Forecasting Of Ambulance Needs In Jakarta, Indonesia.

MAT005 - TIME SERIES AND FORECASTING COURSEWORK 2023

JEYAPRAKASH RAMACHANDRAN
22081255

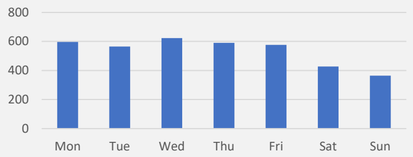
Problem Statement

Jakarta's ambulance service (118) seeks a planning tool to help them schedule and roster their teams. Our task is to forecast the weekly call volume (Monday through Sunday) for the next eight weeks, from October 24 to December 18, 22. MAPE of Linear and ARIMA models are used to forecast the call volume.

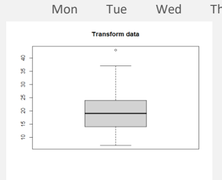
Exploratory Data Analysis

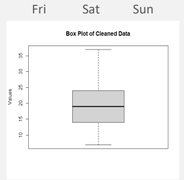
R and MS Excel were used to investigate, assess, and execute forecasting models. The average call is 12/week, maximum of 36, and minimum of 0 calls The call data was cleaned and analyzed to determine the following.

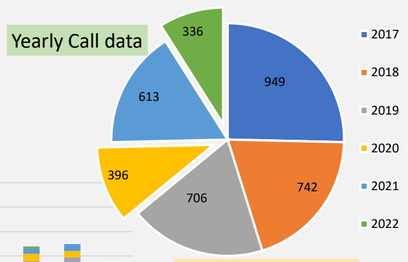
- ❑ Weekly call data
- ❑ Monthly and yearly call data
- ❑ Time of call received

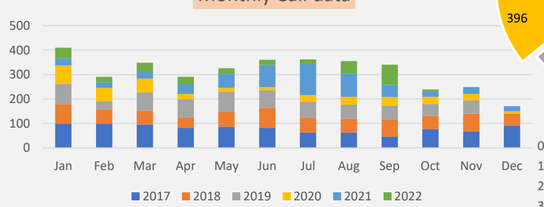


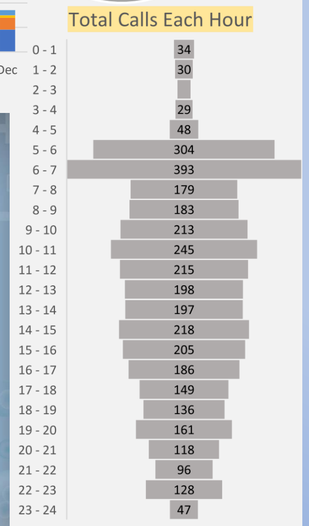
Statistics_of_Data	Value
Mean	12.39
Minimum	0.00
Maximum	36.00
Range	36.00
IQR	10.00
Q1	7.00
Q2	12.00
Q3	17.00
Variance	50.82
Standard Deviation	7.13
Coefficient of Variation	0.58
Mean Absolute Deviation	5.88
Median Absolute Deviation	5.00







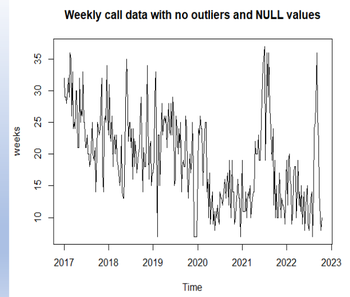




Additionally, the data suggests that there may be a correlation between the time of day and the type of calls received and a possible correlation between the season and the frequency of calls. Overall, these insights could be used to improve call Centre operations and allocate resources better. There are 3742 recorded in the given time frame. Out of it, 2306 calls were received between 8:00 AM to 8:00 PM. But a high volume of calls was recorded between 5 -7 in the morning.

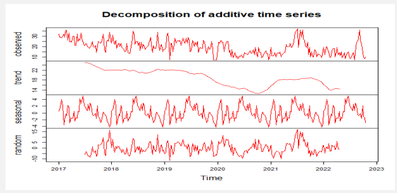
Data Cleaning

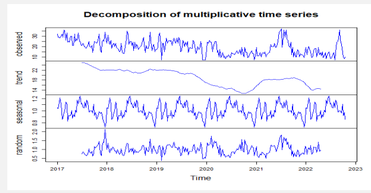
Removing outliers will eliminate any extreme values that may skew the results while imputing missing values will provide a more complete dataset for analysis. The time map reveals numerous abrupt shifts, notably the high call volume during 2021/2022. These changes are due to the Covid19.



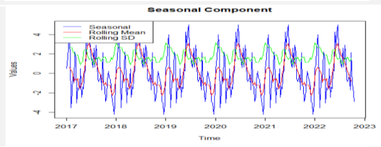
Time Series Components:

Additive and **multiplicative** decomposition finds trend and seasonality in data. There is seasonality and trend in the data.



