Information and Knowledge Management
ISSN 2224-5758 (Paper) ISSN 2224-896X (Online)
Vol.7, No.10, 2017
www.iiste.org

Analysis of U.S. E-Commerce Sales Using Winters' Method

Zehai Zhou
Department of FEMIS, University of Houston-Downtown
One Main Street, Houston, TX 77002, USA

Abstract
More than three billion people around the world have the access to the Internet or around 40% of the world population has an Internet connection, and there are over one billion websites on the World Wide Web as of August 2016 (InternetLiveStas). While these numbers are truly impressive, there is no doubt that they will continue to grow rapidly in the foreseeable future. Business on the Internet has also flourished as a consequence. Electronic commerce or e-business, the process of buying, selling, transferring, or exchanging products, service, or information via computing networks, including the Internet, is positively one of the major driving forces for businesses of all sizes today. The impact of electronic commerce is indeed phenomenal or even revolutionary. The Internet and World Wide Web have changed our society substantially in general, and the ways in which companies and organizations conduct businesses, consumers buy, sell and exchange, and individuals work, communicate, entertain, get educated, and involve in many other activities in particular. The benefits and advantages of electronic commerce are being felt in many different and meaningfully ways. Nowadays a significant portion of businesses rely either heavily or in some cases solely on the revenues or returns generated from the electronic commerce division of their corresponding businesses. More and more companies and organizations of all sizes, are working to make a greater presence in the virtual world and conduct more e-commerce because of the importance of e-commerce to the success, or sustainability, or even the very survival of the organization as well as the advantages e-business brings. U.S. Census Bureau has being conducted survey on e-commerce sales since 1999 and estimated that the total electronic sales (defined as sales of goods and services where an order is placed by the buyer or price and terms of sale are negotiated over an Internet, extranet, Electronic Data Interchange (EDI) network, electronic mail, or other online system) in the U.S. This paper provides a time-series analysis of U.S. e-commerce sales for the period of 1999 - 2015 using Winters' Method. It also includes quarterly U.S. e-commerce sales projections for the period of 2015 - 2017.
Keywords: Electronic Commerce; E-Commerce; Time Series Analysis; Winters' method; Seasonality; Exponential smoothing

1. Introduction

Electronic Commerce or E-Commerce (EC) refers the process of buying, selling, transferring, or exchanging products, services, and/or information via computer networks, mostly the Internet and intranets (Turban *et al.* 2010). It includes digitally enabled commercial transactions between and among organizations and individuals (Turban *et al.* 2010). E-Commerce is sometimes defined from many other perspectives such as business processes, services, community, collaborative, learning, etc. Since its inception around 1995, E-Commerce sales have shown extraordinary growth worldwide. In spite of some turmoil experienced in the e-commerce business environment (e.g., the market crash in 2000-2001), E-Commerce has survived and has been reinvented twice. According to U.S. Census Bureau, E-Commerce sales was $393,003 million, accounted for 8.08 percent of total retail sales ($4,862,853 millions) in U.S. in 2016, compared to 7.24 percent (of $4,727,427 millions) and 6.48 percent (of $4,639,440 millions) in 2015 and 2014, respectively. It is widely accepted and commonly expected that E-Commerce will continue to grow with even more transitions in the future. As Laudon and Traver (2013) asserted, "the e-commerce revolution is just beginning."

In this paper, we attempt a time-series analysis of U.S. E-Commerce sales for the period of 1999 - 2015 using the Winters' Method, one of the forecasting models that take into considerations of trend as well seasonality. We include also quarterly U.S. e-commerce sales projections for the period of 2015 – 2017. The rest of this paper is organized as follows: Section 2 provides a brief literature review, Sections 3 describes the forecasting methods used, Section 4 contains the data analyses and major results, and finally, Section 5 completes the paper with summary, discussion and concluding remarks.

2. Brief Literature Review

There are many different kinds of forecasting techniques available and used for studying the patterns, trend, and making projections. Chambers et al. (1971) pointed out that there were three basic types - qualitative techniques, time series analysis and projection, and causal models and discussed how to choose the right forecasting technique. Another comprehensive survey was provided by Armstrong (2001). U.S. Census Bureau has collected/compiled U.S. E-Commerce quarterly sales data since 1999. Consequently for this study we narrow down to time series analysis and projection, which focuses entirely on patterns and pattern changes, and thus

relies entirely on historical data (Chambers et al. 1971). Some models for time series data may depend linearly on previous data points. Other models may deal with non-linear dependence of the level of a series on previous data. For more information on nonlinear time series analysis one can refer to Kantz and Schreiber (2003).

