Data Analytics Assignment (EB5102)

Time Series Forecasting

Submitted by:

- Aravind Somasundaram (A0163301X)
- Kavya AK (A0163250R)
- Praman Shukla (A0163239A)
- Saravanan Kalastha Sekar (A0163309H)

Introduction

Time series forecasting is a process where a model is built to predict the future values based on previously observed values. It focuses on comparing values of a single time series or multiple dependent time series at different points in time. In this workshop, a model is built to predict the future **GRP (Gross Rating Point)** based on the previous GRP. To implement the same, **4 different time series methods** were used to build models, which are, Decomposition, Exponential smoothing, ARIMA and Time series regression.

Dataset Description

The dataset consists of weekly time series data of GRP ratings for an Indian TV channel for the duration of two years between Jun 2007 and 15th Mar 2009. In dataset, the two variables were "Week" and "GRP". Week variable consists of the number of weeks for the given time and GRP variable consists of ratings for that week. The dataset was split into training and test data. Training set was taken from 17th June 2007 up to 28th Dec 2008 which was used to train the model whereas the test set from 4th Jan 2009 to 15th March 2009 was used to predict the accuracy of the model built.

Tools Used

Methods	**Tool**
Exponential smoothing	R
ARIMA	JMP
Decomposition	Excel
Time Series Regression	Excel

Table 1

Forecast Methods

Decomposition Model

The decomposition of time series is a statistical method that deconstructs a time series into several components, each representing one of the underlying categories of patterns. It is used in time series to describe the trend and seasonal factors in a time series. In this case, a multiplicative decomposition model was built in the Excel tool. The multiplicative decomposition model can be expressed by the following equation:

