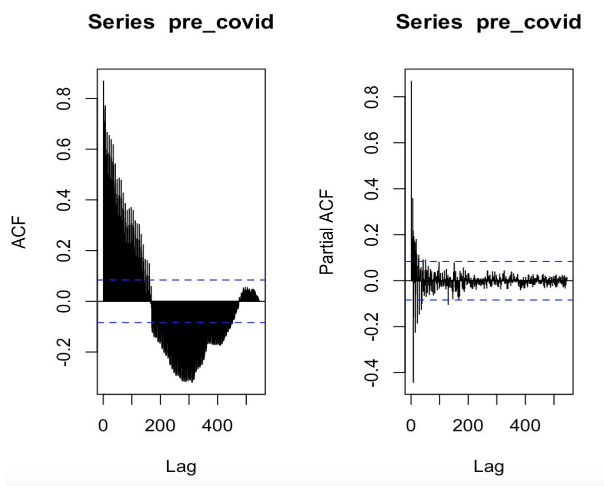


Figure 16

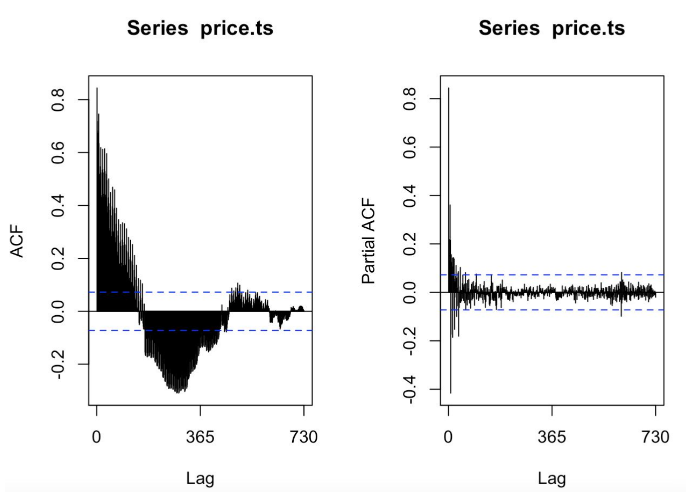


Figure 17

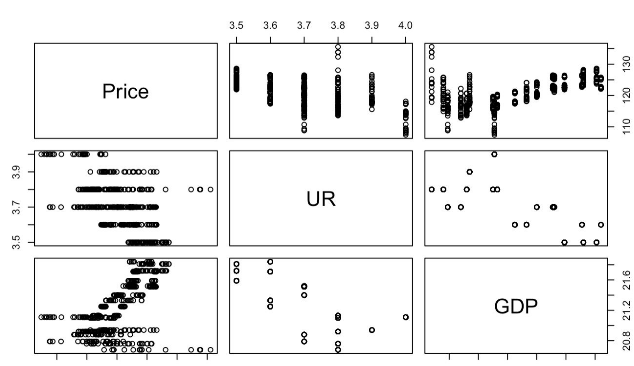


Figure 18

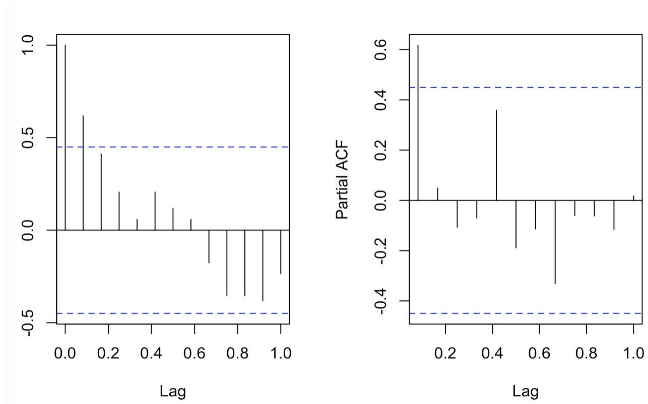


Figure 19

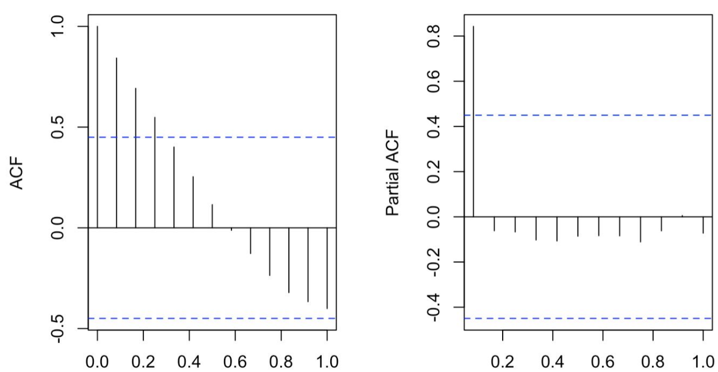


Figure 20

```
node), split, n, deviance, yval
      * denotes terminal node

 1) root 363 6569.00 120.4
   2) gdp.xts < 21.365 210 3138.00 118.0
     4) gdp.xts < 20.72 11  391.90 126.8 *
     5) gdp.xts > 20.72 199 1845.00 117.5
      10) ur.xts < 3.95 178 1266.00 118.0
        20) gdp.xts < 20.93 84  694.00 117.1
          40) gdp.xts < 20.775 18  170.80 119.9 *
          41) gdp.xts > 20.775 66  350.10 116.4 *
        21) gdp.xts > 20.93 94  456.10 118.8 *
      11) ur.xts > 3.95 21  213.60 113.6 *
   3) gdp.xts > 21.365 153  558.30 123.7
     6) gdp.xts < 21.455 21   62.12 121.5 *
     7) gdp.xts > 21.455 132  372.90 124.1 *
```

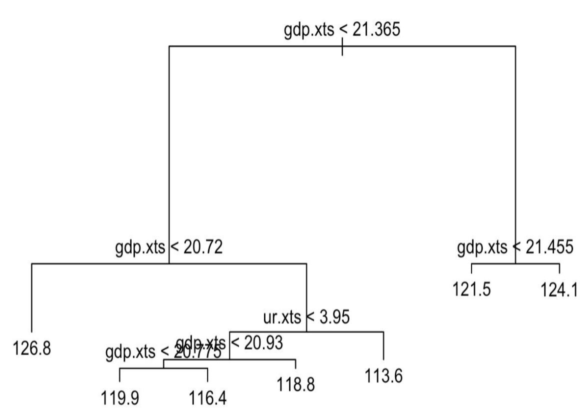


Figure 21

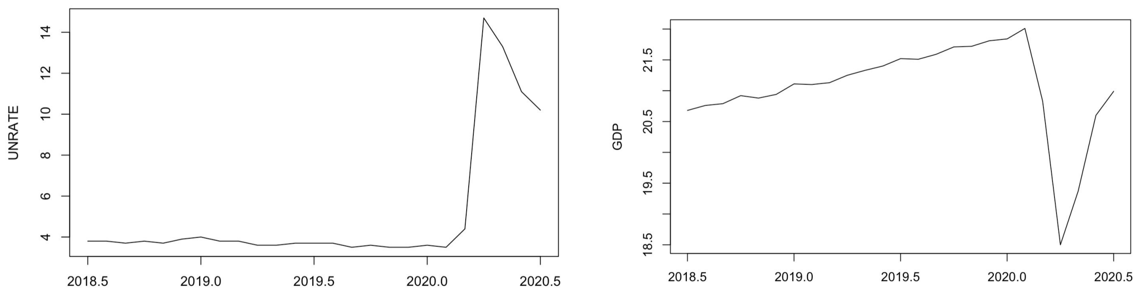


Figure 22

```
node), split, n, deviance, yval
