A Study of Models for Forecasting E-Commerce Sales During a Price War in the Medical Product Industry

Pei-Hsuan Hsieh(✉)

National Cheng Kung University, Tainan 701, Taiwan
peihsuan@mail.ncku.edu.tw

Abstract. When faced with a price war, the accuracy of forecasting sales in e-commerce greatly influences an enterprise's or a retailer's merchandise inventory strategies. When faced with a price war, an enterprise might obtain certain consumption patterns by analyzing previous sales data. This case study research was conducted in collaboration with a medical product company to explore which of the various forecasting models can better inform a company's inventory plan. The study used the company's data from Amazon.com regarding sales volume, number of views, company ranking, etc. between February 7 2016 and March 28 of 2018. Three potential methods of data mining were selected from the literature: the exponential smoothing method, the linear trend method, and the seasonal variation method. Of these, the most suitable was identified for price war situations to forecast the sales volume for April 2018 and to provide concrete information for the company's inventory plan. The results showed that the seasonal variation method is more suitable than the other two sales forecasting methods. To obtain a more accurate sales forecast during a price war, the seasonal variation method is recommended to be used in the following approaches: Adjust the seasonal index by using a simple moving average. Remove the seasonal index from the sales volume, and conduct a regression analysis using the data within the last month. The resulting predicted value (with the seasonal index removed) should be multiplied by each period's corresponding weighted moving average to obtain a more accurate sales forecast during a price war.

Keywords: E-commerce · Price war · Sales forecasting · Inventory plan

1 Introduction

In the e-commerce-driven market today, the preparation for inventory forecast is crucial. Without proper inventory plans, retailers would risk customer disappointment due to merchandise shortage, which causes income loss. However, excess inventory would lead to storage and removal problems, causing surplus-induced increased cost and affecting the overall profit [12]. Situations like these are especially serious for seasonal merchandise that are put through price wars if an enterprise cannot effectively use prior experience to forecast future sales.

F. F.-H. Nah and K. Siau (Eds.): HCII 2019, LNCS 11588, pp. 3–21, 2019.
https://doi.org/10.1007/978-3-030-22335-9_1

Using Amazon as an example, one sees that when two leading companies that sell similar products intend to dominate the market, they usually slash prices as a competitive strategy against each other. As soon as one company begins to lower the price to attract customers, the other will see changes in the incoming orders and will fall into a predicament if it does not respond immediately to the competitor's move. Under such circumstance, traditional forecast models are no longer effective, and new models are needed. To identify more accurate merchandise sales forecast models when faced with a price war, this study looked at some conventional forecast methods, namely, the exponential smoothing method, the linear trend method, and the seasonal variation method.

The exponential smoothing method discards no prior data; more weight is given to the more recent data and less weight to the data from the more distant past [11]. For most data, this is a very suitable and accurate forecast model. However, exponential smoothing is more suitable for forecasting short-term data and less so for long-term data [8]. Therefore, we use the linear trend method to observe the trend of an entire period and the direction of future periods' data, in order to forecast the sales volume of future periods. The seasonal variation method accounts for any notable seasonality of sales. It removes the seasonal factor prior to applying the regression model, then adds it back afterwards [36].

Overall, the study analyzed and forecasted sales data using these three models that are suitable for price wars and aimed to help enterprises accurately calculate the potential sales volume during a price war in order to more effectively control their inventory costs and to increase their overall profits.

2 Literature Review

2.1 E-Commerce in Taiwan and in the United States

E-commerce development is still flourishing in Taiwan. In 2018, the top four sales volumes in the market belonged to Shopee, PChome, ibon mart, and momo. Shopee provides customers shipping fee subsidy, charges no slotting fee or processing fee, and offers a better and more comprehensive user interface. Customers can also chat with the seller in real time, to obtain more information about the products. In addition, Shopee combines different modes of business operations: B2B2C (Shopee Mall), B2C (Shopee), and C2C (Shopee Auction). It is a one-stop shop where users searching for products are presented with items from all three modes to choose from. Compared with the other e-commerce companies operated in Taiwan, Shopee provides a better consumer experience that attracts more sellers and buyers. It has become the largest mobile e-commerce entity in Taiwan since 2017.

PChome has B2C (PChome24), B2B2C (PChome Store), and C2C (Ruten.com, an auction site). Although its user experience is not on par with Shopee, it has a strong shipping system that supports the speed of product delivery. Ibon mart has a network of over 5000 brick-and-mortar stores throughout Taiwan, which provides a comprehensive set of physical-virtual convenience services.

The ibon mart website integrates products from supermarkets and investment companies. According to quarterly financial reports, the overall performance of Q4 are better than the other quarters. This is possibly due to holidays or the Single's Day (November 11) that originated in Mainland China. Momo is a shopping network invested by digital media enterprises. It also has a relatively better Q4 performance.

Although e-commerce originated in the United States, the performance of the US e-commerce has been lukewarm compared to that of Taiwan. Other than the top-earning e-commerce company Amazon, American e-commerce has not had the same level of vigor as that of China. Amazon is currently one of the largest web-based retailers, and it also has a higher performing Q4 compared to the other quarters. The fourth quarter is a major shopping season in the US and is the busiest season for e-commerce. The seasonality is due to the holidays, starting with Halloween in October, Thanksgiving in November and the subsequent Black Friday and Cyber Monday, followed by Christmas in December. The customers of Amazon.com have shown a consistent level of satisfaction, which is higher than other enterprises such as eBay, Walmart, Best Buy, etc. [31]. In recent years, the US has been influenced by the Chinese e-commerce giant Alibaba and has started to promote sales in the trend of celebrating the Singles' Day (November 11) in the last quarter [5].

