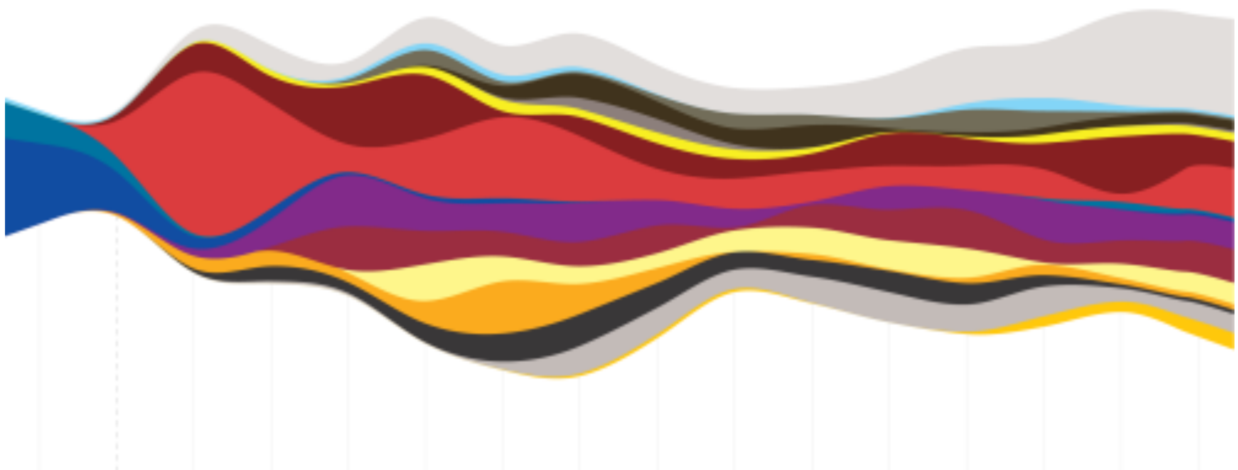

Time Series Forecasting

KE 5108 CONTINUOUS ASSESSMENT

TEAM NAME: OPTIMUS PRIME

TEAM MEMBERS

ANURAG CHATTERJEE (A0178373U)
BHUJBAL VAIBHAV SHIVAJI (A0178321H)
LIM PIER (A0178254X)
LIU THEODORUS DAVID LEONARDI (A0178263X)
TSAN YEE SOON (A0178316Y)

Table of Contents

Objective and data source

In TV advertising, the Gross Rating Point (GRP) metric is one of the primary measures that advertisers use, to determine which TV stations to place their commercials with. With an annual TV advertising market of over US$4 billion, it is of high commercial importance to be able to forecast GRP ratings accurately for India TV stations for all stakeholders.

The objective of this exercise is to demonstrate a suitable Time Series Forecast method that can accurately forecast the GRP TV channel rating data given a limited set of prior data points.

The dataset

The dataset is the actual GRP weekly metric from an Indian TV channel. A total of 92 weekly data points from 17 Jun 2007 to 15 Mar 2009 is available, of which 72 weekly data points from the beginning will be our training period and 20 weekly data points from 26 Oct 08 onwards is our forecast target, named as the validation period.

The following table summarises the variables in the data used for further processing, what they represent and their data type:

No #	Variable Name	Type	Description
1	GRPRatingsDate	Date	Date of the weekly data point
2	GRP	Real number	Gross Rating Point value

Table 1 Dataset Variables

Exploratory analysis

The complete data is plotted as a time series with the GRP ratings in the Y-axis and the dates when the ratings were available on the X-axis, and the below pattern can be observed:

Plotting the available GRP data

Figure 1: Weekly rating trend between 2007-2009

As can be seen from figure 1 above, a general downward trend is observed in the training period of the ratings. The validation period (Oct 2008 onwards) has a slightly increasing trend.

Checking the stationarity of the time series data

By observation, the original series plot shows an overall downward trend, more clearly shown by its rolling mean. To obtain a stationary series, we execute differencing, and the results are shown in the right chart of figure 2.

