MAT005: Forecasting the Demand for Ambulance Services in Jakarta

Bailey Millns: c2098580

Introduction

Accurate forecasting is necessary for ambulance services to provide efficient medical treatment and transportation for patients. These services are time sensitive. If we can build an effective model to predict the number of ambulance calls, we could save time, money and potentially lives. The aim of this project was to predict the future demand of ambulance services in Jakarta.

Data

The data used in this project contained the date, time of call and the city municipality from which a call to the ambulance service was made. The dates run from January the 1st 2019 to May the 31st 2019, having 22540 total number of observations. A table containing the top five most called from city municipalities can be found in table 1.

Table 1: Total calls made in the city municipality

City Municipality	Number of calls
Jakarta Pusat	5683
Jakarta Barat	5220
Jakarta Timur	4554
Jakarta Selatan	3649
Jakarta Utara	2991

We can see from figure 1 that there are more calls during the middle of the day (from 10:00 to 14:00) and we can see from figure 2 that there are more calls on weekdays than on weekends, with Monday being the day with the most calls.

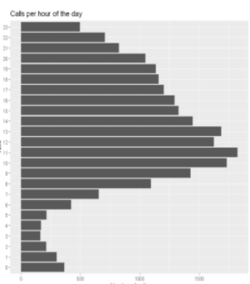

Figure 1: Ambulance calls per hour

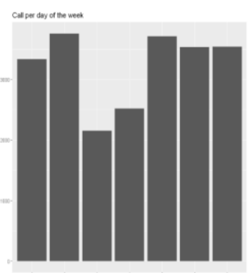

Figure 2: Ambulance calls per day

The plot in Figure 3, which is a scatter plot of the dataset. We can see that the variance may be unstable, most of the observations are not close to mean which is 149.258.

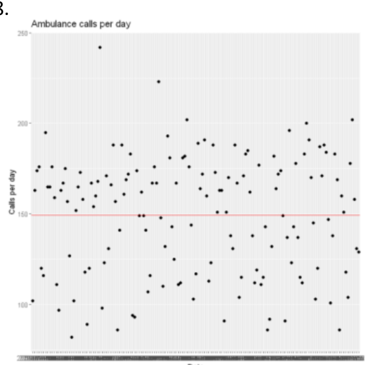

Figure 3: Scatter plot of the calls per day

The data will first need to be cleaned and pre-processed into time series data. Then the data will be split into training data and testing data. The training data contained dates from 01/01/2019 to 16/04/2019 and the testing data contained dates from 17/04/2019 to 31/05/2019 (a 70-30 split). The analysis will be conducted on the training set, then if transformations are needed, they will be applied to the full data set.

Analysis

Decomposition, ACF and PACF plots, of the training data, were plotted to analyse trend, seasonality and whether the time series was stationary. A Dickey-Fuller test was also conducted to test whether the time series is stationary.

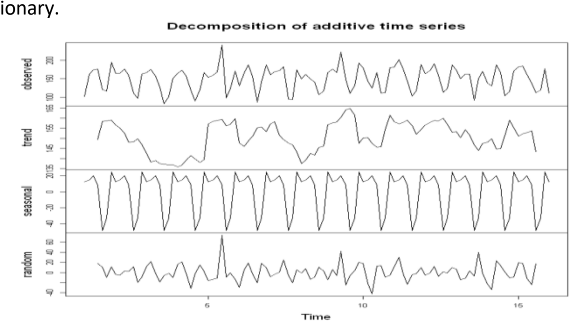

Figure 3: Decomposition plot

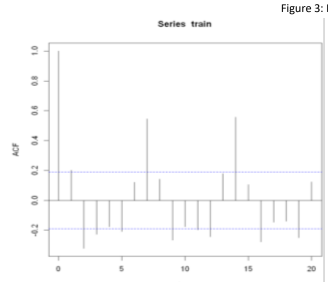

Figure 4: ACF plot of the original training data

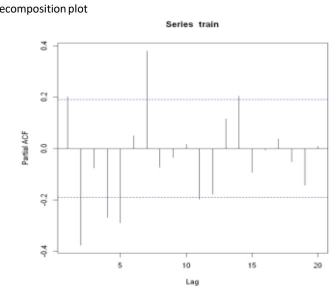

Figure 5: PACF plot of the original training data

The plots shown above suggest we have strong weekly seasonality present, the significant lags 7 and 14 suggest this. The stl() and seasadj() functions in R were used to deseasonalize the full time series data. The data also appears non-stationary, as the lags in the ACF plot to not go to zero quickly. This hypothesis was tested with the Dickey-Fuller test. The data obtained a p-value of 0.01. The p-value is less than the significance level of alpha equals 0.05 so we reject the null hypothesis and conclude the Dickey-Fuller test suggests the time series is stationary. During the first run of the models, it was found that the residuals of the models were not normally distributed. A Box-Cox transformation was applied to the time series data, but the transformation did not assist with making the residuals normally distributed so it was not applied to the full data set.