$$O_t = TC_t \times S_t \times I_t\,.$$

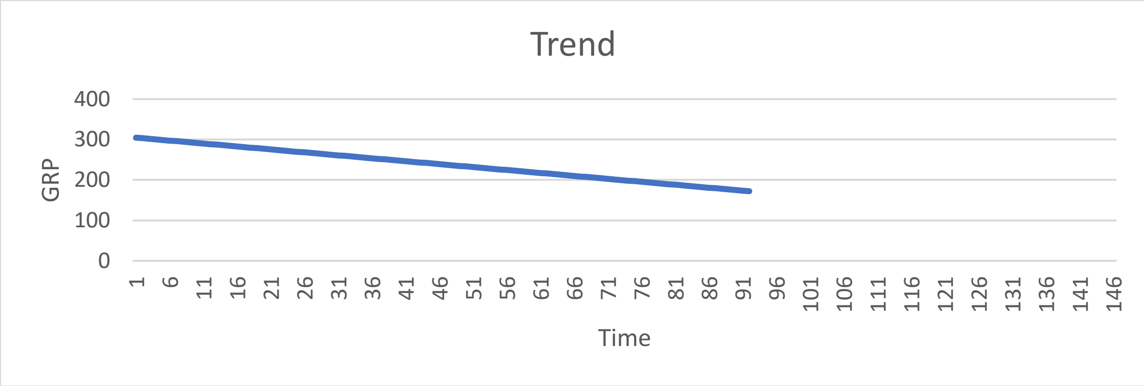

Figure 1

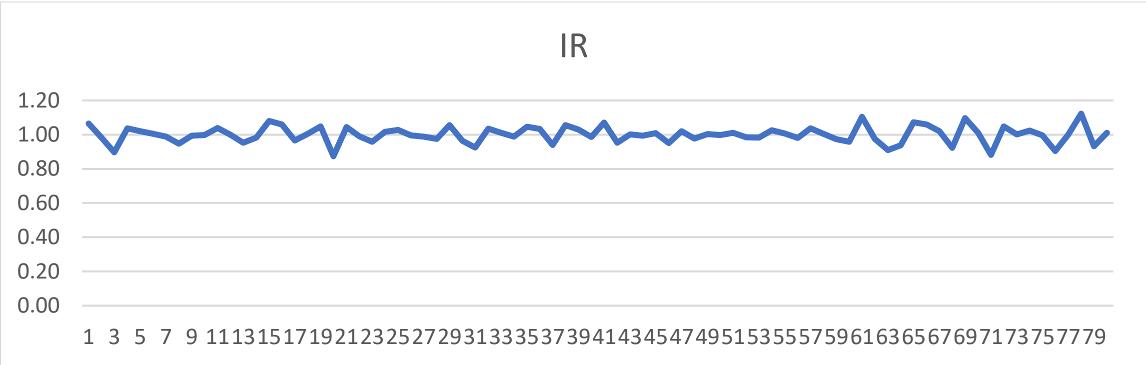

Figure 2

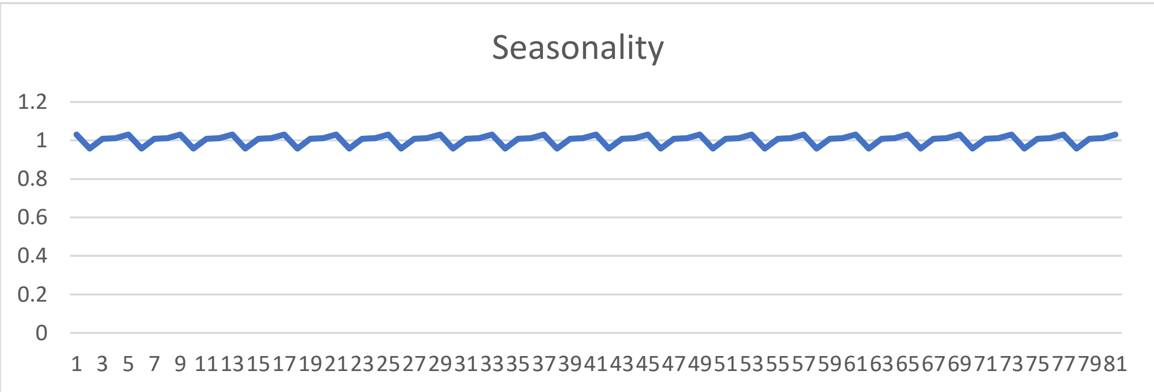

Figure 3

From Figure 1, we can see that the data has a downward trend.

From Figure 2, we can infer that no significant changes in the values of Average (Sn) were seen over a period i.e. all values were nearly 1 for all periodic cycles.

From Figure 3, as for the IR component, it doesn't show any kind of pattern.

Hence, to develop the predictive model we assume the IR and Avg (Sn) value as 1 i.e. considering only the trend component. The forecasted values are then compared with the actual values to determine the accuracy of the model by calculating the Mean Absolute Percentage Error which is given by the formula,

$$\text{MAPE} = \frac{\sum_{t=1}^{n} \left| (A_t - F_t)/A_t \right|}{n}$$

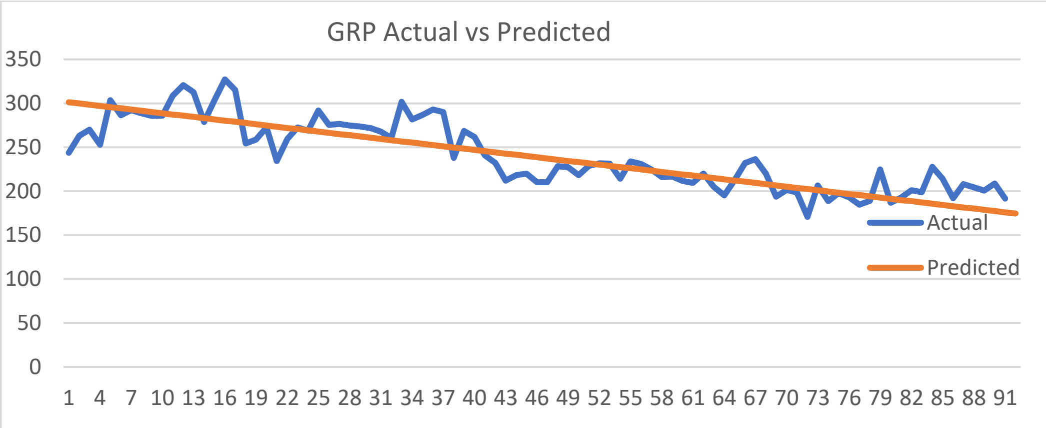

Figure 4

The above graphical representation (Figure 4) of the decomposition model forecasting shows us a flat prediction as well as under prediction of the actual values. A MAPE of **10.30%** has been calculated. Since there is no seasonality and there is not enough data, we consider that this method is not appropriate for the forecasting of this data set.