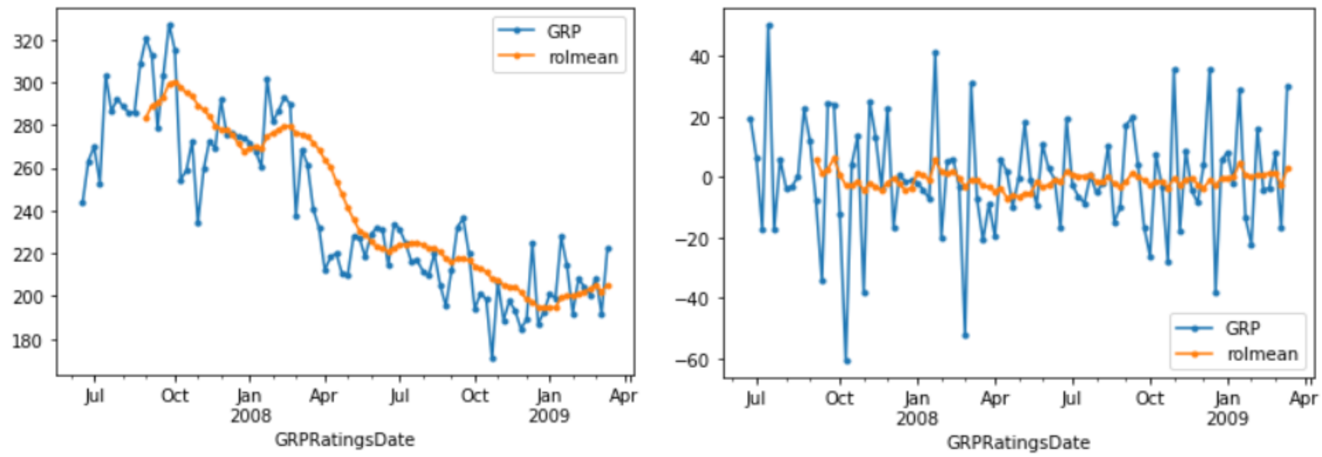

Figure 2 Rolling mean for the original (left) and first-order difference series (right)

The augmented Dickey-Fuller test is a statistical test that checks for stationarity. The null hypothesis, when accepted means the time series is non-stationary and when rejected means it is stationary. We obtained a p-value of 0.66 for the test on the original time series data and a value of 3.1e-18 on the first order difference, suggesting that the differenced series is stationary.

Lag plots and autocorrelation plots

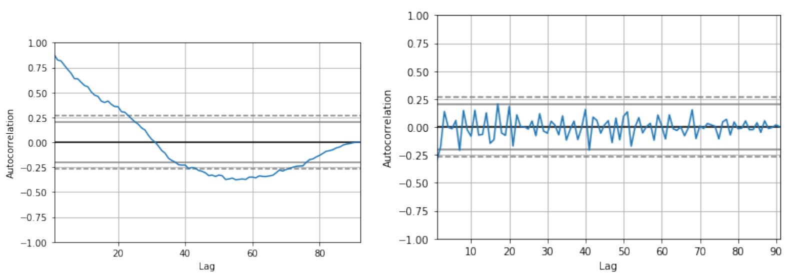

Figure 3 Autocorrelation plot of original series(left) and autocorrelation plot on series after differencing (right)

Autocorrelation plot of the original series shows significant autocorrelation in the series, while the autocorrelation plot of the differenced rating series appears random and have no autocorrelation lying outside the 95% confidence interval. This near zero plot and data randomness suggest that the weekly change in the differenced series has eliminated/reduced trend and seasonality, becoming a stationary time series and no further higher order differencing or transformation is needed for methods requiring stationary series (e.g. ARIMA)

Forecasting Methods

We first evaluated Simple Exponential Smoothing methods for forecasting the TV rating data series. Through visual inspection and computation of Mean Absolute Percentage Error (MAPE) using different values of alpha for Simple Exponential Smoothing, as shown in Figure 3, we determined that alpha=0.3 give a curve that is a reasonable approximation of the train data.

$$s_t = \alpha \cdot x_t + (1 - \alpha) \cdot s_{t-1}.$$

where α is the *smoothing factor*, and $0 < \alpha < 1$.

Exponential Smoothing

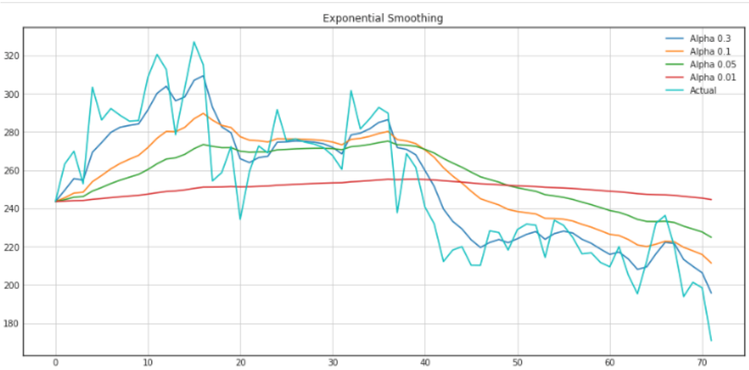

Figure 4 Comparing different values of alpha for Simple Exponential Smoothing

Running the exponential smoothing model (alpha=0.3) on the train and test period produced the following results in JMP.

