**ECONOMETRICS CIA 2**

**Course code : Eco 6.08**

**STAR PLUS FORECAST**

**Name Roll No UID No**

**Divya Somakumar 036 131141**

**Roveena D’Souza 037 131406**

**Sayali Jadhav 044 131004**

**Jerrin Anna Joe 046 131288**

**Ishita Kamath 050 131008**

**INTRODUCTION**

The Indian Media and Entertainment (M&E) industry is a sunrise sector for the economy and is making high growth strides. Proving its resilience to the world, the Indian M&E industry is on the cusp of a strong phase of growth, backed by rising consumer demand and improving advertising revenues. The industry has been largely driven by increasing digitization and higher internet usage over the last decade. With the introduction of digital distribution platforms like direct-to-home (DTH) and Mobile Tv, Indian television industry has undergone a revolutionary change. The small screen industry in India has been doing extremely well big business over the years. It has surpassed even the silver screen industry in attracting the consumers through innovation and new themes appealing the audience. An average person spends about 3-4 hours daily watching the television. Multiple channels in each genre are competing with each other for TRP, increasing pay TV penetration, expanding yet fragmented local as well as overseas viewership of Indian channels. Demand for more specific content clearly set the stage for the next level of growth and transition for players across the television value chain. There are more than 500 channels of different genre which have been increasing over the years. They have been selling the viewership as products to make business.

Through this assignment to attempt the forecast the viewership of the top rated channel Star. Star India is an Indian media and entertainment company, owned by 21st Century Fox. Star India has defined the Indian media landscape for over two decades and today is one of the country’s leading media conglomerates, reaching approximately 650 million viewers a month across India and more than 100 countries. Star generates 20,000 hours of content every year and broadcasts 40 plus channels in 8 different languages, reaching 9 out of 10 C & S TV homes across India. The network’s entertainment channel portfolio includes Star Gold, Star Utsav, Star Plus, Star Gold, Life Ok, Channel V etc. Star Tv is now also present in the Indian movie production and distribution space through Fox Star Studios, an affiliate joint venture company. Star Tv has been at the top since 2000 and we aim to forecast its viewership through analysis using ARIMA (Autoregressive Integrated Moving Average). Also our objective is to study the impact of other channels on the viewership of Star through a Linear Regression.

Methodology

For forecasting of any commodity or service the sources and the method by which data has been collected is of prime importance since it makes the research more clear and transparent to the person reading it. The methodology adopted for this particular project involves collection of data from various secondary sources which include official websites of barcindia.com and the official website of star India. in and including use of ARIMA (autoregressive integrated moving average) for forecasting the various variables.

1. **BARC** (Broadcast Audience Research Council) India ,one of the secondary sources to obtain data is an industry body, to design, commission, supervise and own an accurate, reliable and timely television audience measurement system for India. Guided by the recommendations of the TRAI (Telecom Regulatory Authority of India) and MIB (Ministry of Information and Broadcasting) notifications of January 2014, BARC India brings together the three key stakeholders in television audience measurement - broadcasters, advertisers, and advertising and media agencies, via their apex bodies. BARC India seeks to publish data from both rural and urban areas to get a get a clearer picture on the working of star India. Apart from BARC we also collected data from the website of Star India.We took observation of 52 weeks of the year 2015 to forecast the demand for Star Plus..

**ARIMA**

A time series is a set of well-defined data items collected at successive points at uniform time intervals. Time series analysis is an important part in statistics, which analyzes data set to study the characteristics of the data and helps in predicting future values of the series based on the characteristics. Forecasting is important in fields like finance, industry, etc. In our study , we have used ARIMA. Autoregressive and Moving Average (ARMA) model is an important method to study time series. (ARIMA) model is derived by general modification of an autoregressive moving average (ARMA) model. This model type is classified as ARIMA (p,d,q), where p denotes the autoregressive parts of the data set, d refers to integrated parts of the data set and q denotes moving average parts of the data set and p,d,q is all nonnegative integers. ARIMA models are generally used to analyze time series data for better understanding and forecasting. Initially, the appropriate ARIMA model has to be identified for the particular datasets and the parameters should have smallest possible values such that it can analyze the data properly and forecast accordingly. [14] The Akaike Information Criteria (AIC) is a widely used measure of a statistical model. It is used to quantify the goodness of fit of the model. When comparing two or more models, the one with the lowest AIC is generally considered to be closer with real data. AICc is AIC with a correction for finite sample sizes. Before the application of ARIMA it is necessary to identify if the data is stationary or non stationary. If it is stationary ARIMA be applied directly if not stationary we need to stationarize it and then apply ARIMA. In order to test for stationarity we make use of the Augmented Dickey Fuller test. In statistics and econometrics an **augmented Dickey–Fuller test (ADF)** is a test for a unit root in time series sample. It is an augmented version of the Dickey– Fuller test for a larger and more complicated set of time series models. The augmented Dickey–Fuller (ADF) statistic, used in the test, is a negative number. The more negative it is, the stronger the rejection of the hypothesis that there is a unit roots at some level of confidence.

