Two Months Forecast of Daily Ambulance Runs

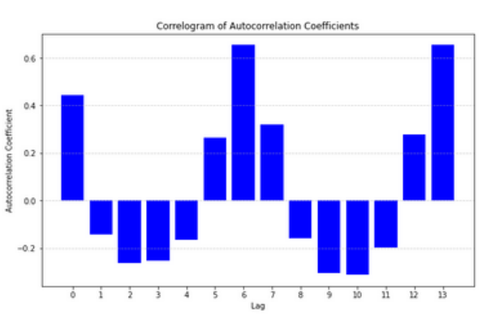
INTRODUCTION

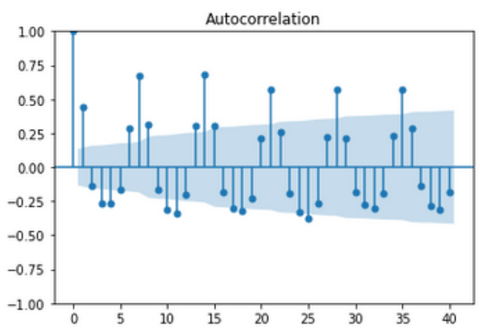
Jakarta Ambulance service Ambulance call data was aggregated Hourly, Daily, and Weekly, and I decided to go with the Daily aggregated data. Data will be split into a training (March – July) and validation (August – September) data set. Time series forecasting models will be used and validated on the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

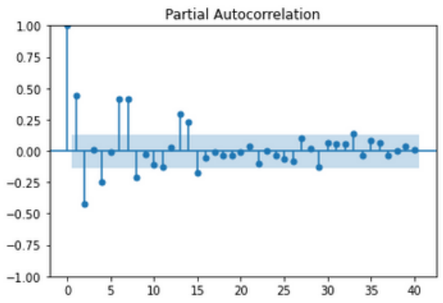
NUMERICAL SUMMARIES

MEAN	Q1	MEDIAN	Q3	IQR	VAR
153.8	119.3	163.5	183.8	64.5	1393

Numerical Analysis of the data shows that it is widely varied around the mean.
A correlogram was plotted for the data and shows a 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 stationality. The ADF stat was -2.132, and the p-value = 0.231, which was not a strong enough rejection of the null hypothesis, and I concluded the data was no-stationery.



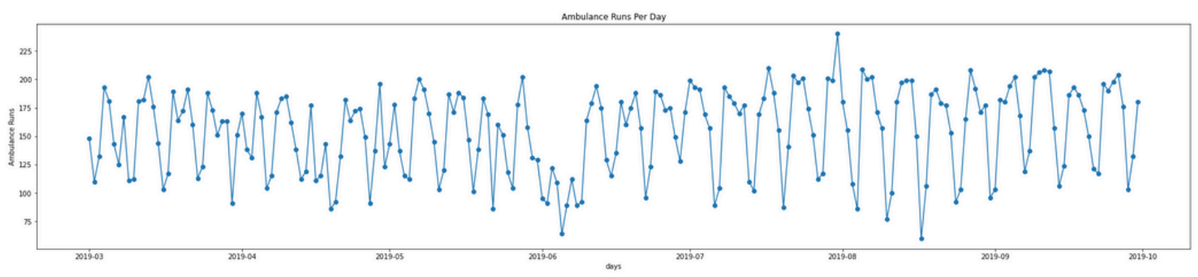


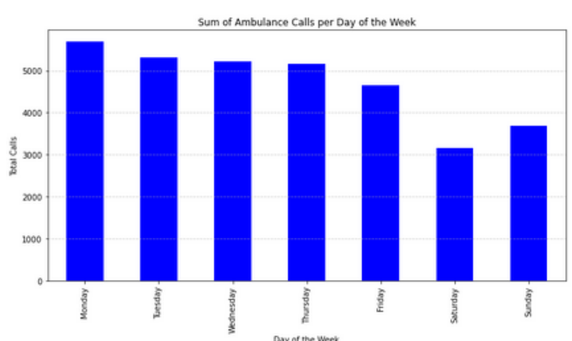


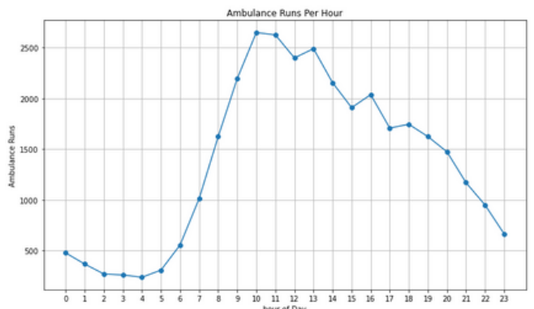
ACF and PACF plots were done on the data, and this shows that the data is not stationary as there is no decay as the lag increases.