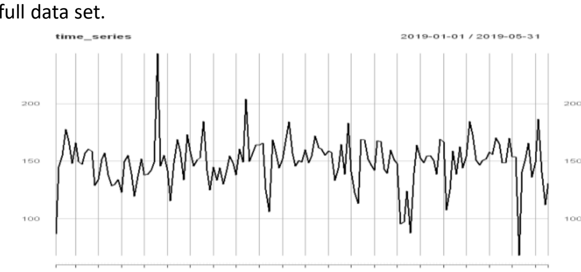

Figure 6: Time series plot after deseasonalizing

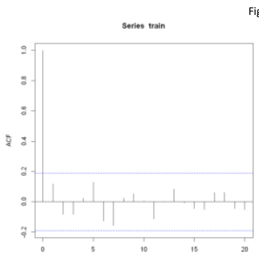

Figure 7: ACF plot after deseasonalizing

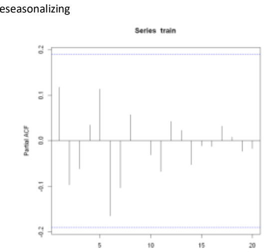

Figure 8: PACF plot after deseasonalizing

We can see, in figure 7, that the significant lags 7 and 14 have been reduced and are no longer significant. All lags are within the threshold, which indicates absence of autocorrelation.

Models

For this project I will evaluate five different time series models, the best model will used to forecast the number of calls for 7 days in the future. The best model will be chosen based upon several factors. The models will be compared by calculating the MSE and MAPE on the test data to find the best model. The residuals of the best model will be analysed, the residuals should be normally distributed, the ACF plot of the residuals will be analysed to see if there is absence of auto-correlation. The Ljung-Box test, is also used to test for autocorrelation. A large p-value (>0.05) will indicate the residuals are not distinguishable from white noise. If these assumptions are met, the model will be used to forecast 7 days. The five different models evaluated are: Naïve, simple exponential smoothing, Holt's linear model, linear regression and ARIMA. In figure 10 we can see an example forecast from the linear regression model.

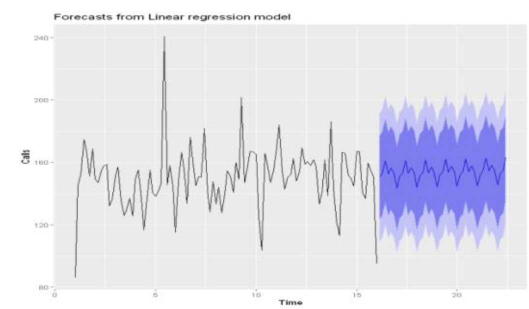

Figure 9: Linear regression forecast

A summary of the model evaluation, using MSE and MAPE, can be found in table 2.

Table 2: Model evaluation

Metric	ARIMA(0,0,1)	Naive	SES	Holt's Linear	Linear Regression
MSE	1106.4285837	3862.3253577	1119.584593	1237.1754781	1289.1123214
MAPE	0.2183842	0.3296079	0.220523	0.2403609	0.2455438

A perfect model has an MSE and MAPE value of 0, so the closer to 0 for both metrics the better the model. We can see that the ARIMA(0,0,1) model performs the best. The residuals will now be analysed, this was completed using the checkresiduals() function in R.

