Journal of the Nigerian Association of Mathematical Physics
Volume27, (*July, 2014*), pp 209 – 214

Comparison of Different Methods for Smoothing And Analyzing Time Series Data

A. H. Bello

Department of Statistics, School of Sciences, Federal University of Technology, Akure – Ondo State. Nigeria

Abstract

This study is on time series analysis on quarterly sales of consumable goods at GlaxoSmithKline Nigeria plc. The data was collected from the consumer sales department in GSK and the aim is to investigate the trends, seasonal variations and forecast for the company. The objectives are toinvestigate the sales performance of the GlaxoSmithKline Nigeria plc. Ilupeju, Lagos State.. The methodology employed the use of least squares method, moving average method and the exponential smoothing method and from each of the method used, we were able to deduce that GSK sales are high during the first quarter of the year (i.e. period prior to winter) and decline drastically during the third quarter of the year (i.e. period prior to summer). A comparison was made among these methods and it was found that the exponential smoothing method has the least sum of squares residuals which makes it the best model created for the company.

Keywords: Time series, seasonal variation, least squares, moving averages.

1.0 Introduction

Time series is a set of data that are successively collected at a regular intervals of its time of occurrence. The regular interval time can be daily, weekly, monthly, quarterly or yearly.
Examples of some time series data include:

- Monthly sales of a company.
- Amount of annual rainfall over a period of time
- Money deposited in a bank on various working days.

It is also, essential to know that when a time series is analyzed it has the following benefits:

i. Understanding the past behavior of a variable and able to determine the direction of periodic fluctuations and also predict future tendencies of the variable.
ii. Determining the impact of the various forces influencing different variables which then facilitate their comparison; such as the difference that may have to do with price of commodities and the physical quantity of goods produced, marketed or consumed in order to make a comparison between periods of time, schools and places.
iii. Knowing the behavior of the variables in order to iron out intra-year variations as control events.

Forecasting in order to predict future occurrence especially in areas of sales and profit making, generation of revenue, dividends etc. has been a major concern and responsibility of most companies and business organizations. Hence, scheming for the future is very important, and for any company or organization to plan effectively and operate efficiently, there is need for accurate and efficient forecasting [1].

1.1 Source of Data

Data can be sourced from different places and in different ways. The source of data depends on what we want and what we are dealing with. The data used for this study was extracted from the consumer sales department in GlaxoSmithKline Nigeria plc. Ilupeju, Lagos State.

Corresponding author: A. H. Bello E-mail: bellab_2011@yahoo.com Tel.: +2347033420672

1.2 Aim and Objectives

The aim is to develop a time series model for monthly sales performance in GlaxoSmithKline Consumer Nigeria plc.
The objectives are:

i. To investigate the sales performance of the GlaxoSmithKline Consumer Nigeria plc. Ilupeju, Lagos State.
ii. To compare and analyze by the method of least squares, moving average and Exponential Smoothing.
iii. To deseasonalize the data and know the period when their sales is at brisk or at lost
iv. To use seasonal adjustment to forecast the sales performance for future plan.

1.3 Motivation

Having a deep knowledge of past events help in making appropriate and balance decision not only for the present growth but for the overall future progress. In fact planning is essential for the survival and development of any company and business organization. The major motivation for this research is the keen interest in the process of using time series analysis for making forecast of monthly sales growth in GlaxoSmithKline Consumer Nigeria plc. Ilupeju, Lagos State and the advantage is that one is able to make sound policies and strategic decisions for the future thereby avoiding past errors or mistakes. The forecast could help the company in making their budget which could help them in having a balanced income and expenditure system.

1.4 Scope and Coverage

This study covers statistical analysis of the data spanning 72 months (January 2005-December 2012) of the monthly consumer sales growth of GlaxoSmithKline Nigeria plc. The sales are in millions of naira.