Graphical Summaries

The data plot shows a sharp increase in Ambulance Runs from the morning to midday, followed by a steady decline through the afternoon and night. Daily plots of the data show seasonality, with dips during 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 the lowest observed day of the week, and picks up on Sunday leading up to the following 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 Idulfitri with the largest dip during Idul Fitri and Idul Adha. This is relevant information 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 when similar dips are observed. The highest observed date was at the end of July, but there was no apparent reason for the high volume of ambulance runs. I was concerned about whether these outliers would affect the forecast.

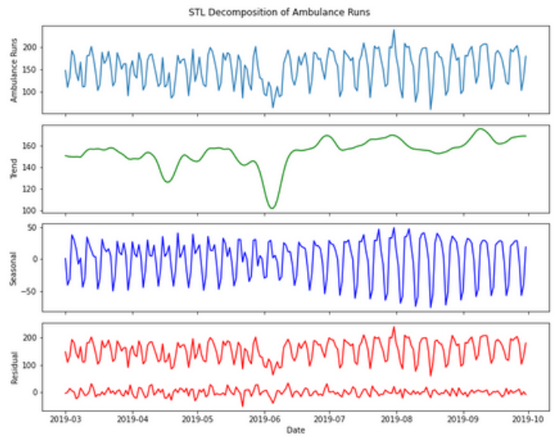


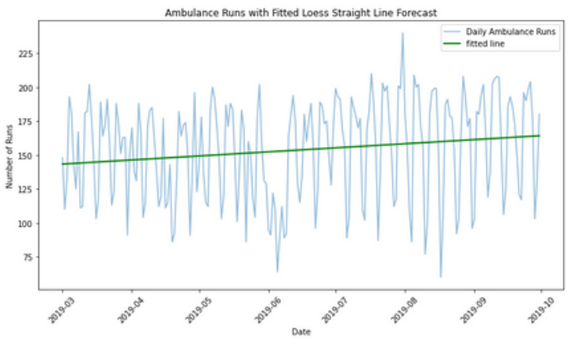




The data was decomposed using STL, which showed evidence of a trend with seasonality and asserted the choice of using Daily data; I was not sure about the trend and fitted a straight line through the data to confirm the direction of the trend using Trend local regression. It showed a slight upward trend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