Exponential smoothing is a commonly used technique for smoothing time series data (Holt 1957). The particular model, the Winters' method, we use in this study belongs in the category of triple exponential smoothing techniques.

Ariguzo and White (2006) provided a historical foundation by documenting the evolution of E-Commerce models for the first decade (1995 -2005). White and Ariguzo (2011) tested the significance of the rate of growth of U.S. e-commerce for the period 2000-2010 and stated that their results indicated the existence of a significant linear trend model. They also offered total U.S. e-commerce sales projections for the 2011-2013 period. White and Ariguzo (2011) asserted that it was clear that the U.S. e-commerce market warrants further macro-level research. There are also other researches using time-series analysis or other techniques for some particular products or within different time frames. For examples, Dai at al. (2014) combine web search data and structure time series model to predict the women's clothing sales volume of Taobao, a popular shopping web site in China, Huang et al. (2015) use a novel trigger model for sales Prediction with data mining techniques, Akter and Wamba (2016) explore big data analytics in E-commerce.

3. Time Series Forecasting Methods

There are many time series forecasting methods. From moving average, simple exponential smoothing, exponential smoothing with trend, exponential smoothing with seasonality, to ad hoc forecasting, from simple linear regression, fitting nonlinear relationships, to multiple regression, to name a few. Each of them may have some special characteristics, advantages and disadvantages, and is suitable or appropriate for some applications. We make use of and focus on the Winters' model because initial exploration and plotting of the data reveal that U.S E-Commerce time series data exhibit some degrees of trend as well as seasonality.

In the following, we briefly review and describe the Winters' model and the forecasting errors and the accuracy criteria to be used. For more details on this technique, one may refer to Winston (2004) or Winters (1960).

3.1 Basic Notation

Let us define the following:

$X_1, X_2, \ldots, X_i, \ldots$ be the observed values of a time series, where X_i is the observed value of the time series at time period i.

F_i be the forecast or predicted value for period i.

$F_{i,1}$ be the forecast or predicted value for period i+1 made after observing X_i.

N be a positive integer.

p be the number of periods in the length of a seasonal pattern (e.g., p = 12 for monthly data and p = 4 for quarterly data).

3.2 Moving Average Method

The moving average method is one of the most commonly used and easiest methods for forecasting purpose. If the moving average method is utilized,

$F_{i,1}$ = average of the last N observed values

$$= (X_i + X_{i-1} + X_{i-1} + \ldots + X_{i-N+1})/N \quad (1)$$

3.3 Simple Exponential Smoothing

Simple exponential smoothing method is more suitable when a time series fluctuates around a base level. Let Y_i be the smoothed average of a time series after observing i. It is the forecast value of the time series during any future period.

$$Y_i = \alpha X_i + (1 - \alpha) Y_{i-1} \quad (2)$$

where α is used as a smoothing constant ($0 < \alpha < 1$) that smooths out fluctuations or variations in a time series by giving a weight α to the last observation and a weight of 1-α to past smoothed average. It can be shown that

$$Y_i = \alpha X_i + \alpha (1 - \alpha) X_{i-1} + \alpha (1 - \alpha)^2 X_{i-2} + \ldots + \alpha (1 - \alpha)^j X_{i-j} + \ldots \quad (3)$$

3.4 Holt's Method

For a time series that displays a linear trend without seasonality, Holt's method may be more appropriate and often results in better forecasts. Let B_i be the base level and T_i be the per-period trend of the series at the end of the i^{th} period.

$$B_i = \alpha X_i + \alpha (1 - \alpha) (B_{i-1} + T_{i-1}) \quad (4)$$

$$T_i = \gamma (B_i - B_{i-1}) + (1 - \gamma) T_{i-1} \quad (5)$$

where α, γ are used as smoothing constants ($0 < \alpha < 1$, $0 < \gamma < 1$).

3.5 Winters' Method

It is believed that if a time series exhibits a linear tread with seasonality, Winters' method often yields good forecasts. Let S_i be an estimate of a seasonal multiplicative factor (i.e., multiplier) for period i, obtained after observing X_i.

