FORECAST and FRONTLINE : Predicting Ambulance Demand in Jakarta for Effective Crew Management

Akash Prakash Bhowad - 23081852

INTRODUCTION

In Jakarta, due to the fast-paced nature of city life, there is an acute need for effective ambulance services. This report presents a forecasting tool designed to maximise crew scheduling efficiency for an ambulance service provider situated in Jakarta.

The goal was to predict the number of calls per day for the next two months by using call data from March to September 2019. The strategy used a variety of time-series models, advancing from simpler Naive and Moving Average models to more complex ARIMA and SARIMA models. By precisely predicting ambulance demand, this forecasting tool aims to help efficient emergency management by optimising resource allocation and enhancing response capabilities.

PRELIMINARY ANALYSIS

Mean	Std	25%	50%	75%	MAD
153.76	37.42	119.25	163.5	183.75	32.43

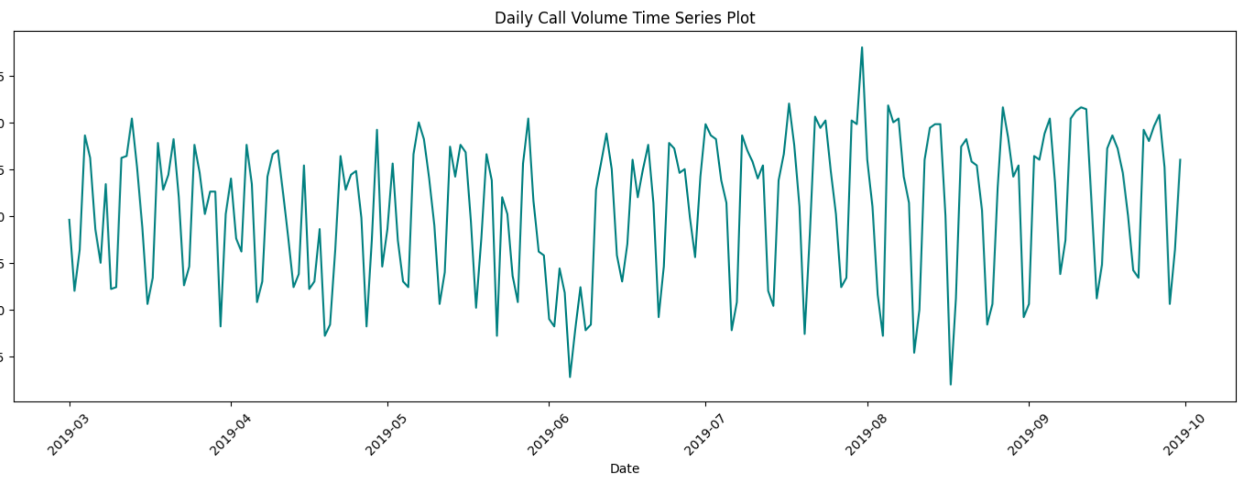
With a standard deviation of **37.42**, the daily call volumes show moderate variability. MAD indicates that on average, the daily call volumes deviate from the mean by about **32** calls.

Total Days	Total Calls	Min Calls	Max Calls	Avg Calls	Active Calling Hour
214	32905	60	240	153	10:00 AM

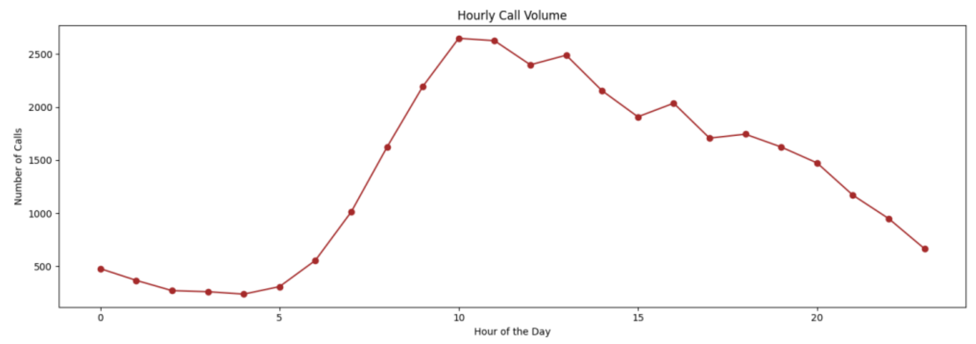
Over the period of **214** days, **32,905** ambulance calls were received in total, indicating a strong demand for emergency services.