**The R Result**

**Objective : To forecast the Viewership of Star Plus for upcoming weeks of February.**

**HYPOTHESIS: The Viewership of Star Plus is going to increase in the month of February.**

The main objective of our paper is to find whether Star Plus is one of the Top viewed Hindi General Entertainment channel. For the validation of our Hypothesis we took the weekly data for the last year. We collected 55 observations for the validation of the same. We have collected weekly data on the viewership of Star Plus for the year 2015. For our analysis we have used ARIMA (Autoregressive Integrated Moving Average) which is also popularly known as the Box and Jenkins Methodology. This approach includes model identification, Parameter estimation and selection model checking and forecasting. Model identification involves determining the orders (p,d,q) of the AR and MA components of the model.

**MODEL IDENTIFICATION**

The first step in ARIMA Model Building is to identify whether the chosen time series is stationary or non stationary. A time series is said to be stationary if its mean and variance remains constant over the period of time. An ARIMA Model cannot be built if the Time series is not stationary. In order to stationarise we need to difference our time series “d” times. By running the Auto ARIMA function in R , the appropriate ARIMA model that we obtained is (1,0,1), where the value of “p” is 1 , “d” is 0 and “q” is 1. The below figure indicates the Times eries plot of the Viewership of Star in the year 2015.



Figure 1 Plot of Star Plus Viewership

A Trend Line can be fitted for this data. We have used excel to fit a trend line against time. The below trendline Clearly shows that there is no cyclical or seasonal trend in the Viewership of Star Plus. We cannot identify any trend.

Thus we can smoothen this curve by using Simple Moving Average . The following Diagram shows the Smoothened Curve. We have considered 3 week moving average.

From the above graph, it can be easily concluded that our time series appears to be stationary. In order to prove that the above series is stationary, we used the Augmented Dickey Fuller test (ADF).

**TEST FOR STATIONARITY: AUGMENTED DICKEY FULLER (ADF) TEST**

In this case we have two Hypothesis .Let our Null Hypothesis Ho be “The Time series Data is not stationary” and let our Alternative Hypothesis be “The Time Series is stationary.” This test was conducted in R and its results are as follows.

Dickey –Fuller =-1.8682, lag order=3, p value =0.624

Thus we reject the Null hypothesis (Ho) that the Time Series is not stationary. Hence our alternative Hypothesis ie. The time series is stationary in its mean and Variance is true. Thus we can conclude that there is no need of differencing our time series data. (The working of the ADF Test is attached at the end of the assignment.)

**PARAMETER ESTIMATION**

The second step is to identify the Parameters “p” and “q”. For this, we have made use of the Correlogram and Partial Correlogram. The Graph of ACF (Autocorrelated Function) is plotted below. The below figure shows significant spikes through lag 1. Hence the value of Moving Average (MA ) term ie “q” is 1. It also shows that it starts decaying exponentially with a Sine wave pattern at lag=1. Thus our Autoregressive term (AR) ie “p” is also 1.



Figure 2 Plot of ACF

From the Plot of PACF (Partial autocorrelation Function) given below we can conclude that our Time series is only significant at lag 1 . There is significant spike through lag 1. Thus our “p value is 1 . The PACF plots declines exponentially after lag =1. Thus our “q” value is 1. Thus at ARMA (1,1) the PACF plot decays exponentially.



Figure PLOT OF PACF

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 0.685 | -0.119 | 0.37 | -0.072 | -0.21 | -0.146 | -0.104 | 0.193 | 0.052 | -0.035 |

THE VALUES OF PACF FOR LAG 10

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 1 | 0.68 | 0.406 | 0.424 | 0.401 | 0.183 | -0.008 | -0.0068 | -0.0010 | -0.005 | -0.0092 |

THE VALUES OF ACF FOR LAG 10

By using auto arima function in R , we directly got our values of (p,d,q) (1,0,1).