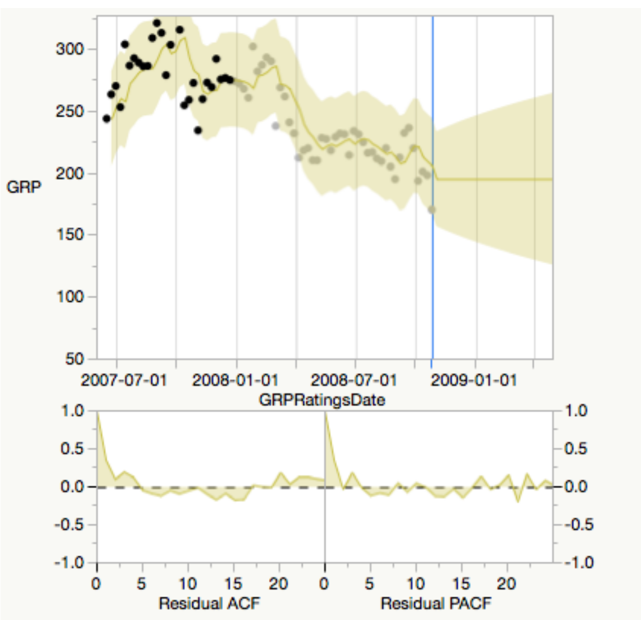

Figure 5 Simple Exponential Smoothing prediction and results

We also explored double exponential smoothing, linear exponential smoothing (which can be used for series which shows trend) and Winters' methods (which can be used for series which shows both trend and seasonality).

	Simple Exponential Smoothing		**Double Exponential Smoothing**		**Linear (Holt) Exponential Smoothing**		**Winters' Method**	
Model Smoothing Parameters	Level = 0.3		Level = 0.33		Level=0.6, Trend = 0.01		Level 0.61, Trend 0.02, Seasonal 0.37	
	Train	Test	Train	Test	Train	Test	Train	Test
MAE	14.57	10.65	14.30	77.80	13.78	34.73	17.03	37.51
MAPE	5.96%	5.09%	5.77%	38.01%	5.55%	16.84%	6.96%	18.22%

Table 2 Comparison of performance of different exponential smoothing methods

Despite the data series showing an obvious downward trend, the Double Exponential Smoothing did not perform well for the test set as the model assumed the downward trend from the train period dataset. Likewise, the Triple exponential methods (Winters) did not perform well also in the test period to account for the changing trend. The Simple Exponential Smoothing gave the best MAE and MAPE out of all the different exponential smoothing methods in both train and test datasets. However, this "good" accuracy of the simple method might be a coincidence as the method is unable to capture trend information and produced a trendless prediction that performed well on the test period data where the trend started reverse, resulting in a flat trend.

ARIMA Method

For the ARIMA method, we first made use of Auto-ARIMA function in Python on the training dataset to generate the best parameters for ARIMA and Seasonal ARIMA respectively. As the training dataset had less than two years of data, we suspected that a seasonal ARIMA model would not be valid. Nevertheless, we ran Auto-ARIMA for both seasonal and non-seasonal cases. As observed earlier in the exploration, first order differencing was also suggested by Auto-ARIMA to make the data series stationary. Using these parameters, we ran the ARIMA models in JMP and got the predictions and compared them against the test dataset.