The day with the least demand saw **60** calls. Whereas, **240** calls were received on the busiest day, which may have put a strain on the resources.

On an average, there were **153** calls per day, and **10:00 AM** is seen to be the busiest time for calls.

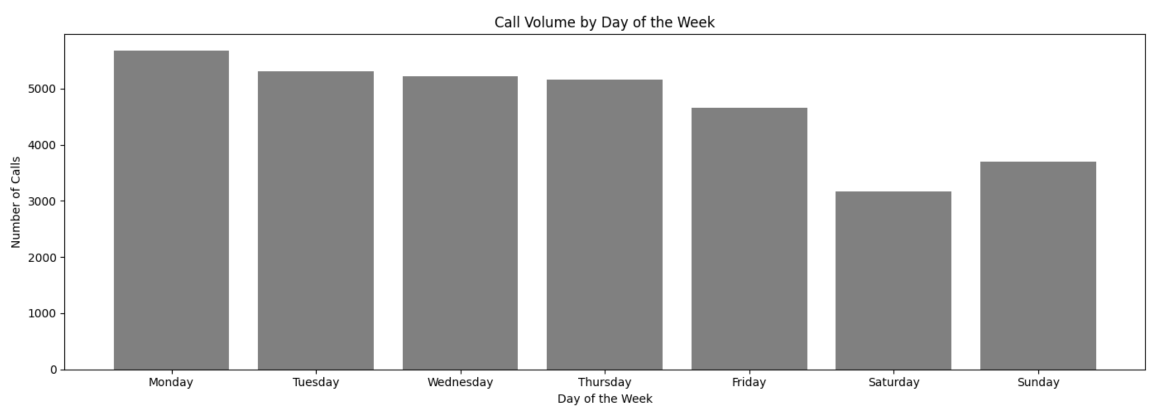


The time series chart of emergency calls indicates fluctuations with a regular pattern, suggesting weekly cycles

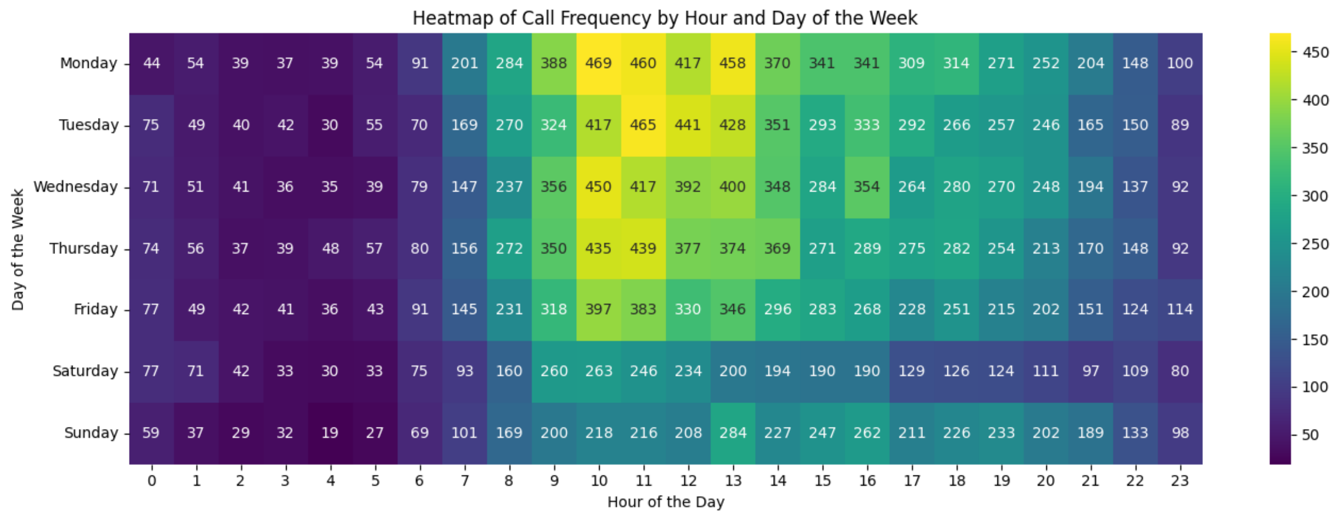


Between midnight and five in the morning, there are comparatively fewer calls. The call traffic increases significantly and steadily starting at roughly **5 AM** and peaks around **10 AM.**