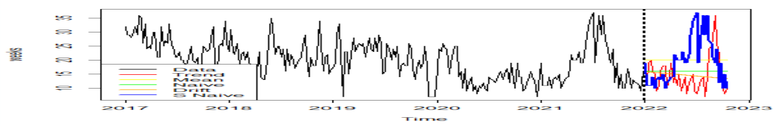


The same has been verified with the **rolling mean** and **SD** methods used. It is evident from the figure.



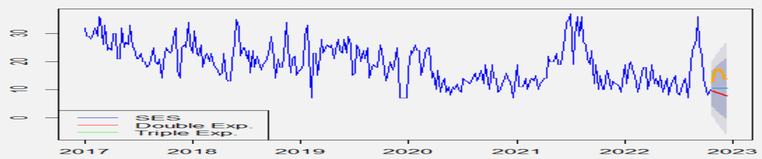
Baseline Model

We have created 4 baseline models, The RMSE on test data by models are Mean model – 7.92, **Naïve Model – 6.19,** Drift Model – 6.24, and Seasonal Naïve – 11.39. The Naïve model has a low RMSE, and the predicted data is according to the previous data.



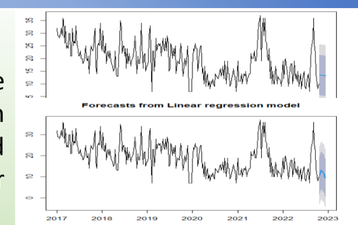
Extrapolation Model

Based on the user's input, it can be concluded that the SES model and the hot linear and exponential models have resulted in a superior forecasting model. The SES model has a low RMSE of 4.94, indicating that it accurately predicted the forecasted values.



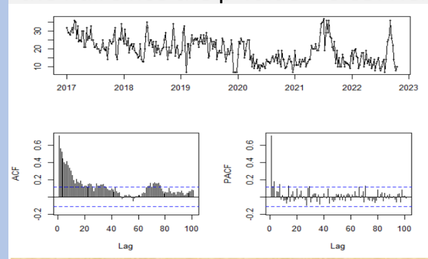
Simple Linear Regression Model

The first model represents only the trend, and the second is fitted with trend and seasonality. The second model better fits data with lower RMSE(5.753) and MAE(4.522) errors.



Dickey-Fuller Test

The test statistic value is -3.9006, and the p-value is 0.01437. reject the null since the p-value < 0.05. Hence, the time series is stationary.

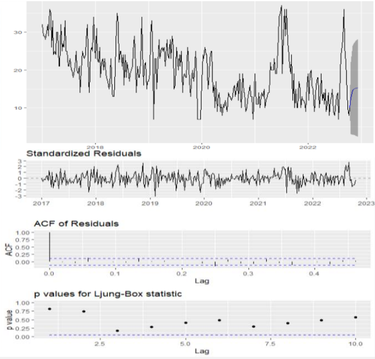


Autocorrelation

By examining the ACF and PACF plots, we can determine the appropriate p and q for an ARIMA model. The data was initially split into 70-30% train-test for the analysis.

ARIMA Model

The ARIMA(1,1,1) model has two coefficients, an AR coefficient of 0.50 and an MA coefficient of -0.89. The standard errors are 0.10 and 0.06, and the variance is 23.2. The log-likelihood, AIC, AICc, and BIC values are -899.63, 1805.26, 1805.35, and 1816.39. The ARIMA model is a good fit.



Model	Naïve	SES	Holt Linear	Holt Winter	imple Linea	ARIMA
MPE	-23.03	-6.84	-4.12	-7.39	-14.96	-8.02
MAPE	38.41	23.00	23.35	30.83	35.02	22.95
RMSE	6.19	4.95	5.14	6.35	6.86	4.79
PREDICTIONS						
Week 1	3.00	3.49	2.93	5.56	6.45	5.64
Week 2	3.00	3.49	2.85	9.22	6.41	6.96
Week 3	3.00	3.49	2.78	10.35	6.37	7.62
Week 4	3.00	3.49	2.71	9.62	6.33	7.95
Week 5	3.00	3.49	2.63	10.41	6.29	8.12
Week 6	3.00	3.49	2.56	8.67	6.25	8.20
Week 7	3.00	3.49	2.49	6.79	6.22	8.25
Week 8	3.00	3.49	2.42	6.92	6.18	8.27

ERROR Analysis

The Naive model has the highest MPE and MAPE values, Holt Winter has the highest RMSE indicating Predictions have higher variability than actual values.

Summary

The ARIMA model is the best for forecasting the data based on RMSE and the MAPE.

The ARIMA method is the best-performing model for this data, with good predicted values for all weeks. However, changes in seasonality, trends, and unexpected events can affect the model's accuracy. Suggest implementing ARIMA for weekly call prediction and resource allocation.