International Journal of Computer Systems (ISSN: 2394-1065), Volume 04– Issue 01, January, 2017
Available at http://www.ijcsonline.com/

A Survey on Promotional and Base Level Forecasting using ARIMA

Aakash Raina, AkashVartak, Prof. P. R. Patil, Yash Katariya, Yash Lahoti

Department of Computer Engineering,
Pune Institute of Computer Technology
Pune, India

Abstract

Forecasting is the process of making predictions of the future based on past and present data and most commonly by analysis of trends [2]. Sales forecasting is the process of estimating future sales. To make informed business decisions and predict short-term and long-term performance an accurate sales forecast is a must. Promotions and targeted marketing is used by many businesses to increase the demand for or visibility of their product. Such promotions normally require increased expenses, loss of revenue and maybe additional costs. But having incurred all the financial and non-financial losses business owners need to determine the value and benefit of these promotions. So one way to evaluate promotions is to analyze the historical data using time series analysis techniques. This paper briefly explores time series analysis for promotional and base level forecasting.

Keywords: *Promotional analysis; Demand Analysis; Forecasting; Sales; Time series analysis; ARIMA.*

I. INTRODUCTION

Companies make their forecasts on past sales data, industry standards, economic trends and consumer needs for a product. Risks and uncertainty are fundamentals involved in forecasting and prediction; it is considered good practice to inform the degree of uncertainty linked to a forecast. In any case, the data must be up to date in order for the forecast to be as accurate as possible [3].

According to [1] sales forecasting allows companies to:

1. Predict achievable sales revenue;
2. Efficiently allocate resources;
3. Plan for future growth.

II. SUPPLY CHAIN AND MANAGEMENT

All universal definitions of "supply chain" encompasses the following three functions:

1. Raw material supply to manufacturer;
2. Manufacturing;
3. Sale of finished goods through a network of distributors and retailers to a final customer.

Companies involved in various stages of this process are linked to each other through a supply chain.

To allow for smooth flow of products and to meet demands, information is shared up and down the supply chain, i.e. with suppliers and clients. Such sharing of information enables all parties to plan accordingly to meet the consumer needs. Achievables through successful supply chain management:

1. Minimised inventory;
2. Reduced costs;
3. Improved product time to market;
4. Enhanced flexibility.

The more the companies within a supply chain are able to integrate and coordinate their activities, the more likely they'll be to optimize the flow of goods from supplier to customer and to react efficiently to changes in demand [4].

III. FOUR FUNDAMENTALS OF PROMOTIONAL FORECASTING[6]

A. Process

A clear, concise and well thought out and developed process must take into account both the promoted and unpromoted sales forecasts and sale histories.

B. People

It is necessary to “break down the silos that exist within an organization and drive collaboration” [6] across the breadth of the chain.

C. Technology

The use of right technology is required so that accurate and targeted insights can be obtained which shall enable a business to take promotional decisions.

When consumers were asked what was the single most effective thing retailers could do from a sales and service perspective to improve their overall perception to that customer, more than 64% responded with “Frequent Promotions” or “Gift/Reward Programs” [5].

D. Metrics

All directly and indirectly used and required metrics for the process must be clearly defined in a consistent manner.

When consumers were asked what was the single most effective thing retailers could do from a sales and service perspective to improve their overall perception to that customer, more than 64% responded with "Frequent Promotions" or "Gift/Reward Programs" [5].

Key Respondent Data

	Size	Average Basket Size in $	Pay % More	Win Rating
Frequent Promotions	38.9%	$71.69	4.2%	93.00
Gift/Reward Programs	25.3%	$60.55	2.6%	93.02
Exchange products easily	18.4%	$59.43	3.0%	93.14
Gift Certificates	10.1%	$87.48	4.7%	95.15
Assist with Child Needs	3.3%	$77.03	3.0%	94.84
Personal Consultation	3.3%	$103.16	5.4%	94.87
Come to home with Products	0.7%	$132.14	5.3%	91.02

Figure 1

IV. Trade Promotion Forecasting

Trade Promotion Forecasting (TPF) is the process that attempts to discover multiple correlations between trade promotion characteristics and historic demand in order to provide accurate demand forecasting for future campaigns [7]. The basic activity in modeling promotion behavior is differentiating the demand due to impact of promotion from the baseline demand of the product. Model determination enables what-if analysis to evaluate different campaign scenarios with the goal of improving promotion effectiveness and ROI at the product-channel level by selecting the best scenario [7].

V. Time Series and Analysis

A discrete-time series consists of a set of random variables, which are observed at equally spaced time periods (say daily or monthly) and which are ordered and indexed by time. Assuming that *T* observations are recorded, then a single realization of a univariate time series can be represented by y(t), where *t = 1,..., T.*

Extracting meaningful data and statistics of use is time series analysis. Using the time series analysis

VI. ARIMA

An autoregressive integrated moving average (ARIMA) model is a combination of an autoregressive and moving average (ARMA) models. Both the models are run over time series data to better understand the data, analyze, to predict/forecast future points in the series[9]. ARIMA models are always applied where data show evidence of non-stationarity[9].

A stationary time series is one whose properties do not depend on the time at which the series is observed. So time series with trends, or with seasonality, are not stationary — the trend and seasonality will affect the value of the time series at different times. On the other hand, a white noise series is stationary — it does not matter when you observe it, it should look much the same at any period of time[10].

To make a time series stationary — compute the differences between consecutive observations. This is known as differencing[10].