2.2 The Importance of E-Commerce

In the past, companies that wanted to run a promotion activity needed weeks or even months to plan, forecast, and calculate carefully the sales volume and target profit. Their methods usually involved buying ads and releasing coupons [28]. In the information age, in the global e-commerce market, it is pivotal to plan the inventory and the logistics based on accurate sales forecast [34]. In the age of e-commerce, sales forecast and inventory control are essential. E-commerce promotion projects can immediately push the newest information out to the relevant consumer groups through emails, social media, direct broadcast, etc. [9]. At any time, sales data should be collected and analyzed; a good forecast can enable a company to ensure that there never is a shortage of merchandise that disappoints customers and leads to profit loss [27]. On the other hand, overstocking not only increases the inventory cost that affects the final profit but also increases labor and processing costs. Price discounts might be needed for inventory clearance. Case in point, the demand for seasonal merchandise will rapidly disappear once the season is over.

The start of a price war implies the reduction of profit. Under the strict control of cost, the product quality may be affected. This leads to a vicious cycle that negatively impacts future sales [26]. Price wars usually start when there are more competitors. Price adjustment is a sales strategy that can be executed easily [15]. The literature shows that enterprises often applies price wars to compete for market share and increase product sales volume when dealing with seasonal merchandise around the various annual holidays [26]. This is especially evident when the target consumer group is highly sensitive to price [25]. Additionally, seasonal products have different attributes depending on whether they are based on daily, weekly, or annual cycles; each has a different level of impact on sales volume [22]. However, one should consider the necessary inventory cost prior to a particular holiday, consumer satisfaction, willingness to re-purchase, and other factors [27]; otherwise, the earnings resulted from promotions may not fully cover the hidden costs, causing an overall loss [32]. Similarly, the competitors do not wish to lose their existing market share and will begin myriad promotion activities including the use of sales representatives' experience-based predictions, but they still need to analyze effectively the distribution point's primary consumer's sale volume [14].

2.3 Methods of Analysis and Effectiveness of E-Commerce Big Data Forecast

Not all historical data are relevant to the current consumers' behaviors. Companies should adjust their current systems and use a large quantity of long-term sales forecast data. They should let the systems consider the historical data appropriately and distinguish the useful data from the obsolete ones, and be able to tell based on current purchase behaviors and models which aspects have remained stable and which have changed and become unstable [10]. The procedure involving big data includes data collection and data processing and analysis, each processing step having an impact on the quality of the big data [1]. During data collection, the data source will affect big data's quality including its veracity, completeness, consistency, accuracy, and security [18]. Big data analysis methods can be categorized as descriptive, predictive, and normative [4]. The purpose of descriptive analysis is to accurately predict what will happen in the future and to provide a rationale for why it will happen. Predictive analysis uses data and mathematical models to confirm and evaluate targets. Normative analysis uses optimization or A/B testing to offer suggestions to staff or managers [33].

Quick response forecasting (QRF) of big data is accepted as more predictive than the traditional point of sale (POS) system [20]. If the expected demand starts to exceed the "most likely scenario," QRF can help adjust the purchase volume to effectively decrease costs. Amazon collects big data such as individual users' habits, the prices of competitors during price wars, consumers' product preferences, records of orders, profit margin, etc. Amazon uses big data to adjust 2.5 million products' prices each day, thereby attracting more customers and resulted in a 25% increase in annual profit [2]. Overall, the key to sales forecast accuracy is the calculation behind the forecast system, since it ensures that individual distributors have enough inventory to meet the demand of potential orders [30] and are able to devise more accurate purchase strategies [1].

3 Methodologies

Utilizing appropriate sales forecast methods is important for enterprises in e-commerce. Based on the previously mentioned literature, a firm grasp of customer types and needs as well as business historical data trends are necessary for choosing better forecasting methods. Common e-commerce sales forecast methods include: naïve method (time series with steady changes, seasonal variations, trend patterns), average method (moving average, weighted moving average, exponential smoothing, double exponential smoothing, triple exponential smoothing), trend-adjusted exponential smoothing, and seasonal variations [3, 7, 17, 23].

This study's data source was a medical equipment seller on the Amazon e-commerce platform (called here the Case Company). The Case Company often conducts sales analysis and forecast for the products they offer. However, during the time of this study, the company suffered the plight of a price war, which led to unstable sales and caused the supply to severely outnumber the demand.

The Case Company provided the current research with sales data including the sales volume data obtained through the sales platform's backend services, the number of views, competitors' sales data obtained via text mining methods, etc. Because the price war for which real sales data could be collected had occurred recently (Fig. 1) and the duration was not very long, three suitable methods for this research was explored: exponential smoothing method, linear trend method, and seasonal variation method. Among these three, the most appropriate method for responding to a price war was identified. The exponential smoothing method was included to investigate price wars due to its ability to focus on recent observation periods while taking into consideration the characteristics of observations in the more distant past. The linear trend method was included because it considers all data and directly reflects the trend of sales. The seasonal variation method was included because there was a noticeable seasonality in the data. We therefore included this method to see if higher accuracy may be achieved by removing any seasonal factors prior to forecasting. This study analyzed the sales data (collected between 2/7/2016 and 3/28/2018) using the three methods to forecast the daily sales volume during the month following the data collection period (4/1/2018 to 4/30/2018).