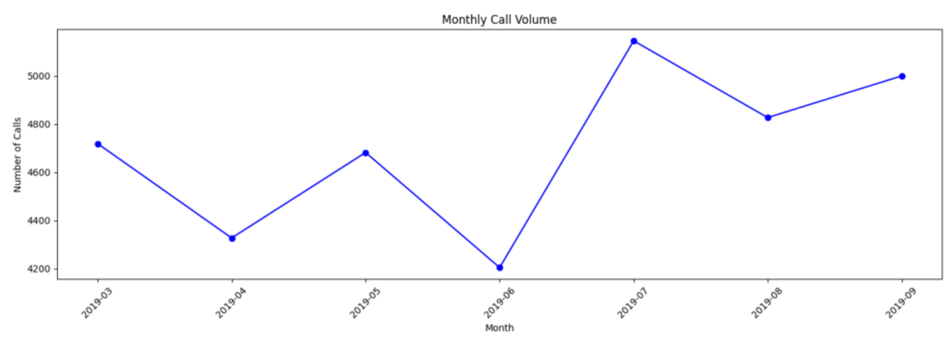
At around 10 AM and 1 PM, there is a plateau in call volume. Additionally, following the plateau, the number of calls gradually drops off and stays that way until the evening.

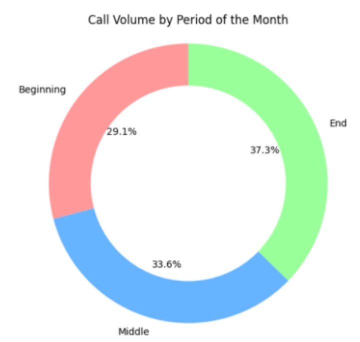


The busiest day of the week is Monday while Saturday being the least.



The heatmap illustrates there's a surge in the ambulance demand during the midday hours of **10 AM to 3 PM on weekdays**, while on weekend there's are notably reduced call volumes.

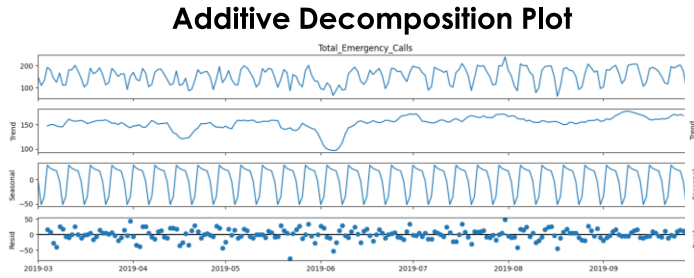


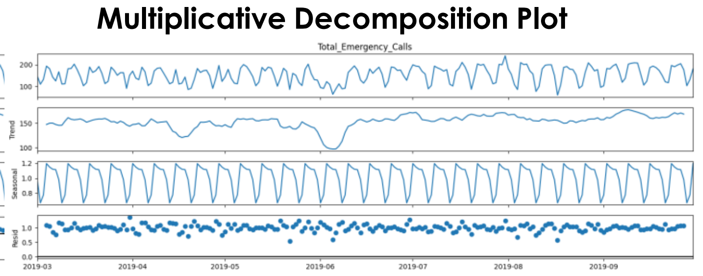


July saw the greatest number of calls (**5146**), while June saw the fewest (**4,204**). Overall, the initial days of the month account for **29.1%** of the calls. The call volumes rise marginally (**33.6%**) in the middle of the month, and peak at **37.3%** towards the month's end.