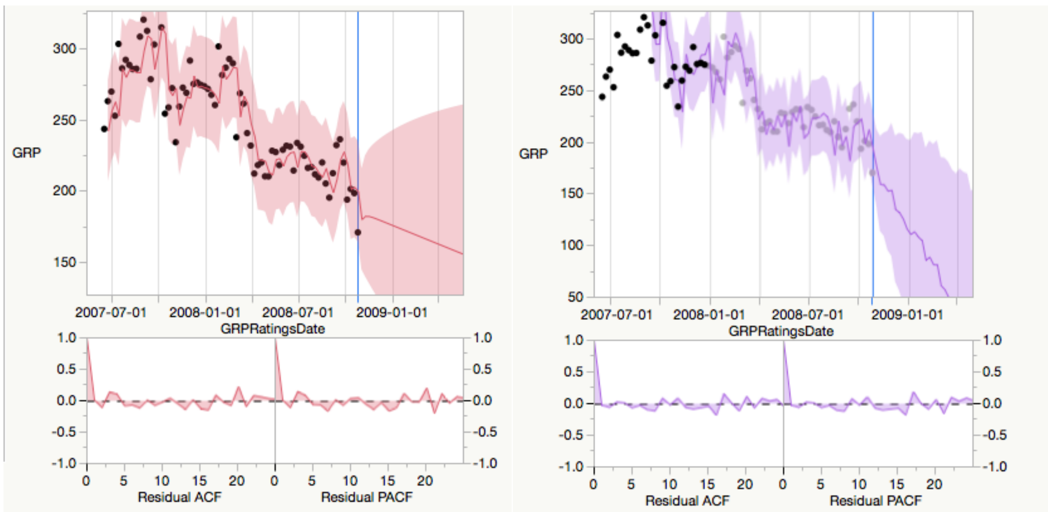

Figure 6 Right - ARIMA (1, 1, 1) result and Left - Seasonal ARIMA (0, 1, 3) x (0, 1, 1)$_{12}$ result

	Non-Seasonal ARIMA (1, 1, 1)		**Seasonal ARIMA (0, 1, 3) x (0, 1, 1)$_{12}$**	
	Train	Test	Train	Test
MAE	12.68	26.85	16.27	87.94
MAPE	5.10%	12.93%	6.60%	43.05%

Table 3 Predictions from ARIMA

As suspected, the seasonal ARIMA did not give a good result, especially on the test result. Due to the lack of data, it is harder to determine seasonality. The non-seasonal ARIMA performed better, giving a better MAPE and MAE. Both ARIMA models also assumed a downwards trend, which explains why it performed poorer on the test set. As can be seen from the residual ACF and PACF, the residuals resemble white noise.

DECOMPOSITION METHODS

Examining the plot of the TV rating data series, we notice that the oscillations in the original data appear to be diminishing over time, suggesting the multiplicative decomposition model with fractional factors should be used. However, the variation in oscillations may not be significant, and we attempted additive, multiplicative and STL decomposition methods using Python and R software for forecasting and compared their results. Due to shorter than 2-years of data points in the dataset, we assumed a quarterly periodicity (12 weeks) for the models. A method to use full year 52 weeks periodicity for decomposition is also attempted. This method suggested by Hyndman [1] allows the estimation of seasonal patterns using Fourier terms, but unfortunately, this method produced poor forecasting results for our dataset and will not be discussed in this report.

Additive Decomposition

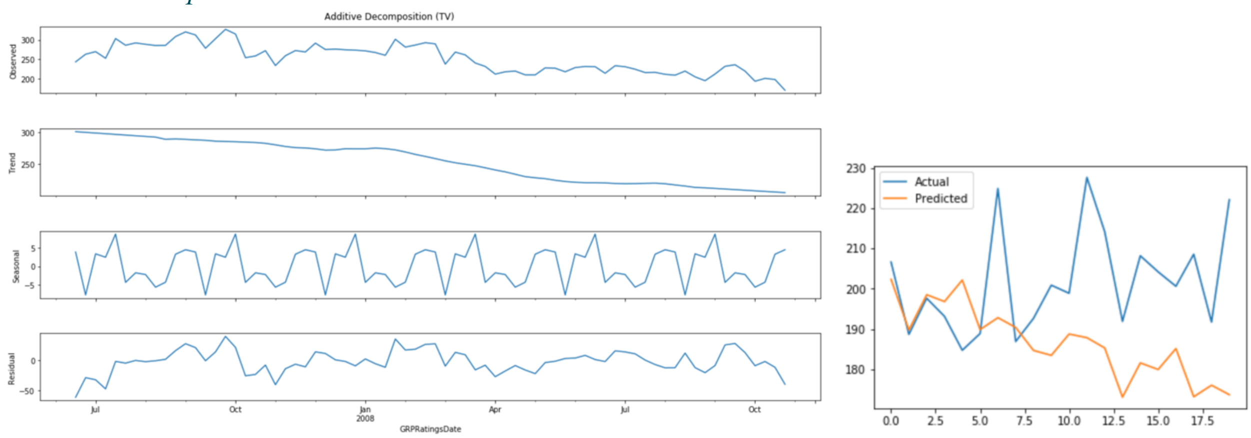

Figure 7 Additive Decomposition (left) and Forecasted Values (right)

We did a power transformation of the observed values by using GRP = GRP^0.7. We were able to get slightly better results by doing this transformation. This gave us a relatively linear trend, which we then used to estimate a linear trend equation. The equation we derived for the linear trend was: **y = -0.204x + 55.52**. We then extrapolated the seasonal values to get the repeating seasonal values. We then got the train and test results after inversing the transformation.

