

Forecasting Daily Ambulance Run of a Wales Ambulance Service

1. Introduction to the problem

Ambulance Runs data was aggregated Hourly, Daily and Weekly, plotted and I decided to go with the Daily aggregated data. Data will be split into a training (March – July) and validation (August – September) data set. Time series forecasting models will be used and validated on the validation data, and one model selected to forecast 61 days for October and November.

2. Numerical Summaries

Mean = 153.76

Q1 = 119.25

Q3 = 183.75

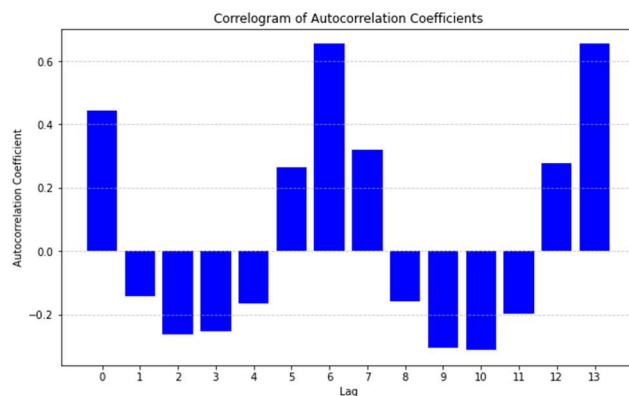
Median = 163.5

IQR = 64.5

Variance = 1393

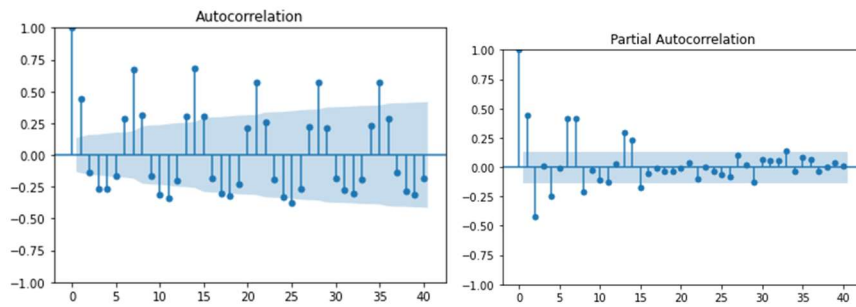
Numerical Analysis of the data shows that it is widely varied around the mean

A correlogram was plotted for the data and shows strong relationship between lag 6 and lag 13 suggesting a seasonal component in the data and similarly on lag (2,3) and lag (9,10)



I performed an augmented Dickey-fuller test to assess stationarity, ADF stat was -2.132 and p-value = 0.231 which was not a strong enough rejection of the null hypothesis and I concluded the data was non-stationary.

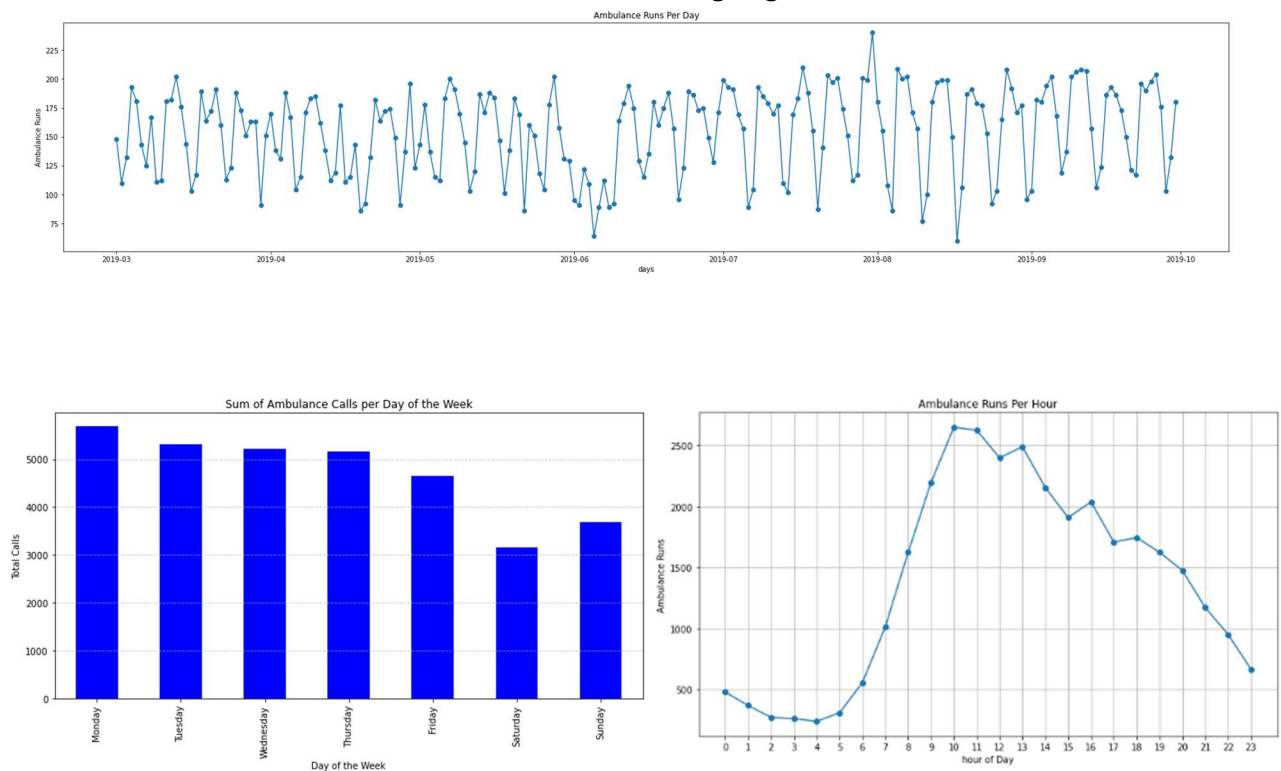
ACF and PACF plots were done on the data and this shows that the data is not stationary as there is no decay as the lag increases,