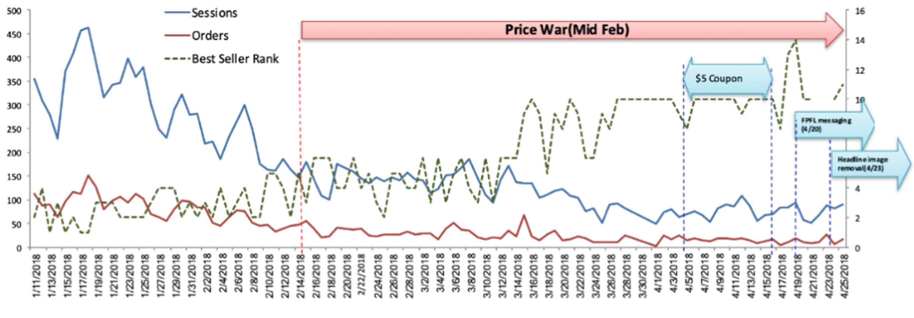

Fig. 1. The estimated period of the price war.

3.1 Exponential Smoothing Method

Smoothing value is the weighted average of forecast values and actual values. In other words, the next forecast value is the weighted average of the previous period's forecast value and its actual value. The smoothing factor (α) plays an important role. It is usually determined based on the products' characteristics and the manager's understanding of the market's sales situations along with any prior experiences with forecasted values. The closer the smoothing factor gets to 1 means that the influence of past observations decreases steeply as it was further in the past; conversely, a factor closer to 0 means that such influence decreases not as steeply. Thus, when the time series is relatively stable, one may choose a larger α, whereas a smaller α should be chosen when the time series contains more fluctuations, as to not ignore the influence of observations from earlier times. In the equation, each period's forecast value requires prior data. Therefore, the very first smoothing value must be defined. There are several ways to do that. If historical data is available from further back in time, then an overall period average from the historical data can be used as the first smoothing value. Otherwise, the first actual value may also be used as the first smoothing value as well as the second forecast value.

3.2 Linear Trend Method

This method uses time as the independent variable (x) and the function of time as the dependent variable (y) and assumes that the independent variable (time) and the forecasted function of time are linearly related. Therefore, using this method requires first a calculation of the correlation coefficient (r) between the two variables. A coefficient of 0 means no linear correlation exists and the linear trend method is not applicable. A coefficient between ± 0.2 and ± 0.3 is a weak correlation; ± 0.6 is a moderate correlation; ± 0.8 is a strong correlation. A coefficient of 1 is a perfect correlation. When the coefficient r is near 1, the linear trend method can be used to analyze the data. When r is close to 1, the linear equation in the form of $y = a + bx$ can be used to predict future changes. Based on historical data, regression can be used to find the values of a and b, thereby finding the linear equation in which the independent variable can be plugged into to produce the forecasted value of the dependent variable y.

3.3 Seasonal Variation Method

As many products' sales are cyclical or seasonal, the seasonal variation method adds to the time series forecast methods the factors that recur periodically and fluctuate regularly based on sale seasons. This enables forecasting the sales of products that are seasonal in nature. First, statistics such as the simple average method are used to compute the forecast seasonal index. Patterns of seasonal changes are also identified. These are then used to forecast the values that are sought after.

In summary, because the Case Company was faced with a price war during a short amount of time, this study used more basic methods for forecasting. Moreover, when it was uncertain whether the product had any seasonality, the linear trend method was first used to find a regression equation that predicts the sales volume; if the data's pattern was affected by the seasonality of the products, then the seasonal variation method would be added as the next step in conjunction with simple smoothing methods (simple moving average, weighted moving average, exponential smoothing) to adjust for the seasonal index.

4 Results

4.1 Application of the Exponential Smoothing Method

The most important step when using the exponential smoothing method is finding the most appropriate smoothing factor (α). We first set the α between 0.1 and 0.9, then fine-tuned the α to be between 0.01 and 0.09 to analyze the data collected between 2/7/2016 and 3/28/2018 in order to predict the sales volume for April 2018. With each month being a period, there were a total of 30 periods usable for forecasting sales volume for just one month (April 2018). Before using any equations under this method, we set the sales volume of the forecast period to be 0. The difference between the forecast value and the actual value was squared for each of the 30 periods, then the 30 squared differences were summed. The results of the equations are listed in Table 1. When $\alpha = 0.1$, the month's overall forecast error was the smallest, especially when $\alpha = 0.06$, the sum of error was the least of all. Then, comparing the April 2018's sum of actual values against the sum of forecast values, it was confirmed that when α was 0.05 and 0.06, the sum of the predicted values was 460 and 374 (only off by 39 and 47), respectively, which were very close to the actual values. On the contrary, when α was set to 0.04 and 0.07, the differences were larger than 100.

Next, each day was treated as a period and was used to forecast the next day (i.e., on April 1 we predicted the sales of April 2, and on April 2 we predicted the sales on April 3, etc.). The actual data from April 2018 was used instead of the predetermined 0. Setting the α value again between 0.1 and 0.9, it was found that $\alpha = 0.2$ yielded the least sum of error. All other α values also yielded smaller errors compared to when computations were based on monthly periods. It is evident, then, that a daily period yields higher accuracy than a monthly forecast and has more tolerance for α errors. Therefore, when facing a short-term price war, the forecast analyst can use this method to avoid severe forecast inaccuracy caused by mistakes in setting the α.