It was observed that the model did not perform as well on the test dataset. We also see that there are peaks in the residual that correspond to the peaks in the original observed values. This suggests that there are additional components in the signal that have not been removed by removing the trend and seasonal components.

Multiplicative Decomposition

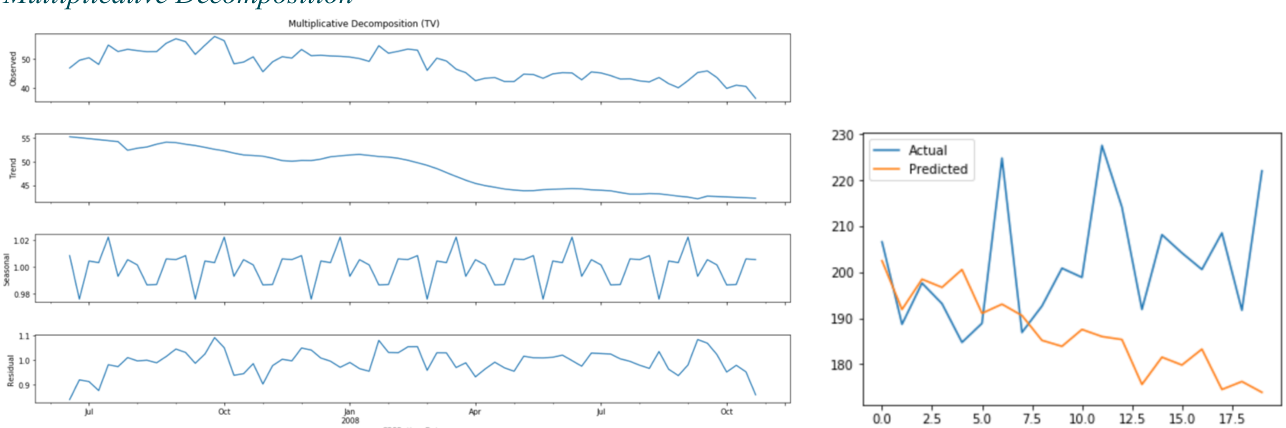

Figure 8 Multiplicative Decomposition (left) and Forecasted Values (right)

The steps to get the train and test values were similar to the additive decomposition, with the difference being that we used multiplication instead of an addition to get the eventual values.

The multiplicative decomposition method performed comparably to the additive method. Again, we see peaks in the residual corresponding to the peaks in the original observed values.

The above additive and multiplicative decomposition method used 12 weeks data as a cycle instead of a full year of 52 weeks data.

Seasonal and Trend Decomposition using Loess (STL)

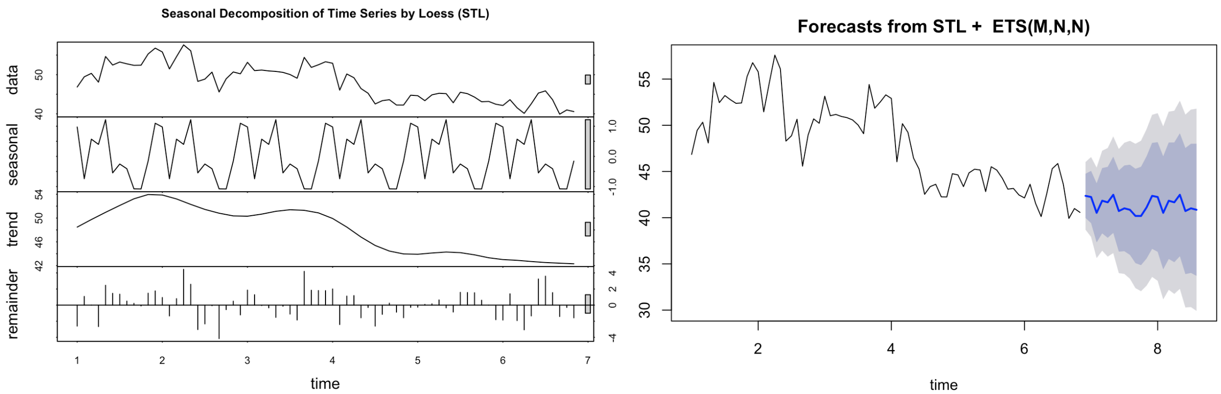

Figure 9 STL Components (left) and Forecast from STL (right)

We also decomposed the time series dataset using Seasonal and Trend Decomposition using Loess to uncover seasonality and trend of Indian TV Ratings. The STL method was developed by Cleveland et al [2] which is based on additive decomposition. The results suggests that STL perform quite well on our dataset.