      * denotes terminal node

 1) root 487 7379.00 120.2
   2) gdp_xts < 21.365 307 3898.00 118.5
     4) gdp_xts < 20.775 96 1192.00 120.7
       8) gdp_xts < 19.985 33  211.80 118.8 *
       9) gdp_xts > 19.985 63  798.00 121.7
        18) gdp_xts < 20.72 40  438.40 122.6 *
        19) gdp_xts > 20.72 23  269.00 120.1 *
     5) gdp_xts > 20.775 211 2040.00 117.5
      10) gdp_xts < 21.19 167 1737.00 117.0
        20) gdp_xts < 21.105 123 1160.00 117.7
          40) ur_xts < 3.85 76  397.80 116.4 *
          41) ur_xts > 3.85 47  457.60 119.7 *
        21) gdp_xts > 21.105 44  400.50 115.3
          42) ur_xts < 3.9 20   34.01 117.4 *
          43) ur_xts > 3.9 24  215.90 113.6 *
      11) gdp_xts > 21.19 44  128.60 119.3 *
   3) gdp_xts > 21.365 180 1012.00 123.2
     6) gdp_xts < 21.825 141  626.30 123.7
      12) gdp_xts < 21.715 98  304.60 122.9
        24) gdp_xts < 21.455 20   50.26 121.2 *
        25) gdp_xts > 21.455 78  180.10 123.3 *
      13) gdp_xts > 21.715 43  115.90 125.5 *
     7) gdp_xts > 21.825 39  196.10 121.2 *
```