**ESTIMATION**

Let “Yt” denote the Viewership of Star Plus in the last year. Thus our identified AR model is

Yt = Delta+ a1Yt-1 + a8Yt-8 +a12Yt-12

Thus after conducting the auto.arima function in R we got the following values. The value of ar1 is 0.4377 with standard error 0.1654 and the value of ar2 is -0.8538 with standard error of 0.09. The value of AIC is 622.3.

Thus our ARIMA with order (1,0,1) can be described as.

Y = 639.9371 + 0.4377 Yt-1 – 0.8538Yt-2

The Akaike Information Criterion (AIC) is 636.16 AIC is widely used to measure the statistical model. Its used to quantify the goodness of fit. The Bayesian information Criterion (BIC) is 643.95.The AIC model is more advantageous than the BIC model.

**FORECAST**

After Conducting the auto . arima, we forecasted the Viewership of Star. The Forecast showed that the Viewership of Star was going to decline in the next couple of weeks. The ARIMA forecasted that the the Viewership of star will decline from 681 million viewers to 647 million Viewers in the first week of February . Further in the second week of February 2016 it will fall to 643 million views and in the third week it would fall further to 640 million. Thus our Hypothesis that Star Plus would be the Top Viewed channel is false. Thus our hypothesis is nullified. The following Table shows the forecast for Viewership of Star Plus at different levels.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Week | Forecast | Low 80 | High 80 | Low 95 | High 95 |
| Week 4 | 647.5575 | 518.3447 | 776.774 | 449.9435 | 449.9435 |
| Week 5 | 643.2725 | 471.1731 | 815.3719 | 380.0692 | 380.0692 |
| Week 6 | 641.3970 | 462.2494 | 820.5446 | 367.4144 | 367.4144 |
| Week 7 | 640.561 | 460.1096 | 821.0426 | 364.5764 | 364.5764 |
| Week 8 | 640.2168 | 459.4987 | 820.9348 | 363.8323 | 363.8323 |

**LIMITATIONS OF THE MODEL**

Like any other predictive model ARIMA also has its own limitations. All the predictions of ARIMA cannot be always accurate. The Viewership of Star Plus may decline or it may increase. There are several other social factors which also affect the Viewership of Star Plus. Entertainment industry is very subjective. Thus our forecasting cannot be considered to be 100 % accurate. Thus inorder to increase the Viewership of Star Plus , Star plus need to come up with shows which have a strong message. They should concentrate more on their content. They should introduce more number of reality shows.The increase in number of Hindi General entertainment channels have also contributed to the fall in the viewership of Star Plus.

**CONCLUSION**

This paper mainly focuses on the forecasting of the viewership of the top rated Star channel using ARIMA model. Star India is an Indian media and entertainment company, owned by the 21st Century Fox. The variable under this study was the Viewership of star Plus in the year 2015.The data regarding the viewership of Star plus was obtained from the Broadcast Audience Research council. The result of this study shows that the forecast for the next week’s viewership of star is going down from 643.2725 million to 641.3970 million. Thus, our hypothesis of star being the next week’s top viewed channel is wrong, this could be because of other socio-cultural factors which we could not include such as; monotonous programmes in Star, better interesting programmes in other channels, lack of new ideas and people preferring to watch online entertainment. Thus this study shows that there is going to be a decline in the number of iewership of Star Plus.

**REFERENCES**

* Coghlan , Avril (2010) , A Little Book for R for Time Series
* S P Gupta(2009) , Statistical methods, Sultan Chand & Sons, New Delhi.
* Damodar Gujarati , Basic Econometrics, Fourth Edition , Tata Mc Graw Hill Education Private limited, New Delhi
* Kumar Manoj , Anand Madhu, “ An Application of Time Series ARIMA forecasting model for predicting sugarcane production in India” available at <http://eccsf.ulbsibiu.ro/articole/vol91/918kumar&anand.pdf>
* Mondal, Goswami, “ Study of Effectiveness of Time Series Modelling (ARIMA) in forecasting stock prices” available at <http://airccse.org/journal/ijcsea/papers/4214ijcsea02.pdf>
* Broadcast Audience Research Council <http://www.barcindia.co.in/>
* Star India <http://www.startv.com/aboutus.aspx>
* Statistical Package R
* Microsoft Excel.