	Additive Decomposition		**Multiplicative Decomposition**		**STL Decomposition**	
	Train	Test	Train	Test	Train	Test
MAE	11.19	17.61	11.09	17.7	10.74	11.56
MAPE	4.39%	8.12%	4.34%	8.48%	4.27%	5.90%

TIME SERIES REGRESSION

For Time Series Regression methods, we first tried the first-degree regression to forecast the trend in the data series. A MAPE of 8.42 % was observed on the validation set and that of 6.58% on the training set. As can be seen from the below graph, the simple regression fails to capture the seasonality and the slightly increasing trend for the validation data.

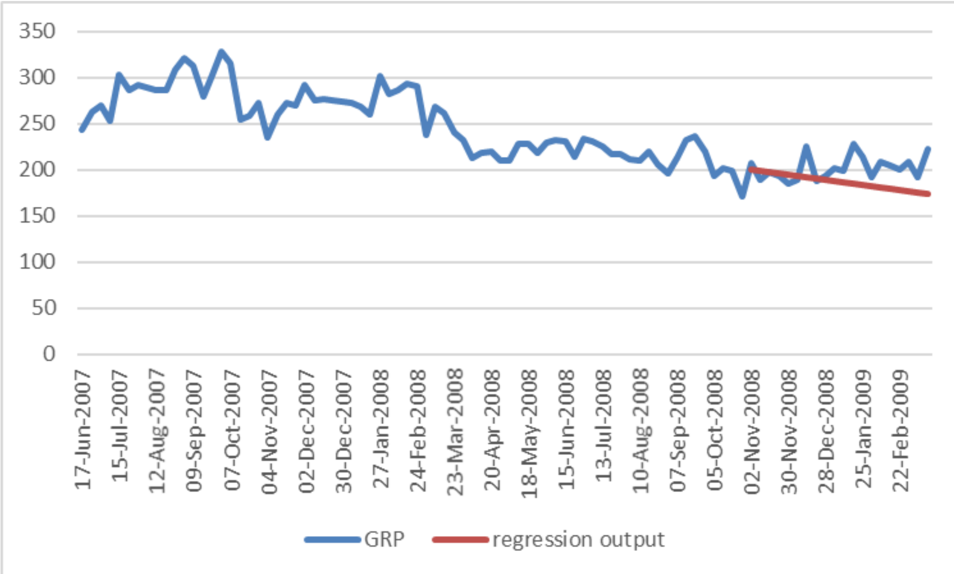

*Figure 10 First order regression output: Y(t) = -1.4*t+302.82*

To capture the seasonality in the regression model, dummy variable regression was leveraged. Since there is a decreasing seasonal variation, a power transformation (y' = y^0.7) is performed on the original GRP ratings, and the below is the new values of the GRP ratings with seasonal variation close to constant.

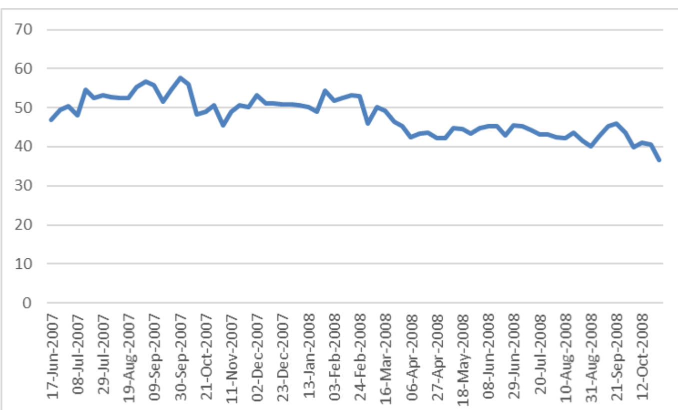

Figure 11 Power transformation of the GRP ratings

Dummy variable regression is performed after reducing the seasonal variation as shown above. Three dummy variables are introduced for the four quarters used to capture the seasonality. The result of the regression is shown below:

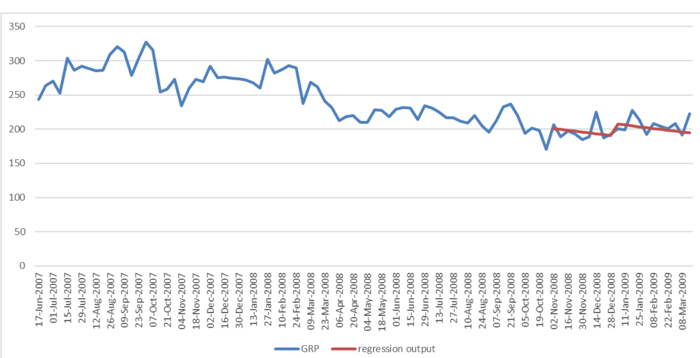

Figure 12 Dummy regression output on original data

The regression equation is **Y'(t) = -0.18017*t+ 2.5595*Q1+ -1.93706*Q2+ 0.7403*Q3+ 54.0689**, where Q1 is 1 if the date lies in the first quarter (i.e. month between January and March), Q2 is 1 if the date lies in the second quarter (April-June) and Q3 is 1 if the date is in the third quarter (July-September). Y(t), the actual predicted output is calculated as Y'(t) to the power of (1/0.7) = 1.428 to undo the initial transformation.

The result of the different types of regressions are shown below:

	Simple time series regression		Dummy variable regression	
	Train	Test	Train	Test
MAE	16.57	17.65	14.83	9.51
MAPE	6.58%	8.42%	5.87%	4.56%

Table 4 Comparison of time series regression results

For the dummy variable regression, we observe that the test errors are less than the train errors, which can be attributed to a greater height of the oscillation peaks in the train data portion and these peaks are not present in the test period data. Although dummy variable regression gives low error values compared to the other methods, we note that since the training data period is less than two years, the seasonality might not be captured well enough and forecasts for unseen future data might not be good enough.

Chosen method for forecasting India TV Channel Ratings

For the given India TV Ratings weekly data time series, we propose using **dummy variable regression method** with three additional dummy variables to capture quarterly seasonality in additional to the trend variable. In comparison to the other models, this time series model built from the method from the training period of 72 weekly data points is able to provide good and consistent forecast results (low MAE and MAPE) across both training period and the test period of 20 weekly data points. Having less than 2 years of training period data may mean our model could lead to less robust predictions. We observed that by assuming a 12-week (quarterly) seasonality, we could create a fairly accurate model using this regression method.

Forecasting Methods Inappropriate for this Dataset

- Seasonal ARIMA and Winters' method are both seasonal methods, and as expected, they performed poorly on the test period, this is most likely due to unexpected trend change on the test period. Simple Exponential Smoothing did produce excellent results during the test period, however the good results could be due to coincidence that the trend change "levelled" the data series in the test period. It is a possibility that this model might fail in forecasting of future time periods.
- Both conventional additive and multiplicative decomposition methods forecasting did not yield good results. On further inspection, the residuals were found not to be random. There are still peaks in the residual corresponding to peaks in the original values, suggesting that not all of the original time series can be explained by the extracted decomposed seasons and trend. Hence, we did not use these models.
- The simple time series regression model does not capture aspects other than the trend that is present in the series and so is a poor method for forecasting.

Poor Prediction Data Points (Over/Under-Prediction with more than 10% error)

Predicting across the entire series with the chosen dummy variable regression model, we identified 12 data points from the train period (No. 1 to 12 is shown in table below) and 3 data points from the test period (no. 13 to 15 from table below) that have absolute percentage error of more than 10 percent.

No.	Date	Absolute % Error	Direction	No	Date	Absolute % Error	Direction

1	17-Jun-2007	15.92	Over Predicted	**9**	30-Mar-2008	12.22	Over Predicted
2	01-Jul-2007	11.36	Over Predicted	**10**	29-Jun-2008	10.17	Under Predicted
3	08-Jul-2007	18.29	Over Predicted	**11**	31-Aug-2008	11.41	Over Predicted
4	02-Sep-2007	10.22	Under Predicted	**12**	26-Oct-2008	18.33	Over Predicted
5	30-Sep-2007	13.76	Under Predicted	**13**	14-Dec-2008	14.06	Under Predicted
6	07-Oct-2007	12.71	Under Predicted	**14**	18-Jan-2009	10.01	Under Predicted
7	04-Nov-2007	15.06	Over Predicted	**15**	15-Mar-2009	12.30	Under Predicted
8	02-Mar-2008	11.76	Over Predicted				

Table 5 Prediction data points with more than 10% absolute error difference from actual

General investigation into the trend of TV ratings of 2007 to 2009

In our readings into the recent reports on TV trends and seasonality of India in general, TV viewership will likely be the strongest towards the holiday season in the last quarter of the year, as there is more festive programming for Diwali, Christmas and New Year. This peak can be shown in the last two quarters of 2007, but in 2008, there is no such peak towards the end. This is most likely due to the negative effects of the 2008 global financial crisis that dampens viewership which only shows minor recovery in the first quarter of 2009.