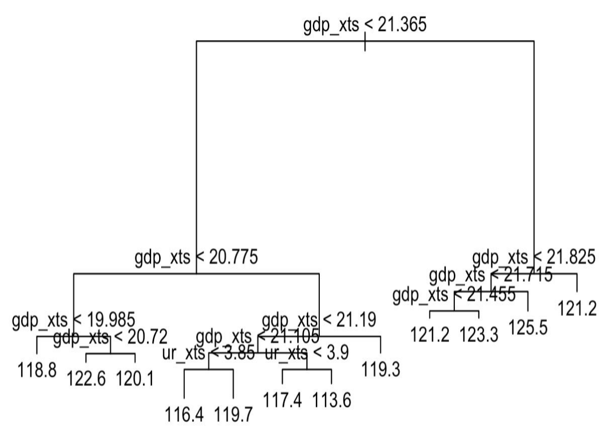


Figure 23

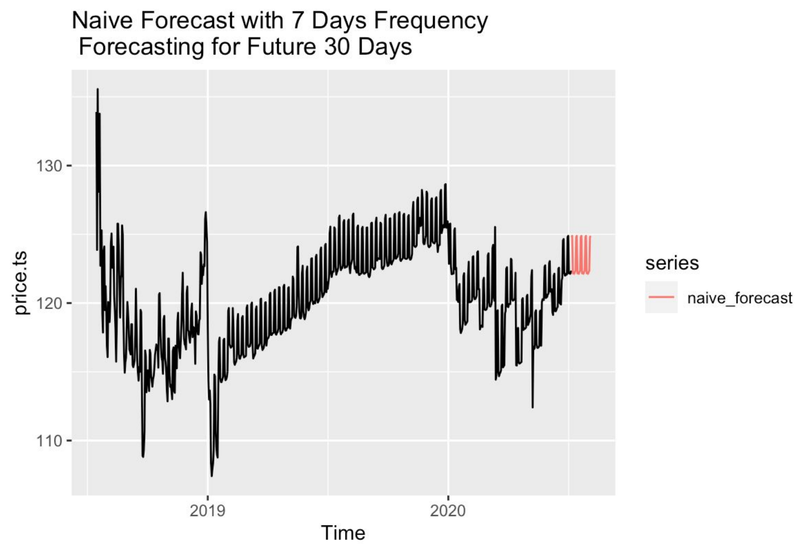
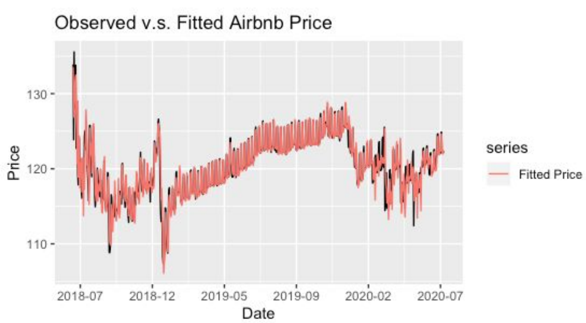
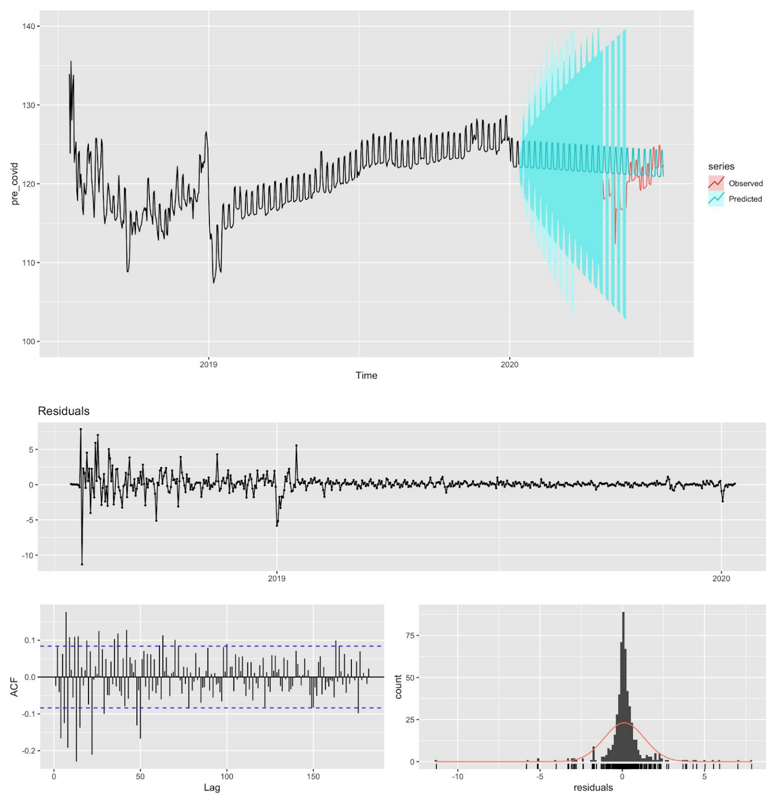