2.0 Literature Review

A time series is a collection of data recorded over a period of time—weekly, monthly, quarterly, or yearly. Two examples of time series are the sales by quarter of the Microsoft Corporation since 1985 and the annual production of sulphuric acid since 1970. An analysis ofhistory—a time series—can be used by management to make current decisions and plans based on long-term forecasting. We usually assume past patterns will continue into the future. Long-term forecasts extend more than 1 year into the future; 5-, 10-, 15-, and 20-year projections are common. Long-range predictions are essential to allow sufficient time for the procurement, manufacturing, sales, finance, and other departments of a company to develop plans for possible new plants, financing, development of new products, and new methods of assembling. Forecasting the level of sales, both short-term and long-term, is practically dictated by the very nature of business organizations. Competition for the consumer's naira, stress on earning a profit for the stockholders, a desire to procure a larger share of the market, and the ambitions of executives are some of the prime motivating forces in business [2 – 5]. Thus, a forecast (a statement of the goals of management) is necessary to have the raw materials, production facilities, and staff available to meet the projected demand. This deals with the use of data to forecast future event. .Almost all businesses tend to have recurring seasonal patterns. Men's and boys' clothing, for example, have extremely high sales just prior to Christmas and relatively low sales just after Christmas and during the summer. Toy sales is another example with an extreme seasonal pattern. More than half of the business for the year is usually done in the months of November and December. Many businesses try to even out the seasonal effects by engaging in an offsetting seasonal business. At ski resorts throughout the country, you will often find golf courses nearby. The owners of the lodges try to rent to skiers in the winter and golfers in the summer. This is an effective method of spreading their fixed costs over the entire year rather than a few months.

It is observed that the quarterly sales, in millions of dollars, of Hercher Sporting Goods, Inc. They are a sporting goods company that specializes in selling baseball and softball equipment to high schools, colleges, and youth leagues. They also have several retail outlets in some of the larger shopping malls. There is a distinct seasonal pattern to their business. Most of their sales are in the first and second quarters of the year, when schools and organizations are purchasing equipment for the upcoming season. During the early summer, they keep busy by selling replacement equipment. They do some business during the holidays (fourth quarter). The late summer (third quarter) is their slow season [6].

3.0 Methodology

3.1 Fitting a Trend: Least-Squares

Given a set of data and the desire to produce some kind of model of those data, there are a variety of functions that can be chosen for the fit. If there is no prior understanding of the data, then the simplest function to fit is a straight line with the data plotted vertically and values of time (t = 1, 2, 3, ...) plotted horizontally.

Once it has been decided to fit a straight line, there are various ways to do so, but the most usual choice is a least square fit. This method minimizes the sum of the squared errors in the data series, denoted the y variable.

Given a set of points in time $\boldsymbol{t}$, and data values $\boldsymbol{y_t}$ observed for those points in time, values of a and b are chosen so that

$$\sum_t \left\{ [(at + b) - y_t]^2 \right\} \quad (1)$$

is minimized. Here *at* + *b* is the trend line, so the sum of square deviations from the trend line is what is being minimized. This can always be done in closed form since this is a case of simple linear regression.
For the rest of this article, "trend" will mean the slope of the least squares line, since this is a common convention.

3.2 The Moving Average Method

This is another important method of calculating the trend of a time series. According to this method the trend is found by smoothing out fluctuations of the data by means of a moving average. The moving average is a series of successive averages secured from a series of values by averaging groups of successive values of the series. A moving average is defined as an artificially constructed time series, from the original series in which each monthly, daily or annually figure is replaced by the means of itself and the values corresponding is a number of preceding and successive periods.[7]
Moving Averages for a" span" of k periods,
y_t= moving average through time

$$t = \frac{y_t + y_{t-1} + y_{t-2} + \cdots y_{t-k-1}}{k} \quad (2)$$

Where seasonal effects are expected, it is standard to use.
k=number of periods per cycle.

3.3 The Exponential Smoothing Method

Exponential smoothing is a procedure for continually revising a forecast in the light of more recent experience. Exponential Smoothing assigns exponentially decreasing weights as the observation get older. In other words, recent observations are given relatively more weight in forecasting than the older observations.This is also known as simple exponential smoothing. Simple smoothing is used for short-range forecasting, usually just one month into the future. The model assumes that the data flunctuates around a reasonably stable mean (no trend or consistent pattern of growth). The specific formula for simple exponential smoothing is [12]