DECOMPOSITION

Additive Decomposition Plot **Multiplicative Decomposition Plot**



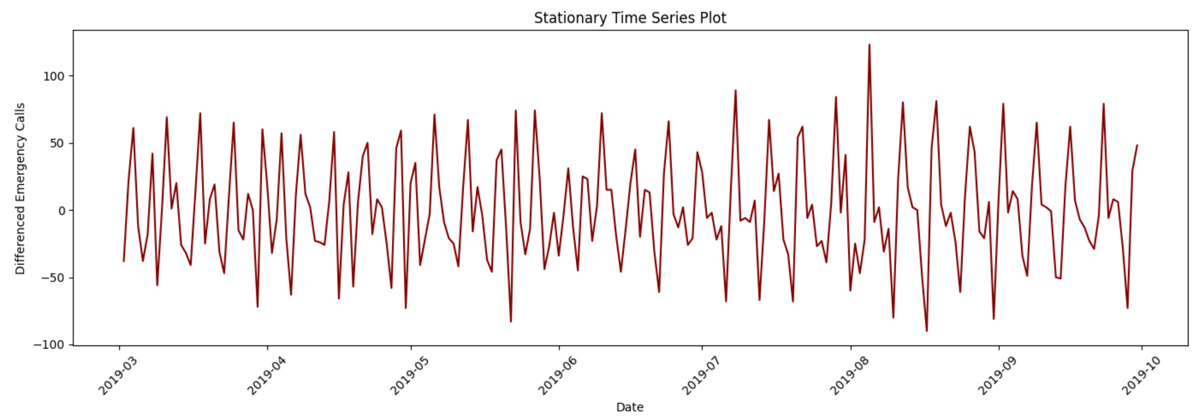


On looking at the Additive Decomposition Plot, the call volume seems to have a stable trend. There is a considerable seasonal effect, and it appears to be consistent with regular fluctuations. Moreover, the residuals indicate randomness without patterns, indicating a good fit.

Initial Stationarity Checks

Test	Test Statistic	p-value	Critical val. (1%)	Critical val. (2.5%)	Critical val. (5%)	Critical val. (10%)
ADF	-2.13	0.23	-3.46	-	-2.87	-2.57
KPSS	0.62	0.01	0.73	0.57	0.46	0.34

ADF and KPSS tests were carried out to check if the data is stationary. The initial tests revealed the data isn't stationary and there is presence of unit root, hence, to achieve the stationarity, First Order Differencing was performed.



Stationarity Checks after First Order Differencing

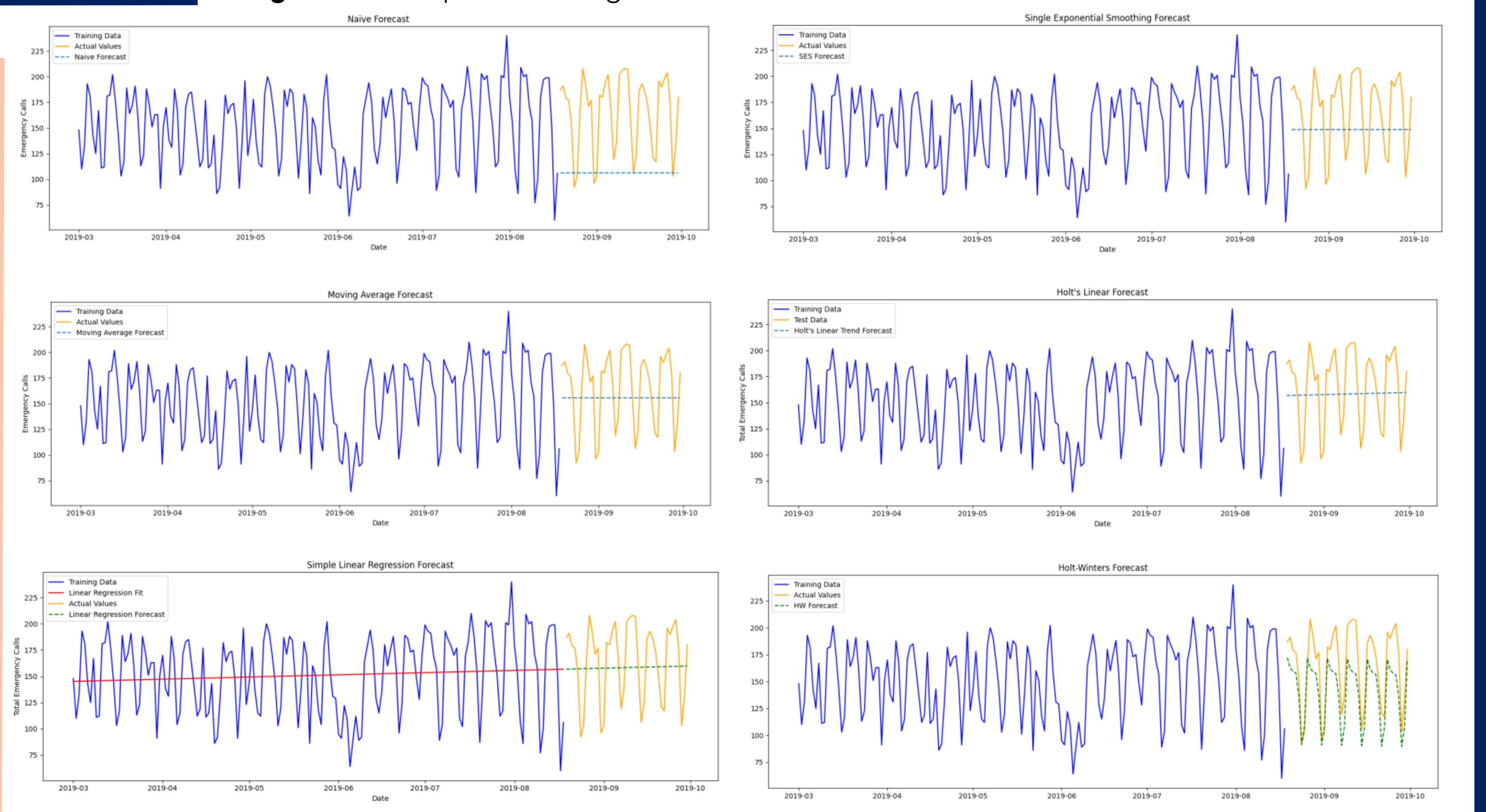
Test	Test Statistic	p-value	Critical val. (1%)	Critical val. (2.5%)	Critical val. (5%)	Critical val. (10%)
ADF	-6.27E+00	3.94E-08	-3.46E+00	-	-2.88E+00	-2.57E+00
KPSS	0.5	0.04	0.73	0.57	0.46	0.34