4.2 Application of the Linear Trend Method

A longer-term price war implies that the forecast period should also be elongated. In this case, the exponential smoothing method is not suitable due to the difficulty of choosing the appropriate α value; hence, a different method is needed. Here, we tested the applicability of the linear trend method. A regression analysis of the data between 2/7/2016 and 3/28/2018 yielded the linear equation $y = -0.0076x + 53.311$, meaning, the Case Company's sales were declining at a rate of -0.0076. Then, this equation was used to forecast the upcoming month's values (between 4/1/2018 and 4/30/2018). The daily differences between actual and forecast sales were squared, and the sum of the squares was 34068. The sum of the predicted values for April 2018 was 1417, which was vastly different from the month's actual sales total of 421 (Table 2). Furthermore, Fig. 2 shows that the data contained a seasonal pattern. The linear trend method was unable to respond to such seasonality, making the forecast sales volume noticeably different from the actual values.

Table 1. Applying the exponential smoothing method to find the most suitable α value.

			Alpha:	0.01	0.02	0.03	0,04	0.05	0.06	0.07	0.08	0.09
KPI	Week	Units Ordered	T=	784	784	784	784	784	784	784	784	784
4/1/2018	112	3		45.81978	42.37974	37,42327	32.84675	29.26438	26.61914	24.6869	23.2548	22,16333
4/2/2018	112	25		45.36158	41.53215	36.30058	31.53288	27.80116	25.02199	22.95882	21.39441	20.16863
4/3/2018	112	17		44.90797	40.7015	35.21156	30.27157	26.4111	23.52067	21.3517	19.68286	18.35345
4/4/2018	113	24		44.45889	39.88747	34.15521	29.0607	25.09055	22.10943	19.85708	18.10823	16.70164
4/5/2018	113	14		44.0143	39.08972	33.13056	27.89828	23.83602	20.78287	18.46709	16.65957	15.1985
4/6/2018	113	18		43.57416	38.30793	32.13664	26.78234	22.64422	19.5359	17.17439	15.32681	13.83063
4/7/2018	113	14		43.13841	37.54177	31.17254	25.71105	21.51201	1S.3G374	15.97219	14.10066	12.53587
4/24/2018	115	7		36.36323	26.62919	18.57341	12.8449	8.994608	6.414085	4.651301	3.416903	2.532719
4/25/2018	116	17		35.9996	26.09561	18.01621	12.33111	8.544878	5.029239	4.32571	3.143551	2.304775
4/26/2018	116	11		35.6396	25.57468	17.47572	11.83786	8.117634	5.667485	4,02291	2.892066	2.097345
4/27/2018	116	16		35.28321	25.06318	15.95145	11.35435	7.711752	5.327435	3.741307	2.650701	1.908584
4/28/2018	116	6		34.93033	24.56192	16.44291	10.90977	7.326165	5.00779	3.479415	2.447845	1.736811
4/29/2018	116	9		34.58107	24.07068	15.94962	10.47338	5.959855	4.707322	3.235856	2.252017	1.580498
4/30/2018	116	14		34.23526	23.58927	15.47113	10.05445	6.611864	4.424883	3.009346	2.071856	1.438253
			SSD	20834.62	11158.62	5161.545	2530.183	1724.355	1704.598	1977.197	2336.242	2701.601
			Min	V								
SUM:	421			1193	963	747	580	460	374	313	267	232

Table 2. Applying the linear trend method to compute the sum of predicted sales volume (SUM) and the sum of squared differences (SSD).

Unit (x)	Date	Sales volume (v)	Forecasting model (yi = −0.0076x +53.311)
782	4/1/2018	3	47.3678
783	4/2/2018	25	47.3602
784	4/3/2018	17	47.3526
785	4/4/2018	24	47.345
786	4/5/2018	14	47.3374
787	4/6/2018	18	47.3298
788	4/7/2018	14	47.3222
789	4/8/2018	12	47.3146
790	4/9/2018	18	47.307
791	4/10/2018	18	47.2994
792	4/11/2018	16	47.2918
793	4/12/2018	19	47.2842
794	4/13/2018	14	47.2766
795	4/14/2018	9	47.269
796	4/15/2018	13	47.2614
797	4/16/2018	17	47.2538
798	4/17/2018	5	47.2462
799	4/18/2018	11	47.2386
800	4/19/2018	18	47.231
801	4/20/2018	10	47.2234
802	4/21/2018	9	47.2158
803	4/22/2018	10	47.2082
804	4/23/2018	27	47.2006
805	4/24/2018	7	47.193
806	4/25/2018	17	47.1854
807	4/26/2018	11	47.1778
808	4/27/2018	16	47.1702
809	4/28/2018	6	47.1626
810	4/29/2018	9	47.155
811	4/30/2018	14	47.1474
			SUM = 1417.728
			SSD = 34068.30735

4.3 Application of the Seasonal Variation Method

Because the data showed seasonal patterns, it was not appropriate to use the linear trend method to produce a linear equation. Therefore, the seasonal variation method was used. Additionally, different moving or smoothing methods were used to adjust for the seasonal index: simple average, simple moving average, weighted moving average, and exponential smoothing.