Transformations such as logarithms can also be used to stabilize the variance of a time series. Differencing can help stabilize the mean of a time series by removing changes in the level of a time series, and so eliminating trend and seasonality[10].

A. Autoregressive model

An autoregressive model of order p can be written as

$$Y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \ldots\ldots + e_t$$

where c is a constant and e_t is white noise. consisting of *lagged values* of y_t as predictors. This as an AR(p) model.

B. Moving Average

A moving average model uses past forecast errors in a regression-like model.

$$y_t = c + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \ldots\ldots\ldots\ldots$$

where e_t is white noise. This as an MA(q) model. Here each value of y_t is a weighted moving average of the past few forecast errors.

C. ARIMA

$$\hat{y}_t = c + \phi_1 y'_{t-1} + \cdots + \phi_p y'_{t-p} + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q} + e_t$$

This would be an ARIMA(p,d,q) model consisting of both AR model and MA model, where:

p : Number of parameters for Autoregressive part

q : Number of parameters for Moving Average part

d : Number of times differencing is done

D. ACF and PACF plots

It is not easy to interpret, simply from a plot of time series data, what should be the appropriate values for p and q. However, it is possible to use the ACF plot, and the closely related PACF plot, to determine appropriate values for p and q.

An ACF plot shows the autocorrelations which measure the relationship between y_t and y_{t-k} for different values of k where K signifies the *lag*.

Similarly we can use partial autocorrelations from the PACF plot. It measures the relationship between y_t and y_{t-k} after removing the effects of other time lags -- 1,2,3,…,k−1.

VII. IMPLEMENTATION

Arima was implemented on the proprietary data-set provided by the company sponsoring our project. First, the data was read in and the column to be predicted was extracted. The frequency was set to 52 weeks as the data was yearly. The algorithm was trained on the following set of features : General store id, General item id, Avg weekly temperature, Avg weekly precipitation, Posqty, Pos amount, Price. The Price feature was the one to be predicted. The dataset contained more than 10 lakh rows representing data from 2009 to 2016. ACF and PACF plots were made to find the values for p,q. Then, the ARIMA algorithm was run on the cleaned data to forecast the values for the next 5 weeks. The Price feature was predicted with an approximate error rate of 20%.

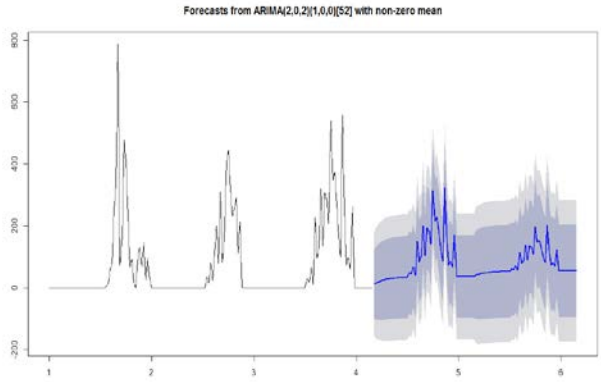

Figure 2 [11]

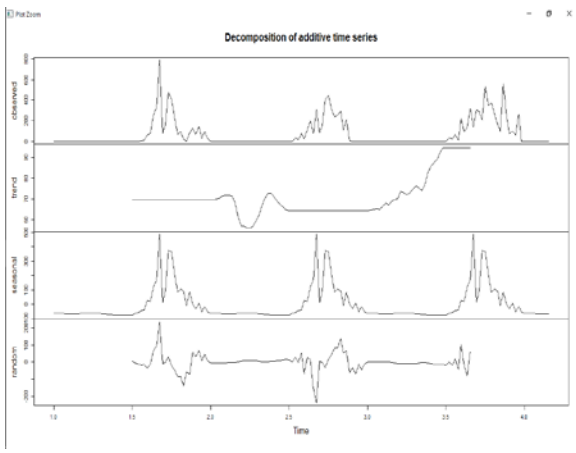

Figure 3 [12]

ACKNOWLEDGMENT

This research was fully supported by our professor and mentor Prof P. R. Patil. Continuously providing us with insight and expertise and greatly assisting us in the writing of this paper, he also provided us with valuable comments that helped improve the quality of the manuscript.

REFERENCES

[1] http://trackmaven.com/marketing-dictionary/sales-forecasting/

[2] https://en.wikipedia.org/wiki/Forecasting

[3] S. Armstrong, F. Collopy, A. Graefe, K. C. Green, “Answers to Frequently Asked Questions”, Retrieved May 15, 2013

[4] http://www.supplychaincanada.org/en/supply-chain

[5] “Four Critical Elements of Retail Supply Chain Success”, Highjump Software

[6] R. Tyagi, J. Morrow, M. Geertsen, “Optimizing Demand Sensing and Promotional Forecasting Accuracy in a Dynamic Environment”, White Paper, Tata Consultancy Services

[7] https://en.wikipedia.org/wiki/Trade_Promotion_Forecasting

[8] M. Leonard, “Promotional Analysis and Forecasting for Demand Planning: A Practical Time Series Approach”, SAS Institute Inc. Cary, NC, USA

[9] https://en.wikipedia.org/wiki/Autoregressive_integrated_moving_average

[10] https://www.otexts.org/fpp/

[11] The data was decomposed to analyze the Trend,Seasonality and Randomness of the data.

[12] Forecasted Output with ARIMA(2,0,2)