On running the stationarity tests on the differenced data, the results from ADF and KPSS tests reveal that the series is now stationary.

BASELINE, EXTRAPOLATION & REGRESSION MODELS

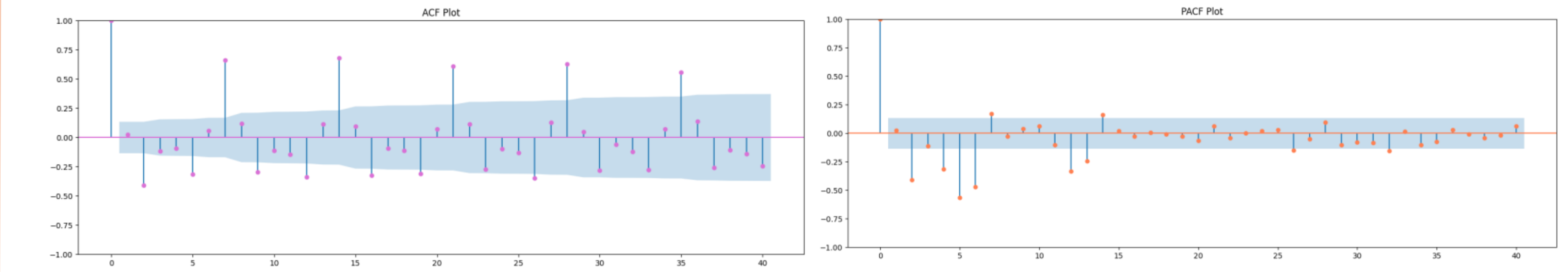
To capture the dynamics of the call volumes, a comprehensive range of time series forecasting models (mentioned-below) were implemented on the initialisation set (80%) and evaluated on the test set(20%).