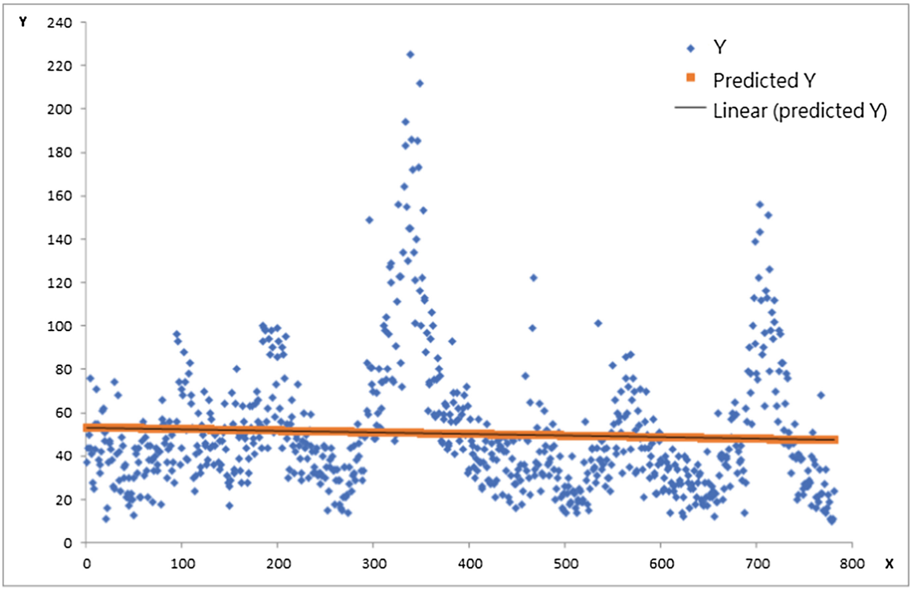

Fig. 2. The data shows a clear seasonal pattern.

Table 3. Applying the simple average method to compute the predicted sales volume.

Unit (x)	Date	Sales volume	ASI	SA (y)	Forecasting Model (yi = −0.0005x + 1.1999)	Predicted sales volume
1	2016/2/7	37	63.333	0.584210526	1.199342788	75.95837656
2	2016/2/B	44	68.333	0.643902439	1.198831652	81.92016292
3	2016/2/9	50	60.333	0.828729282	1.1983205117	72.2986712
4	2016/2/10	76	60.333	1.259668508	1.197809382	72.2678327
5	2016/2/11	44	56.333	0.781065089	1.197298246	67.44780121
6	2016/2/12	41	40.000	1.025	1.196787111	47.87148444
7	2016/2/13	28	38.000	0.736842105	1.196275976	45.45848708
8	2016/2/14	25	42.667	0.5859375	1.19576484	51.01929985
9	2016/2/15	43	49.667	0.865771812	1.195253705	59.36426735
10	2016/2/16	55	58.000	0.948275862	1.19474257	69.29506904
11	2016/2/17	71	57.667	1.231213873	1.194231434	68.86734604
12	2016/2/18	55	39.333	1.398305085	1.193720299	46.95299842
13	2016/2/19	38	35.000	1.085714286	1.193209164	41.76232072
14	2016/2/20	32	39.667	0.806722689	1.192698028	47.31035512
15	2016/2/21	36	45.000	0.8	1.192186893	53.64841018
16	2016/2/22	42	45.333	0.926470588	1.191675757	54.02263434
17	2016/2/23	61	64.333	0.948186528	1.191164622	76.63159069
18	2016/2/24	62	52.000	1.192307692	1.190653487	61.91398131

Simple Average Method. First, each date was converted to an ordinal number such that January 1 is Day 1 and January 2 is Day 2, etc. The mean of the data for the same day of every year was computed; this was the average seasonal index (ASI) for that day. For example, to compute the ASI of Day 37, compute: [(Day 37 of 2016) + (Day 37 of 2017) + (Day 37 of 2018)]/3. Then, the daily sales volume between 2/7/2016 and 3/28/2018 was divided by each day's ASI, which resulted in a set of y-value that has been adjusted for seasonality. The seasonally adjusted (SA) data was analyzed through linear regression, which resulted in the equation $y = -0.0005x + 1.1999$. The final predicted sales volume was computed by multiplying the seasonally adjusted predicted value by each period's ASI (Table 3). The sum of squared differences between the predictions and the actual sales of each day during April 2018 was large (SSD = 11221). Compared to the actual total sales of 421, the predicted sales of 914 was off by quite a bit. This was possibly due to the limited number of years for which data was available, where each period only had two or three prior data to be averaged. If one period had an outlier, the ASI would be severely skewed and unable to properly reflect the true seasonality.

Simple Moving Average Method. Asimple moving average was calculated using the ASI of 10 periods, which was found to be optimal after some calculations. The goal was to smooth out the short-term fluctuations and to reflect any long-term trends or periodicities. For example, to compute the 10-period simple moving ASI for Day 88, every year's ASIs for Days 79-88 will be simply averaged. To calculate the simple 10-period moving ASI for 4/2/2018, the ten ASIs between 3/23/2018 and 4/2/2018 were averaged. Then, the daily sales volumes between 2/7/2016 and 3/28/2018 were divided by the 10-period simple moving ASIs to obtain the seasonally adjusted values. A regression analysis of these values produced the equation $y_i = -0.0005x + 1.2019$. Then, the seasonally adjusted predicted values were multiplied by their corresponding 10-period simple moving ASI to generate the final predicted sales volume (Table 4).

The resulted forecast was considerably larger than the actual sales data of April 2018. The sum of squared differences of actual daily sales values was still relatively large (SSD = 7383). The sum of the predicted values for the month was 849, which was much larger than the actual sales total of 421. It was possible that when the price war started in February 2018, the residuals in the regression model started to deviate from the regression line that was based on the data from the 2/7/2016–3/28/2018 timeframe. Using a 10-period simple moving ASI as an example, we saw that the sum of residuals was 0.557114352 in January 2018, −1.085087416 in February 2018, and −3.474345163 in March 2018.