Figure 24



<u>Figure 25</u>



pre_covid
Time
series
Observed
Predicted
Residuals
ACF
Lag
count
residuals

Figure 26

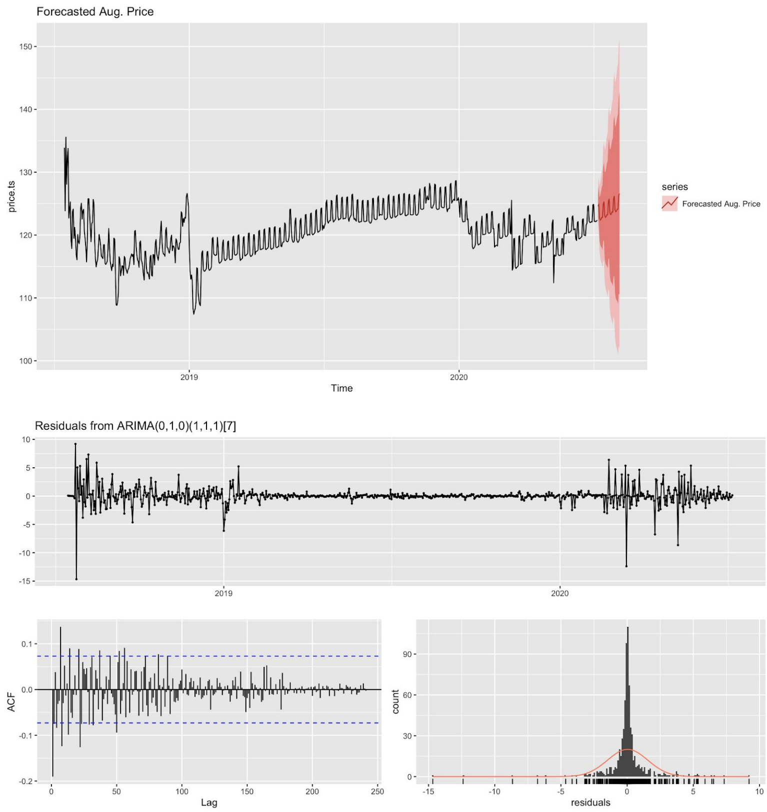


Figure 27

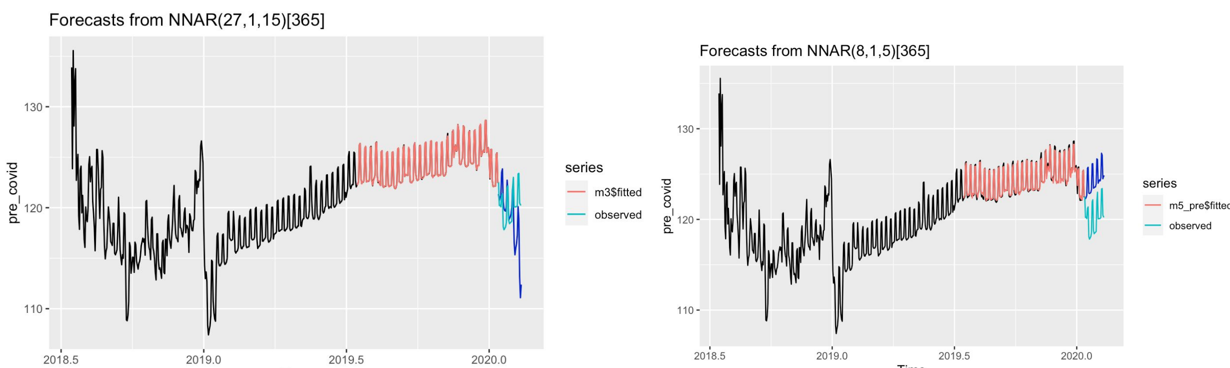


Figure 28

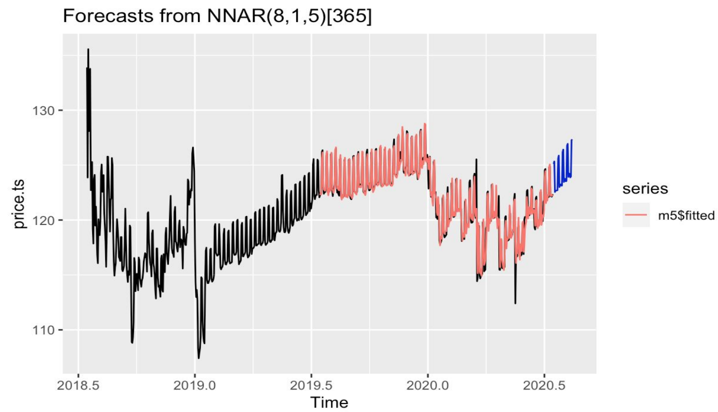


Figure 29

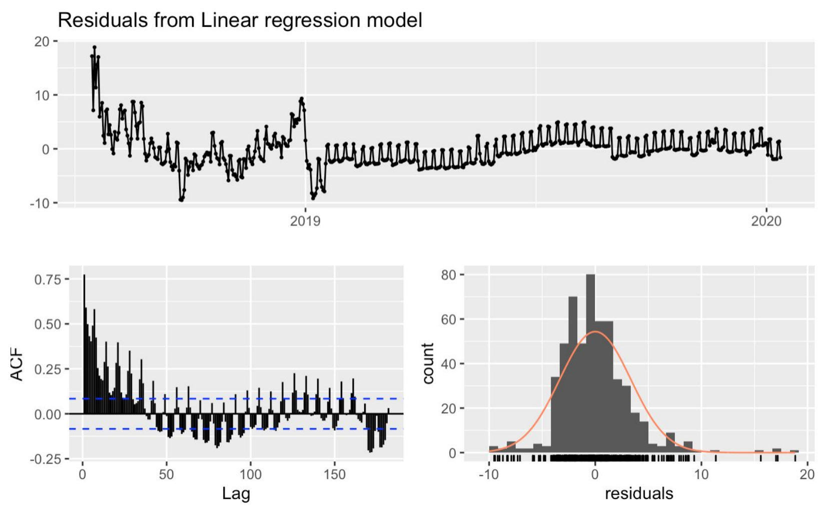


Figure 30

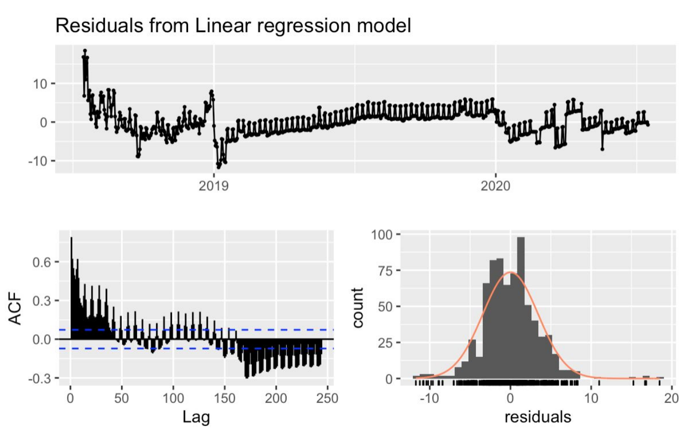


Figure 31

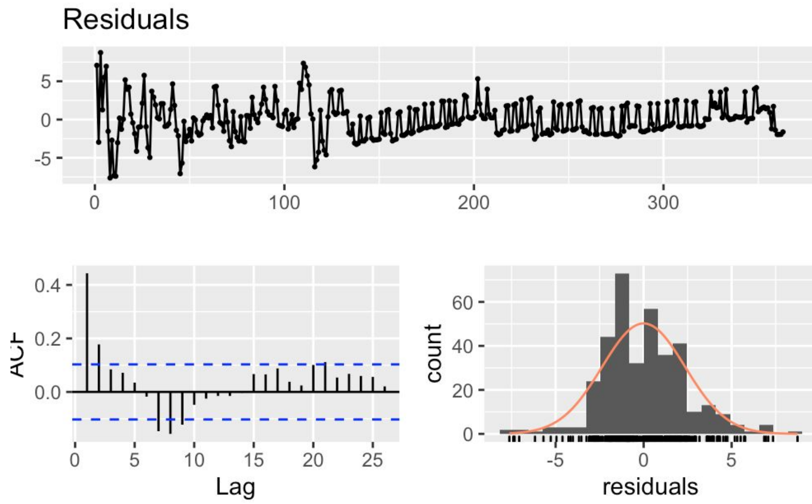
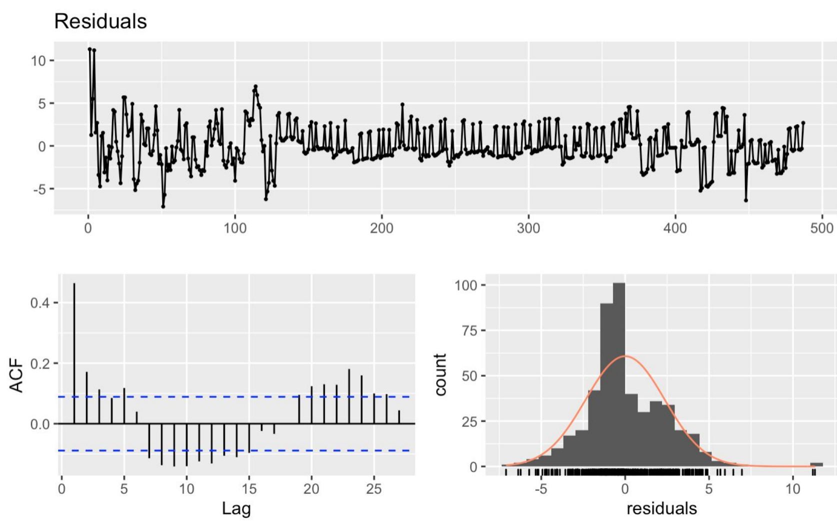


Figure 32



9. References

WHO Timeline - COVID-19.
https://www.who.int/news/item/27-04-2020-who-timeline---COVID-19
Timeline of WHO's response to COVID-19.
https://www.who.int/news/item/29-06-2020-COVIDtimeline
Europe's coronavirus lockdown measures compared.
https://www.politico.eu/article/europes-coronavirus-lockdown-measures-compared/
S-1.
https://www.sec.gov/Archives/edgar/data/1559720/000119312520294801/d81668ds1.htm
Airbnb Quarterly Revenue Drops 67%. IPO Still Planned.
https://www.bloomberg.com/news/articles/2020-08-12/airbnb-revenue-tanks-67-in-second-quarter-ipo-planned
Hosts angry at Airbnb for COVID-19 policies.
https://business.inquirer.net/293430/hosts-angry-at-airbnb-for-COVID-19-policies
Airbnb's Coronavirus Crisis: Burning Cash, Angry Hosts and an Uncertain Future.
https://www.wsj.com/articles/airbnbs-coronavirus-crisis-burning-cash-angry-hosts-and-an-uncertain-future-11586365860?mod=rss_Technology