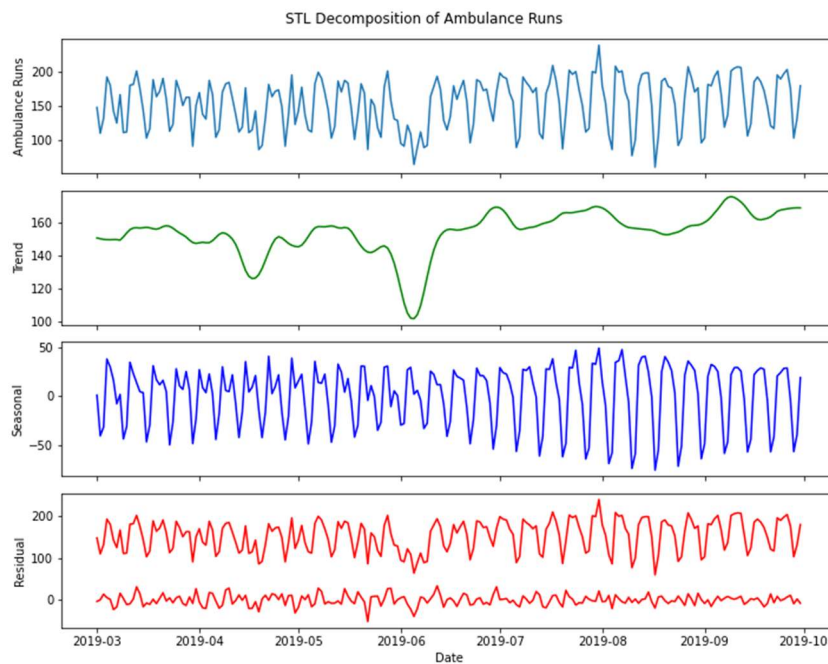
3. Graphical Summaries

plot of the data shows a sharp increase of Ambulance Runs from the morning headed towards midday, followed by a steady decline though the afternoon and night. Daily plots of the data shows seasonality with dips during the the weekends and spikes during the day. Daily analysis of the data shows a large quantity of ambulance calls are made on Mondays and this gradually reduces through the week, with Saturday as lowest observed day of the week, and picks up on Sunday leading up to the next week. This affirms the weekly seasonality shown in the Daily plot.

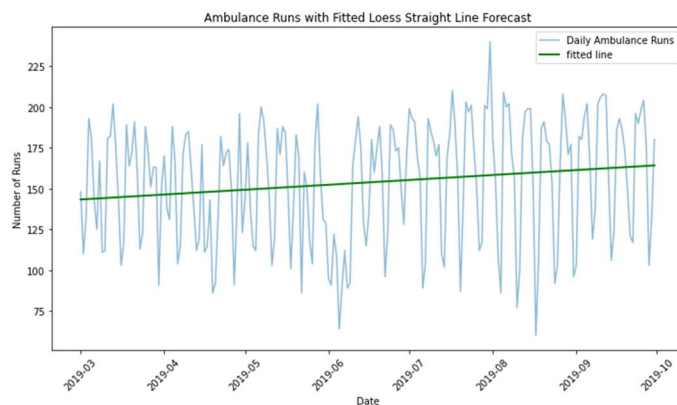
On further observation of the Daily plot shows a dip on April and June, this correlates with two religious holidays, easter holiday and Idul fitri with the largest dip during Idul Fitri and Idul Adha. This relevant information since in Indonesia According to Ministry of Religious Affairs data in 2022, 87.02% of Indonesians identified themselves as Muslim. Another relevant date is during Indonesian Independence Day in August where similar dips are observed. The highest observed date was at the end of July, but there was no clear reason why there was such a high volume of ambulance runs. I was concerned if these outliers are going to affect the forecast.