Consequently, the next step was to analyze only the data during the price war (i.e., the data in March 2018). The seasonally adjusted values (y) were obtained when dividing each day's (row's) sales volume by its corresponding 10-period simple moving ASI. (It can be observed from Table 5 that using 10-period data for simple moving averages yielded the optimal data for this study.) Then, these seasonally adjusted values were analyzed using regression, resulting in the equation $y_i = -0.0099x + 8.2569$. The equation shows that the Case Company's sales volume decreased at a noticeably faster rate once the price war began. Multiplying the seasonally adjusted predicted values by each period's corresponding moving ASI, the final

Table 4. Applying simple moving average method to compute the predicted sales volume.

Unit (x)	Date	Sales volume	ASI {10}	SA (y)	Forecasting model (yi = −0.0005x + 1.2019)	Predicted sales volume
1	2016/2/7	37	74.13333333	0.498764323	1.197769261	88.85451635
2	2016/2/8	44	75.16666667	0.585365854	1.197251898	89.99343432
3	2016/2/9	50	73.6	0.679347826	1.196734534	88.07966174
4	2016/2/10	76	71.03333333	1.069920225	1.196217171	84.97129306
5	2016/2/11	44	67.16666667	0.655086849	1.195699808	80.31117042
6	2016/2/12	41	61.71666667	0.664326222	1.195182444	73.76267653
7	2016/2/13	28	56.36666667	0.496747487	1.194665081	67.33928841
8	2016/2/14	25	54.28333333	0.460546515	1.194147718	64.82231861
9	2016/2/15	43	54	0.796296296	1.193630354	64.45603914
10	2016/2/16	55	53.7	1.024208566	1.193112991	64.07016762
11	2016/2/17	71	53.13333333	1.336260979	1.192595628	63.36658102
12	2016/2/18	55	50.23333333	1.094890511	1.192078264	59.88206482
13	2016/2/19	38	47.7	0.796645702	1.191560901	56.83745498
14	2016/2/20	32	45.63333333	0.701241782	1.191043538	54.35128677
15	2016/2/21	36	44.5	0.808988764	1.190526174	52.97841476
16	2016/2/22	42	45.03333333	0.932642487	1.190008811	53.59006346
17	2016/2/23	61	47.66666667	1.27972028	1.189491448	56.69909235
18	2016/2/24	62	48.6	1.275720165	1.188974084	57.78414051

predicted sales volume values were generated. Lastly, when comparing the forecast against the actual sales of April 2018, the sum of squared differences was relatively small (SSD = 901). Furthermore, the total sum of predicted sales was 442, which was very close to the actual total of 421.

Table 5. Finding the optimal period length for the simple moving average method.

Period	5	7	10	13	15
Total sales volume for the month	575	501	441	384	347
SSD	1383.2	1106	900.77	931.52	1112.6

Weighted Moving Average (WMA) Method. Because ASI data may be off compared with real data, it is advisable to use a 3-period ASI to calculate weighted moving averages in order to smooth out short-term fluctuations and to reflect any long-term trends or periodicities. For example, to calculate the weighted moving ASI for 4/1/2018, the ASI was weighted at 0.2 for 3/27/2018, at 0.3 for 3/28/2018, and 0.5 for 4/1/2018 (Note: No data retrieval activity occurred on 3/29-30/2018 due to staff errors). The three ASIs were then summed. Next, the daily sales volumes between 3/1/2018 and 3/28/2018 were divided by their corresponding weighted moving ASIs to obtain the seasonally adjusted value (y). Then, these y-values were analyzed through regression and resulted in the equation $y_i = -0.0042x + 3.9115$. Each seasonally adjusted predicted value was multiplied by its corresponding weighted moving ASI to obtain the final predicted sales volume (Table 6). Lastly, comparing the predicted values against the actual daily sales volumes of April 2018, the sum of squared differences proved to be relatively small (SSD = 3035). Furthermore, the month's predicted total sales was 653, which was fairly close to the actual sales total of 421.

Table 6. Applying the weight moving average method to compute the predicted sales volume.

Unit (x)	Date	Sales volume	WMA	SA (y)	Forcasting Model (yi = −0.0042x + 3.9115)	Predicted sales volume
754	3/1/2018	26	42.600	0.610328638	0.751438908	32.0112975
755	3/2/2018	29	42.300	0.685579196	0.747247873	31.60858505
756	3/3/2018	28	41.367	0.676873489	0.743056838	30.73778455
757	3/4/2018	17	39.233	0.433305013	0.738865803	28.98816835
753	3/5/2018	40	37.633	1.062887511	0.734674768	27.64826044
759	3/6/2018	51	41.500	1.228915663	0.730483733	30.31507492
760	3/7/2018	38	51.067	0.744125326	0.726292698	37.08934711
761	3/8/2018	36	50.367	0.714758438	0.722101663	36.36985376
762	3/9/2018	21	48.233	0.435383552	0.717910628	34.62722262
763	3/10/2018	16	47.167	0.339222615	0.713719593	33.66377413
764	3/11/2018	21	39.467	0.532094595	0.709528558	28.00272708
765	3/12/2018	19	32.967	0.576339737	0.705337523	23.252627
766	3/13/2018	34	33.700	1.008902077	0.701146488	23.62863663
767	3/14/2018	23	36.800	0.625	0.696955453	25.64796065
768	3/15/2018	68	45.900	1.481481481	0.692764417	31.79788676
769	3/16/2018	23	40.233	0.571665286	0.688573382	27.70360242
770	3/17/2018	15	33.767	0.444225074	0.684382347	23.10931059
771	3/18/2018	27	30.100	0.897009967	0.680191312	20.4737585