$B_i = \alpha X_i / S_{i-p} + \alpha (1 - \alpha) (B_{i-1} + T_{i-1})$ (6)

$T_i = \gamma (B_i - B_{i-1}) + (1 - \gamma) T_{i-1}$ (7)

$S_i = \delta X_i / B_i + (1 - \delta) S_{i-p}$ (8)

where α, γ, and δ are used as smoothing constants ($0 < \alpha < 1$, $0 < \gamma < 1$, $0 < \delta < 1$).

3.6 Forecasting Errors and Accuracy Criteria

Given a forecast for X_i, let us define E_i be the error in the forecast for X_i.

$E_i = X_i - (\text{forecast for } X_i) = X_i - F_i$ (9)

Many measures of forecast accuracy can be defined such as mean error, mean absolute error, mean square, root mean square error, mean percentage error, mean absolute percentage error, maximum absolute error, maximum absolute error, etc. Measures like mean error, mean percentage error are rarely used because the positive values and negative values cancel each other out. We chose mean absolute error (or mean absolute percentage error) as the goodness-of-fit measure.

4. Data Analysis and Results

4.1 Data Sets

U.S. E-Commerce sales data can be found at U.S. Census Bureau Web site or Federal Reserve Bank of St. Louis Web site. Both not seasonally adjusted and seasonally adjusted data of U.S. total (retail) sales and E-Commerce retail sales are available. U.S. Ecommerce retail sales as a percentage of the total sales can also be found at the Web sites as well. Because Winters' Method takes into account both tread and seasonality, we logically use the not seasonally adjusted data set displayed in Table 1.

Table 1. U.S. E-Commerce Sales

Year	Sales ($millions)			
	Quarter 1	*Quarter 2*	*Quarter 3*	*Quarter 4*
1999				5,266
2000	5,563	6,071	6,902	9,080
2001	7,889	7,822	7,759	10,808
2002	9,630	10,080	10,766	14,177
2003	12,369	12,990	13,926	17,927
2004	16,210	16,513	17,403	22,563
2005	20,179	20,985	22,218	28,167
2006	25,539	25,883	26,954	35,210
2007	30,484	31,682	32,444	42,246
2008	34,367	34,367	33,582	39,653
2009	32,353	32,991	34,554	45,872
2010	37,108	38,542	40,143	54,391
2011	44,217	45,353	45,997	64,138
2013	57,953	59,825	60,788	82,103
2014	65,517	68,858	70,351	93,530
2015	74,920	78,750		

4.2 Model Statistics

Some of the widely available software, commercial or free, can be used to conduct the data analysis for our purpose: Dell Statistica and IBM SPSS are just two notable examples. Table 2. provides the major model fit statistics. The R-squared is 0.997 and the sig. is 0.007, which indicates that the Winter's model fits extremely well. The Mean Absolute Error (MAE) and the Mean Absolute Percentage Error (MAPE) are 862.465 and 3.008 percent, respective, which again shows that the model we chose fares well with reasonable forecasting errors.

Table 2. Model Statistics

Model	Number of Predictors	Model Fit statistics						Ljung-Box Q(18)			Number of Outliers
		R-squared	RMSE	MAPE	MAE	MaxAPE	MaxAE	Statistics	DF	Sig.	
Model_1	0	.997	1325.038	3.008	862.465	13.163	5111.697	31.883	15	.007	0

4.3 Exponential Smoothing Model Parameters

Exponential smoothing model parameters are summarized in Table 3. The t values for Alpha (Level), Gamma (Trend), and Delta (Season) are 6.032, 2.468, and 3.997, respectively, and the corresponding sig. values are 0.000, 0.016, and 0.000, respectively. This reveals that U.S. E-Commerce sales exhibit (up) trend and seasonality.