Exponential Smoothing

Exponential Smoothing is an effective method when the components of the time series may be changing over time. Forecasts produced using exponential smoothing methods are weighted averages of past observations, with the weights decaying exponentially as the observations get older.

As the single exponential smoothing method cannot be implemented due to the presence of trend, Holt's double exponential linear trend method has been performed in R Studio. This method involves a forecast equation and two smoothing equations (one for the level and one for the trend):

- Forecast equation $y' = \ell_t + hb_t$
- Level equation $\ell_t = \alpha y_t + (1-\alpha)\ (\ell_t - 1 + b_t - 1)$
- Trend equation $b_t = \beta*(\ell_t - \ell_t - 1) + (1-\beta*)\ bt-1$

where ℓ_t denotes an estimate of the level of the series at time t, b_t denotes an estimate of the trend (slope) of the series at time t, α is the smoothing parameter for the level, $0 \le \alpha \le 1$ $0 \le \alpha \le 1$ and $\beta*$ is the smoothing parameter for the trend, $0 \le \beta* \le 1$.

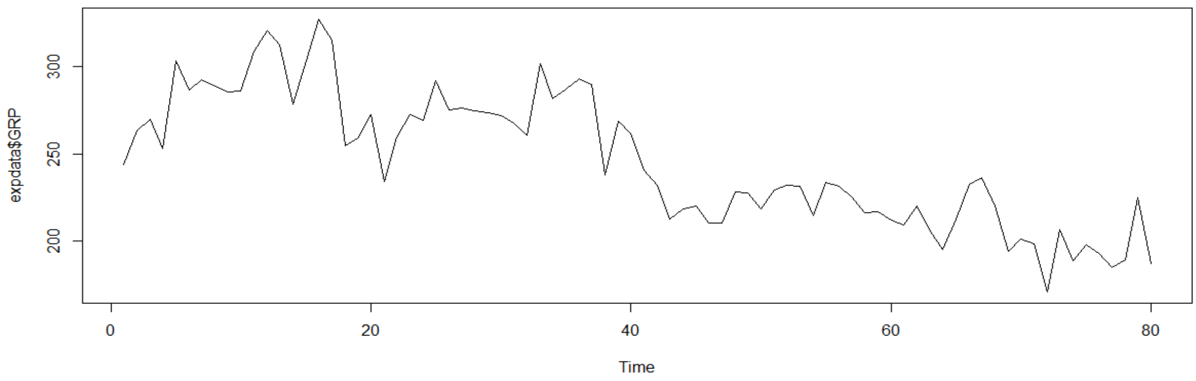

Figure 5

Figure 5 is a graphical representation of the actual GRP values. The training data (16-Jun-07 to 28-Dec-08) has been used to build the predictive model in R with Holts linear function where we let the function estimate the preferred alpha and beta parameters for the given dataset. Figure 6 shows the alpha, beta value and other statistics.

```
  Smoothing parameters:
    alpha = 0.537
    beta  = 1e-04

  Initial states:
    l = 255.571
    b = -0.7622

  sigma:  17.6573

     AIC     AICc      BIC
831.0766 831.8766 843.0488

Error measures:
                       ME    RMSE      MAE        MPE     MAPE      MASE       ACF1
Training set -0.005613243 17.6573 12.59495 -0.4211145 5.211311 0.8921438 0.04062075
```

Figure 6

The model predicted the values for the entire time and forecasted the values for the upcoming weeks of data for the test set which has been highlighted in blue colour in Figure 7. The two red lines denote the upper and lower bounds for the forecasted values. These forecasted values were observed to be in par with the actual values. The MAPE for the exponential predicting model was **5.21131%**