$$St = \alpha * Xt + (1 - \alpha) * St{-}1 \qquad 0 < \alpha < 1 \quad (3)$$

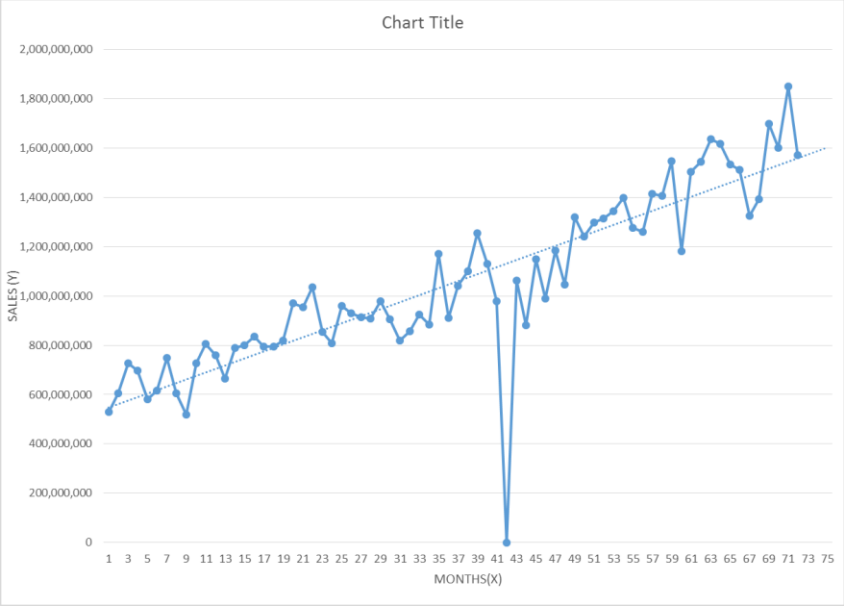

Fig 1: Time plot of the monthly data;

4.0 Analysis

Table 1: The least squares method

YEARS	Q1	Q2	Q3	Q4	SUM
2007	1.076	1.018	0.942	1.083	
2008	1.004	1.022	1.096	1.125	
2009	1.015	0.967	0.861	0.941	
2010	1.037	0.920	0.875	0.879	
2011	1.018	1.033	0.975	0.989	
2012	1.088	1.057	0.968	1.075	
TOTAL	6.238	6.017	5.717	6.092	24.064
AVERAGE	1.039666667	1.002833333	0.952833333	1.015333333	4.010666666
ADJUSTMENT	1.036901596	1.000166223	0.950299201	1.012632978	3.9999999998
INDEXES	103.6901596	100.0166223	95.0299201	101.2632978	

Table 2: Moving average method

YEARS	Q1	Q2	Q3	Q4	TOTAL
2007	-	1.008635573	0.927628257	1.071102635	
2008	0.970138749	0.980380461	1.012284604	1.043936143	
2009	0.982811317	1.022431714	0.933675539	0.991538847	
2010	1.073714586	0.97682436	0.981250575	0.949880717	
2011	1.040217166	1.024353467	0.984798233	0.970463048	
2012	1.042964412	1.015695688	0.939633277		
TOTAL	5.10984623	6.028321263	5.779270485	5.02692139	21.94435937
MEAN	1.021969246	1.004720211	0.963211747	1.005384278	3.995191205
Adjusted mean	1.023199335	1.005929538	0.964371113	1.006594405	4
Indexes	102.3199335	100.5929538	96.4371113	100.6594405	

Table 3: Exponential smoothing method

YEARS	QUARTER1	QUARTER2	QUARTER3	QUARTER4	TOTAL
2007	1.177631	1.154427	1.128199	1.134443	
2008	1.108934	1.097038	1.09245	1.069233	
2009	1.054086	1.034105	1.006105	1.00447	
2010	1.003789	0.976686	0.958581	0.947348	
2011	0.95265	0.942739	0.925015	0.91557	
2012	0.915059	0.90072	0.881517	0.881908	
MEAN	1.035358167	1.0176525	0.9986326667	0.992162	4.043805334
ADJUSTED MEAN	1.024142	1.0066286	0.9803307	0.9814142	3.9925155
INDEXES	102.4142	100.6629	98.0331	98.1414	

4.3 Results of Analysis

The adjusted quarterly index is, therefore, each mean multiply by the correction factor. Each of the means is adjusted downward so that the total of the four quarterly means is 4.00. Usually indexes are reported as percentages, so each value in the last row of the tables has been multiplied by 100.

From result of the Tables [1 – 3] above, it is observed that all the seasonal indexes are almost the same in value and also giving the same interpretation of the data which gives us the conclusion that GlaxoSmithKline Consumer Nigeria plc. Sales are high during the first quarter of the year (i.e. period prior to winter), decline or increase with a small percentage during the spring period and decline drastically during the summer and then go back to normal during the fall period.

In comparing the sum of square residuals below, the exponential smoothing method gave a better result for the analysis because it involves all the data, there are no data lost at the end and at the beginning of the series. Exponential smoothing method happen to be the best model for forecast because, it has the least sum of squares residual and the procedure also allow

the forecast function to be updated easily every time a new observation become available and they are easy to implement and quite effective.