4. Decomposition



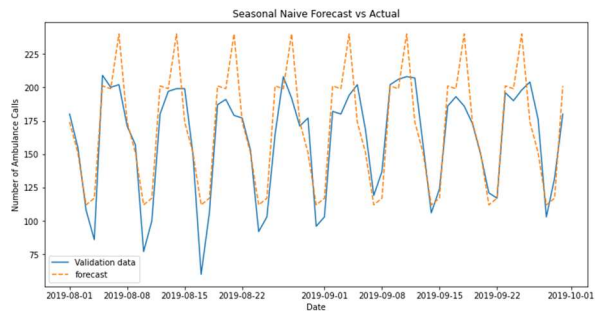
The data was decomposed using STL which showed evidence of a trend with seasonality, and asserts the choice of using Daily data, I was not sure about the trend and fitted a straight line through the data to confirm if the direction of the trend using Trend local regression. It showed a slight upward trend through the data. The residual component shows a systematic fluctuation around 0 and suggest that decomposition may not adequately capture the underlying pattern, differencing should be considered for the data.



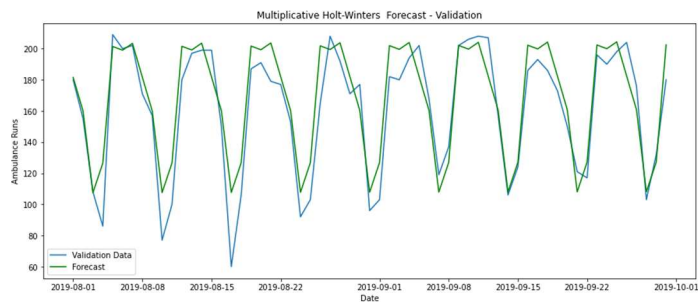
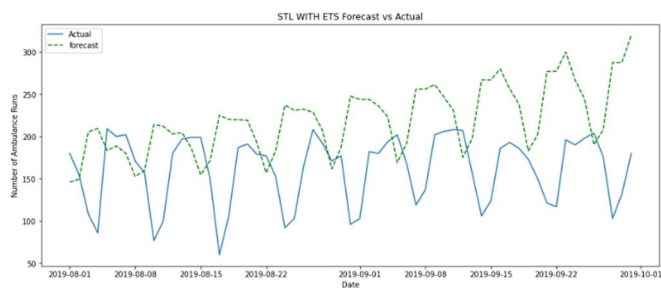
5. Forecasting Models

5.1. Baseline Model

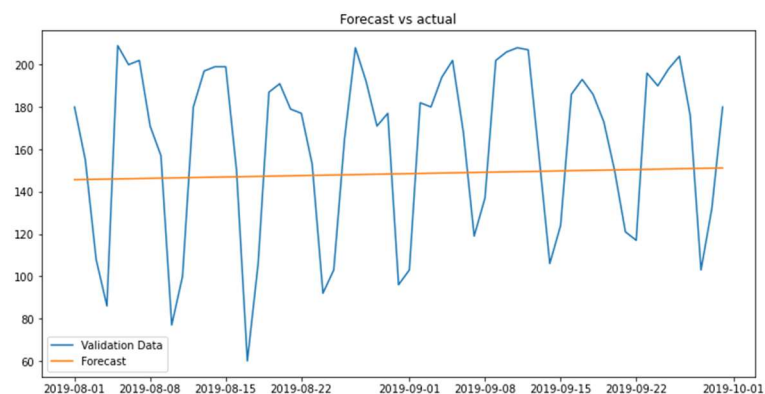
I used a seasonal naïve model as the base line model