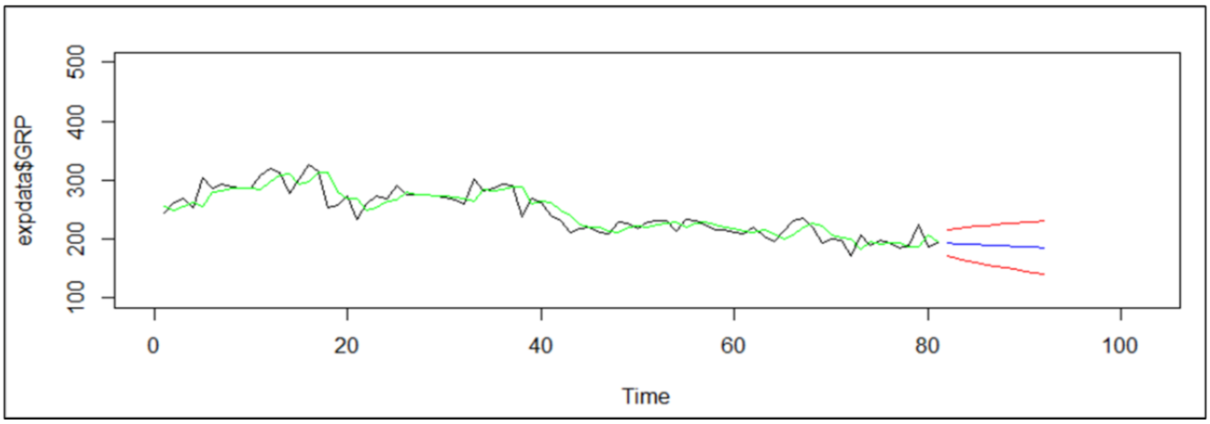

Figure 7

The actual and predicted values of the forecast period were differenced to identify the MAPE of the predictive model. The results indicated that the model prediction for the test data consisted of an error of **7.950%.** And to further reduce the value of MAPE and for accurate measures, ARIMA modelling was performed. Refer Figure 8 for model summary.

```
summary(MAPE)
 Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.9451  3.5590  7.6530  7.9500 10.6000 16.4600
```

Figure 8

ARIMA (Auto Regressive Integrated Moving Average)

An autoregressive integrated moving average (ARIMA) model is fitted to time series data to predict future points in the series (forecasting). It aims to describe the autocorrelations in the data and applied in cases where data show evidence of non-stationarity. ARIMA has been performed in JMP tool.

The time series actual GRP in Figure 9 shows that the model contains a trend and it requires differencing as its ADF values does not meet the required Augmented Dickey Fuller criterion (ADF < -4).

.

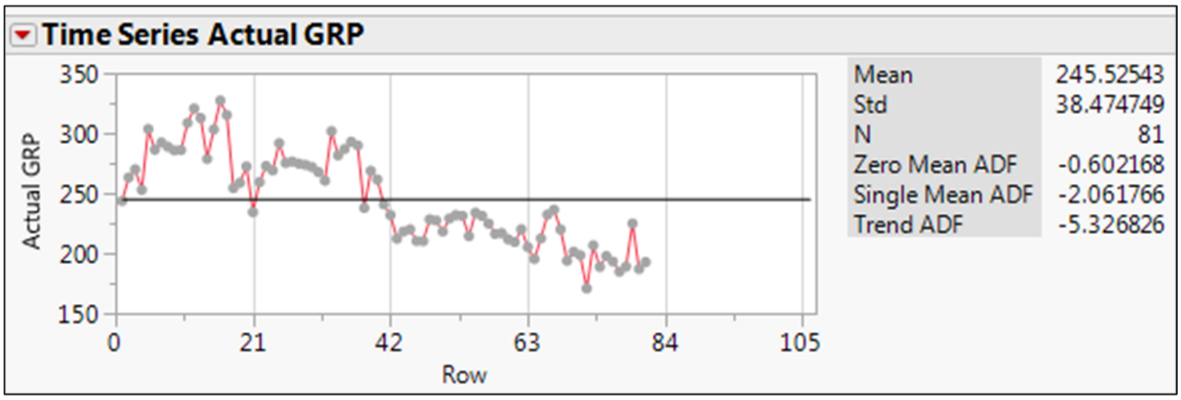

Figure 9

The forecast model is built with the training data and the suitable ARIMA model is recognized by comparing the different possible models. Refer to Figure 10 for the summary of model.

The **ARIMA (3,1,1)** is established to be the perfectly fit model with a least percentage of error (MAPE). The model (3,1,1) also proved that it is suitable for forecasting than the other models since it's the only model which captures the crest and trough of the actual variations. Thus, we determined forecasting values with this model.