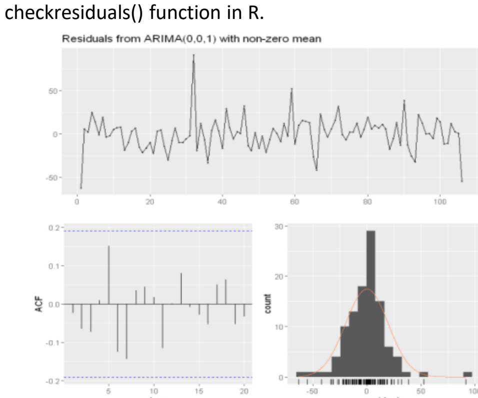

Figure 10: Diagnostic criteria of the ARIMA(0,0,1) residuals

The ACF plot of the residuals from the ARIMA(0,0,1) model in figure 10 indicates the absence of auto-correlation, demonstrating that the residuals are behaving like white noise. In addition, the Ljung-Box test returns a p -value of 0.4879, which is larger than the significance level of 0.05 also suggesting that the residuals are behaving like white noise. Testing the normality assumption of the residuals, with the Shapiro-Wilk test, we obtain a p-value of 3.689e-05 which suggests the residuals do not follow a normal distribution. All models evaluated had p-values which suggested that the residuals are not normally distributed. Due to the central limit theorem the residuals will have approximately a normal distribution. The ARIMA(0,0,1) will be the model used to forecast, as it is the best model evaluated with the considerations above.

Forecast

A summary of the 7-day forecast with the ARIMA(0,0,1) model can be found below in figure 11 and table 3. The diagnostic criteria is found in figure 12.

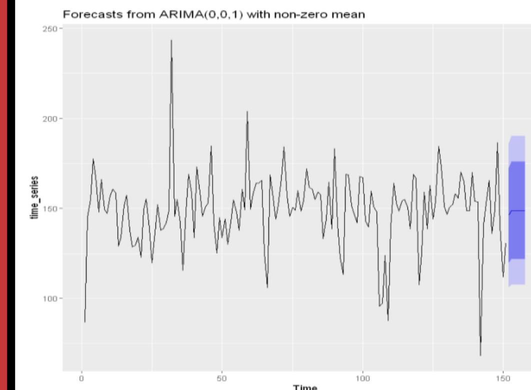

Figure 11: ARIMA(0,0,1) 7 day forecast

Table 3: ARIMA(0,0,1) 7 day forecast, with 80% and 95% prediction intervals

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
152	146.2590	120.0916	172.4265	106.2393	186.2787
153	148.8022	121.7994	175.8050	107.5050	190.0995
154	148.8022	121.7994	175.8050	107.5050	190.0995
155	148.8022	121.7994	175.8050	107.5050	190.0995
156	148.8022	121.7994	175.8050	107.5050	190.0995
157	148.8022	121.7994	175.8050	107.5050	190.0995
158	148.8022	121.7994	175.8050	107.5050	190.0995

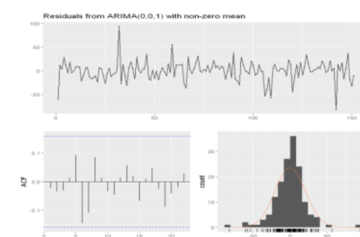

Figure 12: Diagnostic criteria of the ARIMA(0,0,1) residuals

The forecast results show a mean number of calls of 148.44 from June the 1st to June the 7th. The diagnostic criteria indicates the absence of autocorrelation, as seen in figure 12. The Shapiro-Wilk test on the residuals gives a p-value of 9.99e-07, which suggests the residuals are not normally distributed. Due to the central limit theorem the residuals will have approximately a normal distribution.

Conclusion and future work

We chose the ARIMA(0,0,1) as the best model to forecast, as this model obtained a MAPE value of 0.2184 and MSE value of 1106.429 on the testing data set, the best out of all models evaluated. ARIMA models allow both autoregressive (AR) components as well as moving average (MA) components and are more flexible than the other models tested. It was no surprise the ARIMA model performed the best. We cannot say that this is the accurate forecast, due to the size of the error statistics. In future work I would try more advanced models, such as Facebook Prophet or a neural network, to forecast this data. I would also try the technique of dynamic harmonic regression for deseasonalizing the data, as the method used in this project did not completely remove the seasonality. More data (longer dates and more features) would also be useful to further investigate the demand in ambulance services.

References

- Rob J Hyndman and George Athanasopoulos 2018. Forecasting: Principles and Practice. Monash University. Available at: https://otexts.com/fpp2/.Steins, K. et al. 2019.
- Forecasting the Demand for Emergency Medical Services. Linköping University. Available at: https://core.ac.uk/download/pdf/211327292.pdf.