Exponential Smoothing Method. The next attempt was applying the exponential smoothing method to the seasonal index. A damping parameter of 0.7 was used to smooth out short-term fluctuations and to reflect any long-term trends or periodicities. Dividing the daily sales volumes between 3/1/2018 and 3/28/2018 by the exponential smoothing seasonal index generated the seasonally adjusted values (y). Then, these values were analyzed through regression and resulted in the equation $y_i = -0.003x + 3.0488$. Multiplying each seasonally adjusted predicted value by its corresponding weighted moving ASI, the final predicted sales volume data was obtained (Table 7). Lastly, comparing the predicted values against the actual daily sales volumes, the sum of squared difference was very large (SSD = 3608). Furthermore, the month's predicted total sales was 704, which was quite different from the actual sales total of 421. It can be concluded that using the exponential smoothing method to adjust for the seasonal index is not effective, possibly due to the size of the damping parameter

Table 7. Applying the exponential smoothing method to compute the predicted sales volume.

Unit (x)	Date	Sales Volume	DP (0.7)	SA (y)	Forcasting Model (yi = -0.003x + 3.0488)	Predicted Sales Volume
754	3/1/2018	26	40.494	0.642075003	0.754946637	30.57059137
755	3/2/2018	29	43.648	0.664404431	0.751904374	32.81920743
756	3/3/2018	28	41.561	0.673706892	0.748862111	31.1235336
757	3/4/2018	17	40.702	0.417673346	0.745819848	30.35610852
758	3/5/2018	40	38.810	1.030648947	0.742777585	28.82756877
759	3/6/2018	51	37.076	1.375534997	0.739735322	27.42678413
760	3/7/2018	38	43.323	0.877133354	0.736693059	31.91571282
761	3/8/2018	36	54.997	0.654582546	0.733650796	40.34850736
762	3/9/2018	21	48.932	0.429163513	0.730608533	35.75042781
763	3/10/2018	16	45.946	0.348231958	0.727566271	33.42904084
764	3/11/2018	21	48.084	0.436736477	0.724524008	34.83795142
765	3/12/2018	19	36.592	0.519241428	0.721481745	26.40034559
766	3/13/2018	34	30.111	1.129159742	0.718439482	21.63284916
767	3/14/2018	23	35.867	0.641265149	0.715397219	25.65886522
768	3/15/2018	68	38.527	1.765012177	0.712354956	27.44464749
769	3/16/2018	23	48.425	0.474964609	0.709312693	34.34822641
770	3/17/2018	15	37.627	0.398645687	0.70627043	26.5751187
771	3/18/2018	27	29.955	0.901355455	0.703228167	21.06511967

(DP) being too large at 0.7. Thus, excessive focus on the ASIs of the current period and the period immediately before would not produce satisfactory results in smoothing out extreme values.

5 Discussion and Conclusion

The current study chose a medical equipment company on the American e-commerce platform Amazon as the object of the research. This company had recently been impacted by a price war that caused significant overstocking and increased the company's inventory cost. Because of the impact of the price war, the existing mathematical model of the overall sales trend was not as reliable as it had been in the past when there was no price war. Thus, this study focused on exploring how a company can respond in a price war in terms of choosing sales forecast models to analyze forecast data. Three forecast methods were studied: the exponential smoothing method, the linear trend method, and the seasonal variation method. The results indicate that the seasonal variation method was the most suitable method for the data in this study because the original sales volumes data contained patterns that reflected seasonality and holiday influences. In the process of applying the seasonal variation method, it was found that the simple moving average method that used a 10-period simple moving ASI offered the best results for removing the seasonal influence from the data and was therefore identified as the most suitable forecast method for this case. The results are expected to help the medical equipment vendor accurately calculate the sales volume during a price war, reduce inventory cost, and increase overall profit.

5.1 Discussion of Results

This study found that the exponential smoothing method is constrained by the antecedent variable and must use the prior period's forecast. Thus, the forecast values in this method will trend toward 0 in long-term situations.

While the linear trend method does reflect the company's overall sales trend, the data in the study had an annual periodicity. If the goal is to forecast values for the upcoming month, this method would no longer offer accurate predictions due to its negligence of seasonality.

Among the three potential methods, the seasonal variation was the most suitable for the data in this study. We first used a simple average to calculate the seasonal index, then removed the seasonal index from all past sales volume values before doing the regression analysis, then multiplied the predicted values by their corresponding seasonal indexes. However, the results of this approach showed that because of insufficient data points (each period only had 2 or 3 values to be averaged), it was unable to approximate the true seasonal index. Any unusually high or low outliers would cause it to deviate considerably. Therefore, the seasonal index was calculated with a moving average or an exponential smoothing average to smooth out the short-term fluctuations and to reflect any long-term trends or periodicities. As a price war had started recently, using the overall historical data for regression would overlook the near-term trend of rapid sales decline. Thus, the study tested several scenarios and found that the data that yielded the most accurate regression equation for sales trend was the data from the month immediately prior to the forecast period that was also within the price war duration. After exploring the moving average, weighted moving average, and exponential smoothing methods to smooth out seasonal short-term fluctuations, it was found that the simple moving average of 10 periods was the most effective in eliminating seasonal influence from the data in this study and was the most suitable forecast method for this case. From the comparison presented in Table 8, it is clear which method is the best for companies facing a price war.

Table 8. Comparing the various methods for computing predicted sales volume and SSD.