Table 3. Exponential Smoothing Model Parameters

Model			Estimate	SE	t	Sig.
Model_1	No Transformation	Alpha (Level)	.484	.080	6.032	.000
		Gamma (Trend)	.390	.158	2.468	.016
		Delta (Season)	.999	.250	3.997	.000

4.4 Forecasts

The model statistics and the t values and the corresponding sig. values for the exponential smoothing model parameters indicate that Winters' methods fits the data very well and that the Winters' model that takes into account level, trend, and seasonality is expected to yield good forecasting results. U.S. E-Commerce sales projections for the next two years (8 quarters starting from Quarter 2 of 2015) after the last quarter (Quarter 2 2015) of the data set we used in model fitting are contained in Table 4. The actual U.S. E-Commerce sales (in millions) (U.S. Census Bureau) for the corresponding 8 quarters turned out to be $80,198, $107,433, $85,431, $90,397, $92,644, $122,515, $98,292, and $105,096, respectively, which all falls right in the range of [LCL, UCL] listed in Table 4. Figure 1. displays the U.S. E-Commerce sales from Quarter 4 of 1999 to Quarter 2 of 2015 and the forecasts yielded from the Winters' model from Quarter 3 of 2015 to Quarter 2 of 2017.

Table 4. Forecasts

Model		Q3 2015	Q4 2015	Q1 2016	Q2 2016	Q3 2016	Q4 2016	Q1 2017	Q2 2017
Model_1	Forecast	78990	105166	84308	88074	88074	116922	93476	97398
	UCL	81641	108680	88085	92737	94508	125807	101888	107083
	LCL	76340	101652	80531	83410	81640	108037	85065	87713

For each model, forecasts start after the last non-missing in the range of the requested estimation period, and end at the last period for which non-missing values of all the predictors are available or at the end date of the requested forecast period, whichever is earlier.

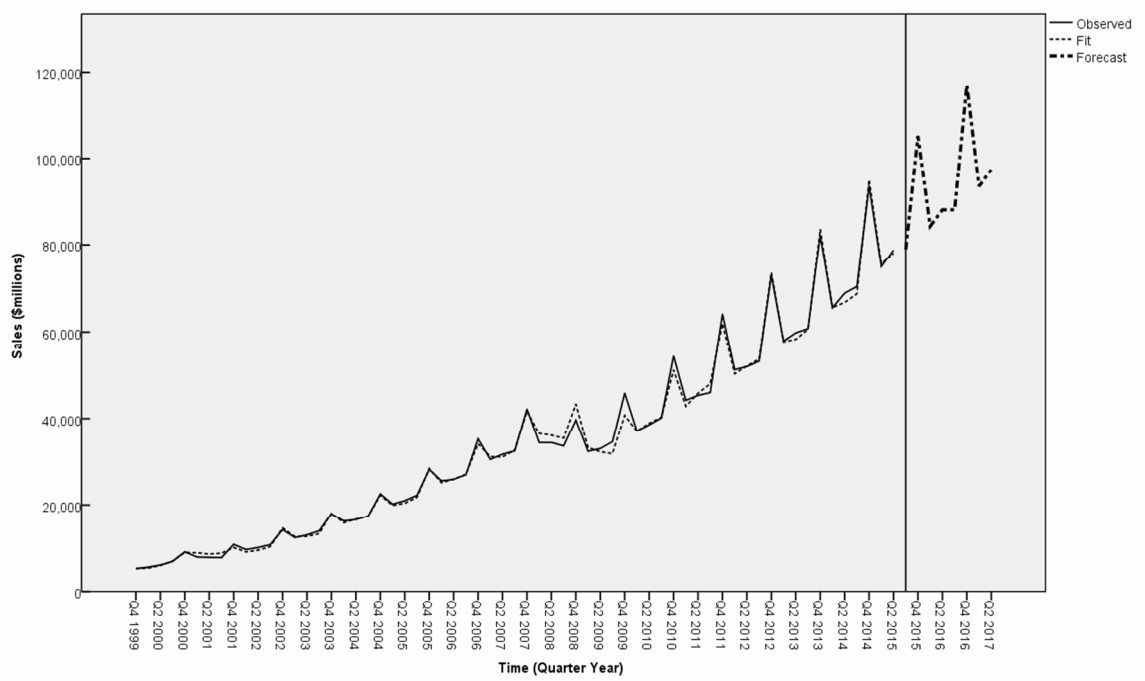

Figure 1. U.S. E-Commerce Sales (and Forecasts) 1999-2015 (2015-2017)