Decomposition



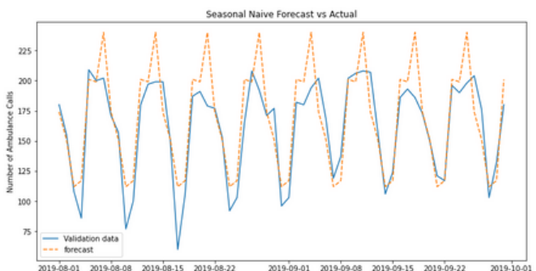


Forecasting Models

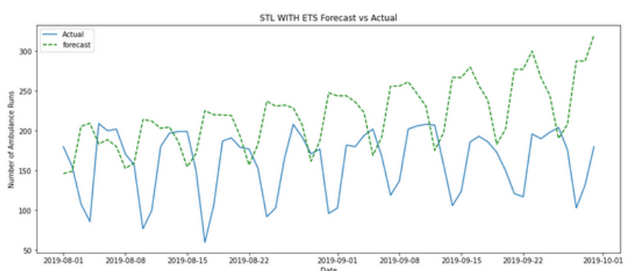
Baseline Model : seasonal naïve model

Extrapolation Models : STL + ETS, Multiplicative Holt Winters,

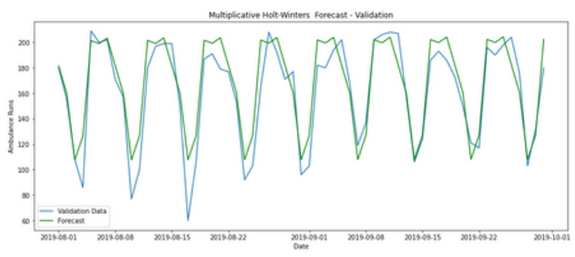
seasonal naïve model



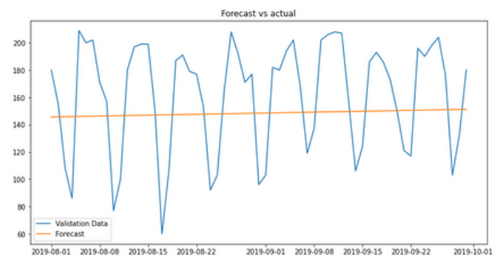
STL + ETS



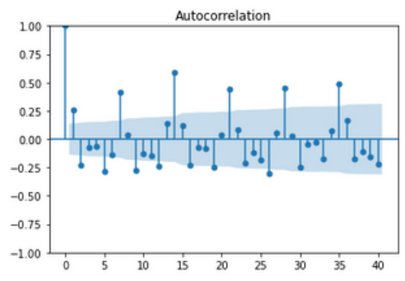
Multiplicative Holt Winters

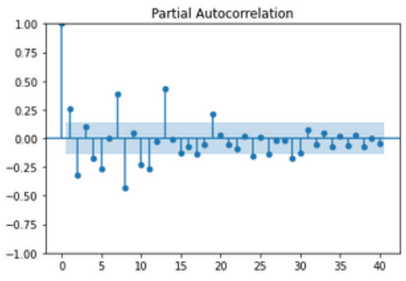


Simple Liner Regression Model



ACF and PACF graphs are plotted with the daily_runs data, showing correlations at different lags. I performed first-order differencing coupled with a seasonal seven differencing on the original data and plotted the ACF AND PACF graphs below, but stationarity was not archived. I tried a second-order differencing and still could not archive the 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 shows slight rises throughout the data.

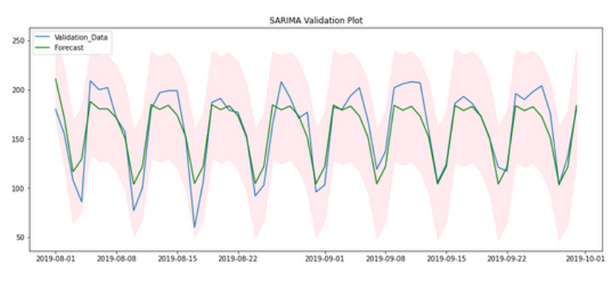


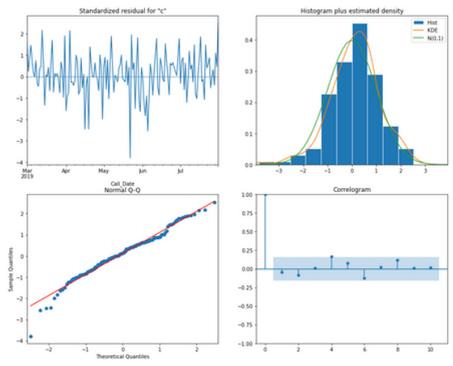


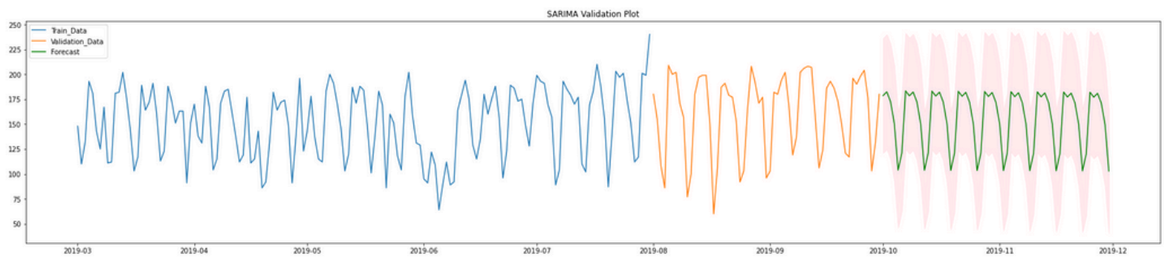
ARIMA

Various Arima models were fitted, including SARIMA(1,0,1) and SARIMA(2,0,1). The best-fitting model was SARIMA(2,0,1), which was used for forecasting.

	√MSE	MAE
Seasonal Naïve	23.3	17
STL -ETS	86.2	64
Mul-Holt Winters	23.86	28
SARIMA (1,0,1)	19.87	18
SARIMA (2,0,1)	17.83	14







Ambulances are underutilized in Jarkarta; research shows 9.3% of patients go to the hospital using an ambulance. A large number of people get to hospitals using motorcycles. This is highly attributed to the high medical bills that the patients get if they use the ambulance. Other factors are the 24-minute response time IQR of 54 minutes, patient awareness of the emergency services, and a fatalistic culture that highly impacts healthcare-seeking behavior.
Following the forecast of the data, I would advise the ambulance team to expect low volumes of calls during the first day of September, which is 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 the use of ambulances are advised to increase the utilization of these services.