5.2. Extrapolation Models: STL + ETS, Multiplicative Holt Winters,

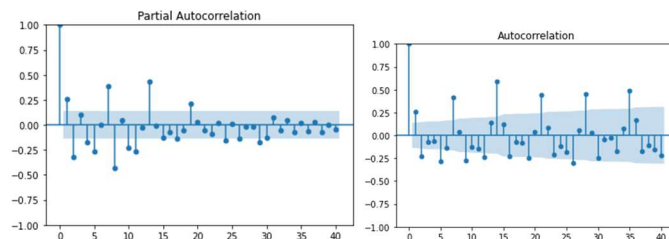


6. Simple liner regression model



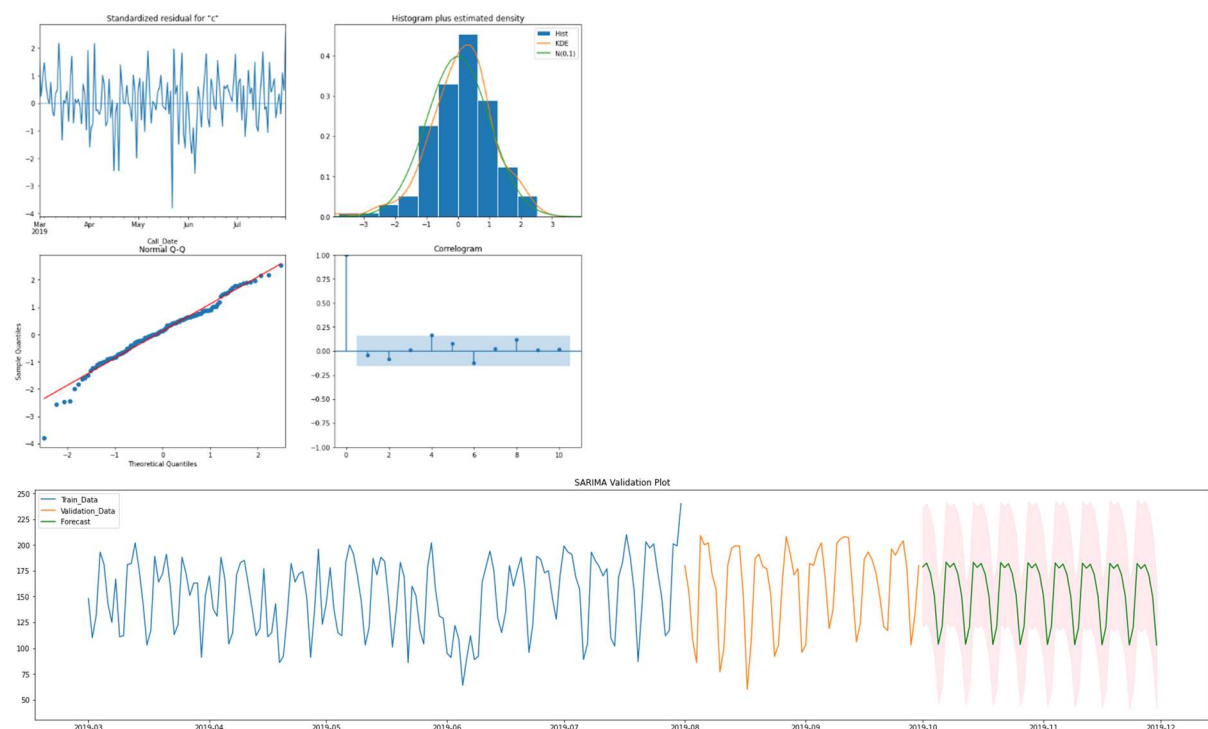
7. Arima

ACF and PACF graphs plotted with the daily runs data, and it shows correlations at different lags. I performed first order differencing coupled with a seasonal 7 differencing on the original data and plotted the ACF AND PACF graphs as shown below but stationarity was not archived, I tried a second order differencing and still could not archive stationarity of the data, I also removed all outliers from the data and the same result was observed. I also tried box-cox transformation on the data and concluded that this is probably because the trend on shows slightly rises all through the data.



I decided to use Seasonal Arima

Various Arima models were fitted, SARIMA (1,0,1), and SARIMA (2,0,1) the best fitting model was SARIMA(2,0,1) and it was used for forecasting.



Ambulance is underutilised in Jakarta; research shows 9.3% of patients go to the hospital using an ambulance. Many people get to hospitals using motorcycles. This is highly attributed to by the high medical bills that the patients get if they use the ambulance, other factors are the 24min response time IQR of 54 minutes, patient awareness of the emergency services and a fatalistic culture which highly impacts healthcare seeking behaviour.

Following the forecast of the data I would advice the ambulance team to expect low volumes of calls during the first day of September which is a Muharram the Islamic New Year and also a weekend, they should expect a steady number of ambulance calls during the two months. Public awareness campaigns on symptoms and cases that warrant use of ambulance are advised to increase the utilisation of these services.