Model: ARIMA(3, 1, 1)

Model Summary

DF	75	Stable	Yes
Sum of Squared Errors	22570.2093	Invertible	Yes
Variance Estimate	300.936123		
Standard Deviation	17.3475106		
Akaike's 'A' Information Criterion	690.927182		
Schwarz's Bayesian Criterion	702.837316		
RSquare	0.80467908		
RSquare Adj	0.79426196		
MAPE	5.13085522		
MAE	12.3966835		
-2LogLikelihood	680.927182		

Figure 10

Initially, the model was built with the training data and then test data values were being forecasted. The actual vs predicted, Figure 11 reveals the best fit model are identical to actual values. MAPE is estimated to be **5.130%** for the prediction model and the MAPE for the forecast model is **6.4043%** for forecasting values.

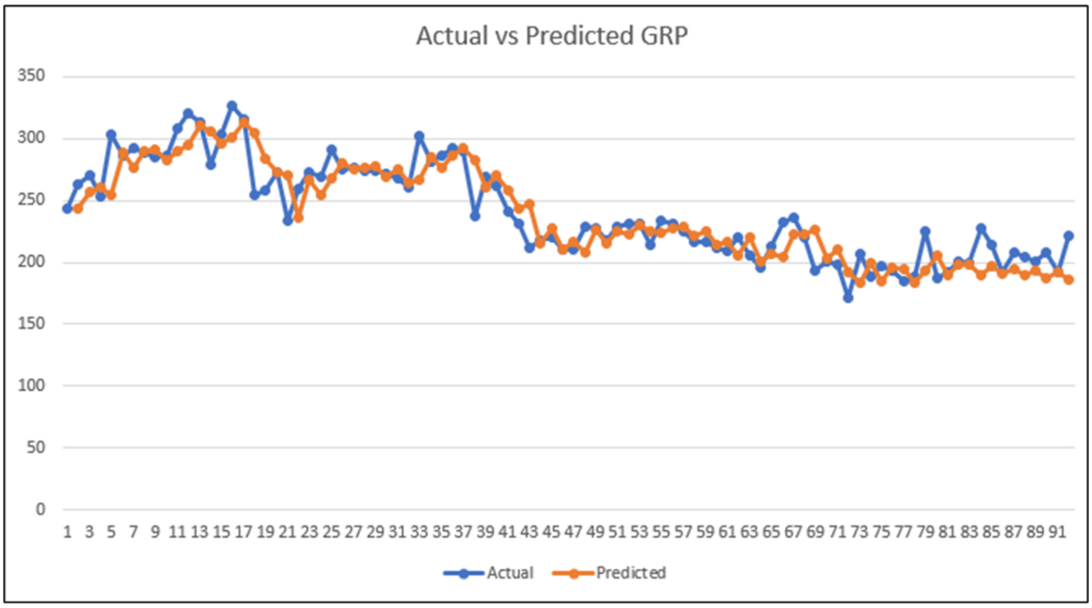

Figure 11

Time Series Regression:

Time series regression is a statistical method for predicting the future response based on the response history and the transfer of dynamics from relevant predictors. It predicts the behaviour of dynamic systems from experimental or observational data.

A Time Series Regression was performed on the given dataset to identify the expected values for the GRP rating in Excel tool. As the given data set consisted only of linear trend along with no seasonality and no pattern in IR, the equation is

Predicted $Y_t = TR_t + \varepsilon_t$

Where,

$TR_t = \beta 0 + \beta 1_t$

Y_t= value of the time series in period t

TR_t= the trend at time t

ε_t = the error term in time t

A time series regression model was built in R using the ‘prophet’ package launched by Facebook, which implements a procedure for forecasting time series data based on an additive model where non-linear trends are fit with yearly and weekly seasonality, plus holidays. It works best with daily periodicity data with at least one year of historical data. Prophet is robust to missing data, shifts in the trend, and large outliers.