Baseline – Naïve Model
Averaging – Moving Average
Extrapolation – Single Exponential Smoothing, Holt's Linear, Holt Winters
Regression – Simple Linear Regression

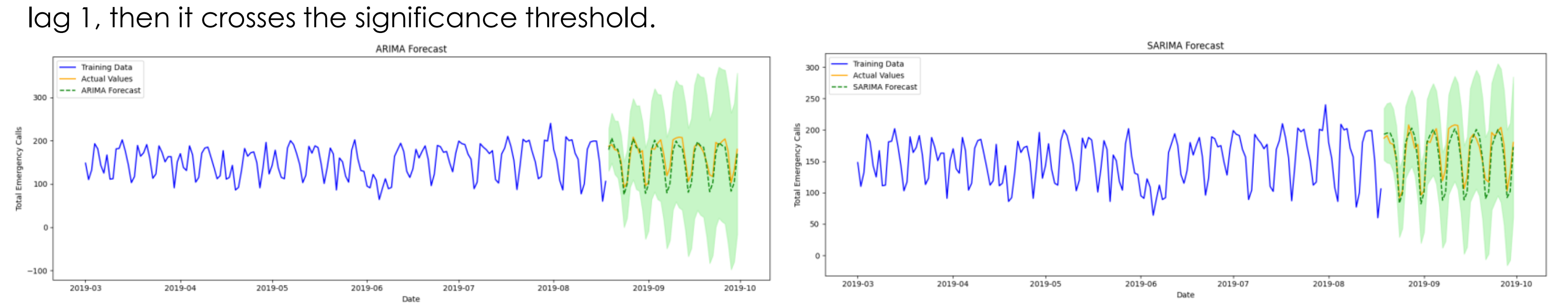


The forecasting began with the baseline Naïve model, which yielded high error margins, implying its limitations to adjust to the fluctuations. Simple Linear Regression and Moving Average Models captured trends more effectively. Whereas it was observed that Holt's Linear and Holt-Winters models provided improved predictions by adjusting for trend and seasonality. However, the data is non-stationary and to capture complex patterns required advancing to the ARIMA and SARIMA models.

ARIMA MODELS & ACF, PACF ANALYSIS



The ACF plot shows spikes at regular intervals indicating seasonality in the data. The initial ADF tests confirmed non-stationarity and the same was confirmed by the ACF plot as slow decaying of correlations are observed. First order differencing was performed to attain stationary data. In the PACF plot, a significant spike appears at lag 1, then it crosses the significance threshold.

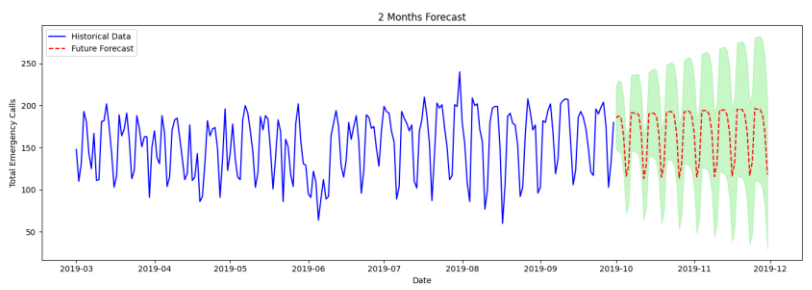


ERROR STATISTICS

Model	Naïve	Moving Avg.	Linear Reg.	Single Exp.	Holt's Linear	Holt Winters	ARIMA	SARIMA
MSE	4757	1385	1331	1543	1331	782	369	202
MAPE	33.36	22.42	22.3	23.13	22.3	14.31	10.73	7.15

The ARIMA and SARIMA models seems to be better capturing the nuances in the data, however SARIMA's improved error statistics, MSE(**202**) & MAPE (**7.15%**), indicates it well handles both the trend and seasonality.

FUTURE FORECASTS



To predict the number of calls for next two months, **SARIMA** model was chosen since it effectively captures underlying trend and seasonal fluctuations.

CONCLUSION & INSIGHTS

With thorough analysis, the complexities of predicting emergency call volumes have been examined. Trend and Seasonality patterns were found in the data and was validated by the decomposition plots and stationarity tests. The presence of the trend and seasonality supported the need for sophisticated modelling. Significant improvements were observed in the forecast accuracy on implementation of the ARIMA and SARIMA models, as evidenced by lower error metrics. SARIMA model proved to be good at capturing the data's seasonality, indicating that it may be used for efficient operational planning. These observations are critical for ambulance services to maximise resource allocation, accelerate response times, and ultimately raise the standard of emergency care.