Table 4: The sum of squares residual for the three models used in this analysis

Square of residuals on moving average method	Square of residuals on least squares method	Square of residuals on exponential smoothing method
1604.643	0.347651744	0.181683
79.08545	0.35520408	0.176033
5121.979	0.115281979	0.183946
151.1424	0.009643044	0.205485
2352.153	0.42405811	0.065744
130.0284	0.100834191	0.077825
10983.67	0.18481057	0.124527
251.9521	1.055499391	0.079425
3633.437	0.856087562	0.009593
172.3181	0.765887522	0.008384
9864.661	0.803911285	0.007654
251.9521	1.920971052	0.006337
5333.235	0.812478693	0.005126
216.7667	0.318953857	0.011488
13958.48	0.547230062	0.006272
298.1838	1.338107579	0.015913
6881.2	2.255433276	0.07244
361.8365	1.82750383	0.059767
22761.76	3.15808441	0.046259
491.1986	3.724409796	0.06764
10151.37	0.348931215	0.191493
479.3034	4.7095163	0.18611
28463.74	2.120192264	0.166957
725.8714	1.019112459	0.174399
124719.9579	29.11979427	2.130499333

Mean square residuals on least square method = $\frac{29.1197427}{22}$ = 1.323624668

Mean square residuals on moving average method = $\frac{124719.9579}{22}$ = 5669.088995

Mean squares residual on exponential smoothing method = $\frac{2.130499333}{22}$ = 0.09684087877

Forecast for the first quarter of year 2013 using the three different models created.

For least square method, the model is:

Y = 1602.455906+ 128.7993609x

and the forecast for the first quarter of 2013 is:

Y = 1602.455906 + 128.7993609(25)

= 4822.439929

The required sales = N4,822,439,929

For the moving average method, the model is:

Y = 1596.836648 + 129.6097348x

and the forecast for the first quarter is: Y = 1596.836648 + 129.6097348(25)

= 4837.080018

The required sales for the first quarter of the year 2013 is:4,837,080,018

For the exponential smoothing method, the model is:

Z = 3.486982491 + 0.06173042382x

The forecast is: Z = 3.486982491 + 0.06173042382(12.5)

= 4.258612789

Which implies that sales for the first quarter in 2013 is ₦4, 258,612,789

5.0 Conclusion

The, least squares method, moving average method, and exponential smoothing method were used to fit the data of GlaxoSmithKline Nigeria plc. Consumer sales income. From the analysis, it was deduced that the period prior to winter (the first quarter) is when GSK sales are high. After the winter (the spring period), the sales remains the same that is, they neither loose nor gain. GSK consumer sales decline after the spring period and then go back to normal during the fall period.

This analysis gives an overview on different methods and the best model for the forecast turns out to be the exponential smoothing method because it gave us the least sum of square residuals and also because the procedure allow the forecast function to be updated easily every time a new observation become available and they are easy to implement and quite effective.

5.1 Recommendation

GlaxoSmithKline Nigeria plc. With the help of the best model for forecasting, they would be able to have an estimated value of sales income for each quarter of the year; it will also help them to plan their expenditure and reduces over budgeting or under budgeting to some extent.

References

[1] Wood, W.A, Gray, H.L and Elliot, A.C (2012) Applied Time Series, New York – U.S.A. CRC Press Publication.
[2] Durblin. J and Koopman S.J (2001) Time Series Analysis by state space methods London, United Kingdom. Oxford university press. Pp21 – 123
[3] Boashash, B.A (2003) "Time frequency signal analysis and processing" Journal of the Royal Statistical Society.A142, pp18
[4] Shasha, D (2004) High performance Discovery in Time series,New York. U.S.A John Wiley & Son
[5] Enders. W. (1995) Applied Econometric Time Series. New York. John Wiley & Son pp221 – 300
[6] Bowerman. B and O "Conneli R.T (1993) Forecasting and Time series.
[7] Belmont, CA: Wadsworth. Pp437 – 486
[8] Cook T.D and Campbell D.T (1979) Quasi – Experimentation: Design and Analysis Issues for Field setting. Boston. U.S.A. Houghton – Miffin pp233-252
[9] Weigend A.S, Gershenfield, N.A (1994) "Time series Prediction: Forecasting the future and understanding past" Biometrika pp507-516
[10] Paul Jedamus and Robert Taylor (1991). "Statistical Analysis for Business Decision" New York. John Wiley & son
[11] David Groebner and Patrick Shannon (1994). "Statistical for Business and Economics" New York. John Wiley & son
[12] Cleveland, W.S (1972). "The inverse Autocorrelation of a Time Series and their Application" Technometrics, 14, pp277-293