5. Summary and Concluding Remarks

In this paper, we review one of the forecasting models that take into considerations of level, trend as well seasonality, i.e., the Winters' method. We also conduct a time-series analysis of U.S E-Commerce retail sales for the period of 1999 - 2015 using the Winters' Method. The paper also includes quarterly U.S. E-Commerce sales forecasts for the period of 2015 – 2017. We argue that E-Commerce has played an important role in U.S economy and is expected to continue to grow with even more transitions and gain significance in the foreseeable future. As the importance of e-commerce to the success, or sustainability, or even the very survival of the organizations increases as well as the advantages e-business brings, it is of interest or necessary to study the trend and make reasonably accurate and practically useful projections or forecasting for U.S E-Commerce sales. Thought this study and especially by the goodness-of-fit measures resulted from this analysis Winters' method is shown to be robust because it takes into accounts level, trend, and seasonality. It is therefore recommended as a viable and practical approach that can be used to analyze and project U.S. E-Commerce sales. Future work includes investigation of E-Commerce in the global market or for some important products, alternative methods to analyze and forecast U.S. or global E-Commerce activities, etc.

Acknowledgement

The author would like to thank the anonymous reviewers for their helpful and insightful comments and suggestions that improve this paper in meaningful ways.

References

Akter, S. & Wamba, S.F. (2016), "Big data analytics in E-commerce: a systematic review and agenda for future research", Electron Markets 26(2): 173-194. https://doi.org/10.1007/s12525-016-0219-0.

Armstrong, J.S. (2001), "Selecting Forecasting Methods." In Principles of Forecasting: A Handbook for Researchers and Practitioners (Ed. J. Scott Armstrong). Kluwer.

Ariguzo, G.C., Mallach, E.G. & White, D.S. (2006), "The First Decade of E-Commerce," International Journal of Business Information Systems 1(3): 239-255.

Chambers, J. C., Mullick, S. & Smith, D.D. (1971), "How to choose the right forecasting technique," Harvard Business Review 49: 45-71.

Dai, Wei, Peng, Geng, Liu, Ying & Li, Shuaipeng. (2014), "A Prediction Study on E-commerce Sales Based on Structure Time Series Model and Web Search Data", Proceedings of IEEE 26th Chinese Control and Decision Conference (CCDC), 5346-5351.

Dell, Statistca. Availabe at: http://statistica.io/products/compare-versions/.

Federal Reserve Bank of St. Louis, Quarterly Retail E-Commerce Sales. Available at: https://fred.stlouisfed.org/release?rid=179.

Holt, Charles C. (1957), "Forecasting Trends and Seasonal by Exponentially Weighted Averages", Office of Naval Research Memorandum 52. reprinted in Holt, Charles C. (2004), "Forecasting Trends and Seasonal by Exponentially Weighted Averages," International Journal of Forecasting 20 (1): 5-10. doi:10.1016/j.ijforecast.2003.09.015.

Huang, W., Zhang, Q., Xu, W., Fu, H., Wang, M. & Liang, X.,(2015), "A Novel Trigger Model for Sales Prediction with Data Mining Techniques," Data Science Journal, 14(15), 1–8. DOI: http://dx.doi.org/10.5334/dsj-2015-015.

IBM, SPSS Software. Available at: http://www-01.ibm.com/software/analytics/spss/.

InternetLiveStas, Internet Users. Available at: http://www.internetlivestats.com/internet-users/.

Kantz, H. & Schreiber, T. (2003), Nonlinear Time Series Analysis. Cambridge, UK: Cambridge University Press.

Laudon, K.C. & Traver, C.G. (2013), E-Commerce: Business, Technology, Society, 9th ed. Upper Saddle River, NJ: Prentice Hall - Pearson.

Turban, E., King, D, Lee, J., Liang, T.P. & Turban, D. (2010), Electronic Commerce: A Managerial Perspective. Upper Saddle River, NJ: Prentice Hall - Pearson.

U.S. Census Bureau, Monthly & Annual Retail Trade. Available at: http://www.census.gov/retail/index.html#ecommerce.

White, D.S. & Ariguzo, G.C. (2011), “A TIME-SERIES ANALYSIS OF U.S. E-COMMERCE SALES,” Review of Business Research 11(4): 134-140.

Winston, W.L. (2004), Introduction to Probability Models, 4th ed. Belmont, CA: Brooks/Cole - Thomson Learning.

Winters, P. R. (1960). "Forecasting Sales by Exponentially Weighted Moving Averages," Management Science 6 (3): 324–342. doi:10.1287/mnsc.6.3.324.