Possible explanations for the poor predictions

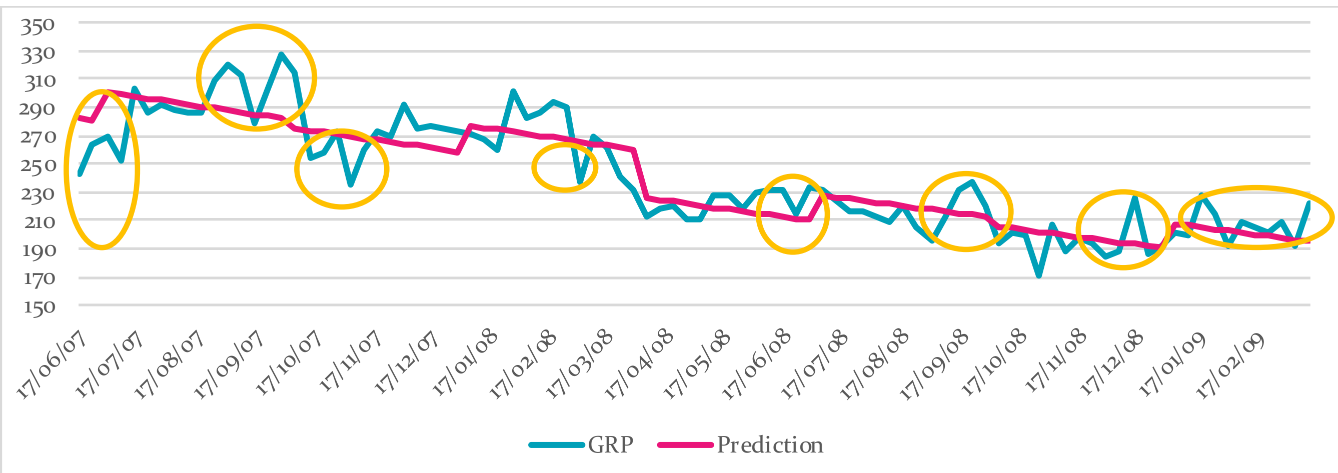

Figure 13 Prediction across the entire series with chosen model - Orange circle marks the poor predicted points

- The poor prediction at the start of the data series (June 2007) is most likely due to not having enough data points of that quarter (2nd Q ending June) for computing the model.
- The spike in TV Ratings in September 07 is likely due to festive programming during the Janmashtami and the exciting ODI Cricket series where India team was matched-up against the England team. The ratings spike in early October is also likely due to festive programing for Mahatma Gandhi Jayanti gazette holiday.
- The sharp drop in TV Ratings in early November 07 could be due to a series of earthquake and aftershocks happening in the Gujarat region.
- The overprediction occurring in the two weeks of March could be due to the increasing stress of the US subprime crisis affect the mood even in India. In March 2008, bear sterns, a large US Bank almost collapsed. During that month, stock markets around the world suffered the most year-to-date.
- In June 2008, the Olympics in China could have gathered more viewership, therefore the model underpredicted during this period.
- From August through the end of the year of 2008, the global financial crisis is in full swing, stock markets around the world suffered huge losses, with some losing more than 50% of total share market capitalization

(S&P 500 lost 36% from June to August and another further 30% to the end of the year). This is very likely to dampen the mood in India and thus viewership of TV programs in India too.

- An unfortunate major terrorist attack happened in Mumbai on 26 Nov 2008, and the subsequent reporting and increased coverage on the event could have led to increase viewership in early December 2008.
- The mood and thus viewership in India most likely have bottom out around January 2009 and morale of the people started recovering, resulting in increasing TV viewership. This change of trend caused the model to under-predict for 2 weeks in January.

Reference

1. Hyndman, R.J. *Seasonal decomposition of short time series*. 2018 14 July 2018 5th Sept 2018]; Available from: https://robjhyndman.com/hyndsight/tslm-decomposition/.

2. Cleveland, R.B., et al., *STL: A Seasonal-Trend Decomposition.* Journal of Official Statistics, 1990. **6**(1): p. 3-73.