The data underwent a logarithmic transformation with an assumption of producing better result than with just the actual data but the model built seemed to have a very small MAPE as shown in Figure 12, wherein the values were completely over predicted and hence there was a need for a change in tool and therefore it was finally performed in Excel.

```
MAPE <- abs((GRP_test1-cut))/GRP_test1*100
summary(MAPE)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.2633  1.5769  2.6609  2.3682  3.2432  3.7371
```

Figure 12

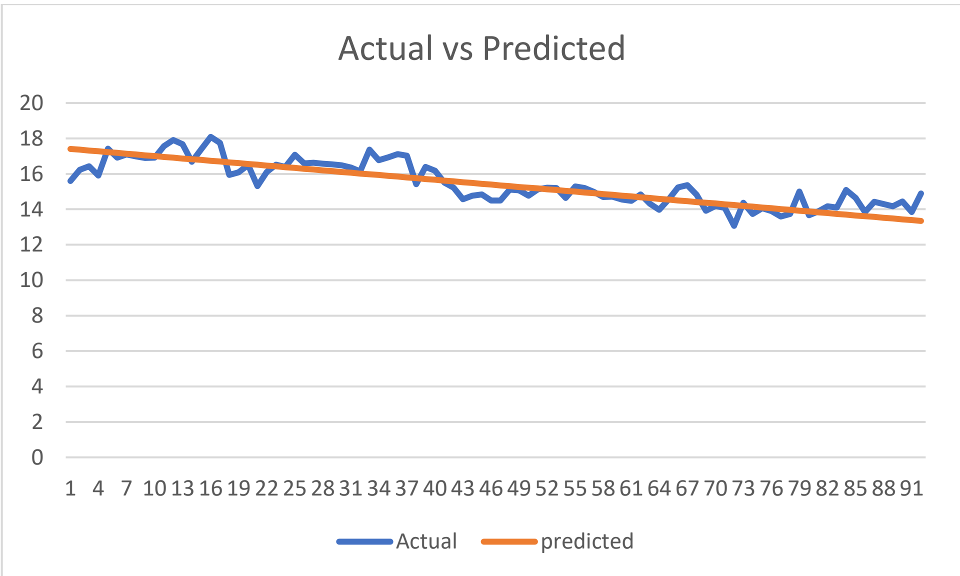

Figure 13

The model was built in excel and from Figure 13, we considered this as a multiplicative model and to reduce the effect, we brought in a dampening factor of 0.5 in the regression variable. With that value, the regression had been carried out to know the parameter estimates of trend variable and intercept.

	Coefficient	tandard Err	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	17.45254	0.146439	119.1794	6.23E-91	17.16106	17.74402	17.16106	17.74402
X Variable 1	-0.04467	0.003103	-14.3971	8.57E-24	-0.05084	-0.03849	-0.05084	-0.03849

Figure 14

X Variable 1 – The trend variable coefficient (β1) as shown in Figure 14

The forecast model was built, and the prediction was observed to be a flat one and the derived values which were under-predicting were compared to the actual values. Though MAPE was as less as **5.43%**, this method was inferred to be not suitable for forecasting since the variations in actual values were not clearly explained and due to under prediction.

Interpretation

In general, the purpose of the GRP metric is to measure impressions in relation to the number of people in the target for a TV show or any advertising medium. GRP values are commonly used by media buyers to compare the advertising strength of components of a media plan. The calculation of GRP always involves a business cycle, so if a forecasting must be made, there needs to be a large period of data. In our case, there is not enough data to see the seasonality or cyclic components. The given data has a downward trend across the time. Moreover, all the models have under predicted than the actual GRP value in the forecasting period. The reason behind this could be the lack of data and the only presence of trend variable. As there is a negative linear trend most of the model failed to pick up if any spikes in the forecast.

The **decomposition** and **time series regression** is not an appropriate model to forecast with this given data and performed poorly in terms of predicting the values as most of the values are under-predicting and remains flat. These two models also require a large dataset to build an effective predictive model. Short-term forecasting is performed by **exponential smoothing** and the predictions made by the model is quite good on coinciding with the actual model but under-predicting at places and having a comparably higher MAPE.

ARIMA was identified as the most suitable and appropriate model to perform forecasting for the given data set. Moreover, it holds a very low prediction error of 6.404% MAPE for forecasted values when compared to all the other forecasting model. And ARIMA holds good for long term prediction and the very reason behind going with this model as it senses the spikes in the forecast and shifts to an extent.

Conclusion

S.No.	Methods used	Error % (MAPE)
1	Exponential smoothing	7.95
2	**ARIMA**	**6.40**
3	Decomposition	10.30
4	Time Series Regression	5.43

Table 2

All the 4 methods were compared in terms of MAPE and captured. From Table 2, the ARIMA model had performed well in terms of forecasting and comparatively possessed a lower MAPE which gives enough support to conclude that ARIMA can be considered for forecasting for this dataset.