Forecast period/Predicted Sales Volume and SSD	Linear trend method	Simple ASI	Simple moving 10-period ASI	Weighted moving ASI	Exponential smoothing seasonal index
2016/02/07–2018/03/28	1417 SSD = 34068	914 SSD = 11221	849 SSD = 7383	910 SSD = 9659	930 SSD = 9852
2018/02/01–2018/03/28	−21.3654 SSD = 8323	594 SSD = 2932	550 SSD = 1526	608 SSD = 2359	655 SSD = 2701
2018/03/01–2018/03/28	153.1 SSD = 3844	649 SSD = 3852	442 SSD = 900	652 SSD = 3035	704 SSD = 3607

Currently, the most common manufacturing method in industries is the "just in time" (JIT) model, which means that the manufacturers responsible for production would only produce items when demands arise. This is a manufacturing and purchasing strategy for minimizing or eliminating inventory that helps reduce inventory cost and increase overall profit. JIT was first proposed in 1953 by Taiichi Ohno, an executive of the Japanese company Toyota. The idea is to keep information flow and logistics parallel during production, in order to have the exact quantity of necessary materials to produce the exact quantity of necessary products at the right time [29]. This was done to shorten labor hours, reduce inventory, decrease production cost, and increase manufacturing efficiency [16].

As manufacturers adopt the JIT model, the pressure of managing inventory falls on the distributors. Generally, from the order of merchandise to product delivery takes about three months (a lead time of 90 days). This means that a distributor would need to predict at least three months' worth of merchandise sales volume when placing an order with the manufacturer. Otherwise, most manufacturers would not accept the order. For this reason, sales forecast is very important for distributors. If the three-month sales forecast is inaccurate, the distributor would be faced with inventory-related stress. If the forecast is larger than actual sales, the distributor would suffer a large inventory cost. When the distributor has originally planned for a low inventory, having overstock due to erroneous forecast would result in even more damaging consequences. In addition, consumers may not like buying products that have been sitting in the inventory for a long time, especially if the products have expiration dates [6]. For example, some products use lithium batteries, which may lose the battery power as they are kept in inventory for a prolonged period and result in lower durability than what the consumers expect. On the other hand, if the forecast is less than the actual sales, then the distributor is faced with the problem of merchandise shortage. Not only does the distributor miss out on sales opportunities, but it might also cause mistrust in its customers.

Moreover, more enterprises are using different social media to meet the customer's needs. Social media are now deeply rooted in modern life and have become a new way for companies to communicate with consumers [13]. According to prior research, including public sentiment variables in sales forecast models can increase the significance of the models. Because public sentiment variables are significantly related to certain products' sales outcomes, they are also valuable for these products' sales forecast [21, 35]. However, consumers often ignore social media ads because the ads and the consumer's preferences are mismatched [19]. Accordingly, enterprises should properly utilize social media, not only providing the target consumers with ads contents but also collecting and conducting time-series analysis on the big data around consumer's social media affective variables. Of course, social media platform administrators should also ensure the security of information and transactions, so that companies may properly use social networks to contact other companies and customers and so that customers can trust the newest information about the distributed merchandise [24].

5.2 Research Limitation and Future Research Suggestions

From the above discussion, it can be concluded that the importance of sales forecast is unquestionable. However, in the past, mathematical forecast models were not usually discussed when a company faced price wars to see how the forecast method might be adjusted. For this reason, this study provides some potential quantitative methods (exponential smoothing method, linear trend method, and seasonal variation method) that would allow companies to accurately predict sales in a price war and place orders accordingly.

The main challenge that this study encountered was that the Case Company was unwilling to make available its social media interaction data. Thus, we were unable to access data such as the level of product popularity or customer feedback. These missing data lowered the accuracy of our forecast results. If the Case Company could provide earlier historical sales data, we would be able to find the optimal ASI for the available dataset. If it could provide additional data such as number of orders placed, inventory cost, ideal quantity of inventory reserve, and delivery time, etc., we would be able to calculate the economic ordering quantity (EOQ) as a way to provide the company with the best reorder point. Future research might consider using public opinion analysis to understand the public's response to certain target products and to adjust the mathematical model accordingly in order to increase the overall accuracy of the forecast.

Acknowledgement. I would like to express my special thanks of gratitude to my four undergraduate students (Pin-Yuan Chen, Yun-Che Lin, Jun-Da Shu and Hirosato Song) who had tried so hard to manage a large dataset provided by the Case company. I would also like to thank the representative of the Case company for his clear direction in advising us to adopt certain data analysis methods. Without his help, my students and I would not be able to find a better forecasting sales model during a price war on Amazon e-commerce website for the Case company.

References

1. Avinash, B., Babu, S.: Big data technologies for e-business- Future opportunities, challenges ahead and growing trends. Int. J. Adv. Res. Comput. Sci. **9**(2), 328–332 (2018)
2. B2C, written by Volan, P., August 18, 2018, Price wars in the e-commerce industry: How big data helps businesses to gain market share. https://www.business2community.com/big-data/price-wars-in-the-e-commerce-industry-how-big-data-helps-businesses-to-gain-market-share-02109204. Accessed 22 Jan 2019
3. Brown, R.G.: Exponential smoothing for predicting demand 1956. http://legacy.library.ucsf.edu/tid/dae94e00. Accessed 21 Jan 2016
4. Brynjolfsson, E., Geva, T., Reichman, S.: Crowd-squared: amplifying the predictive power of search trend data. MIS Q. **40**(4), 941–961 (2016)
5. Business Insider, written by Green, D., November 24, 2017, These are the most popular items sold online for Black Friday so far, according to the data. http://www.businessinsider.com/black-fridays-most-popular-items-online-2017-11. Accessed 21 Jan 2019
6. Chan, T.K.H., Cheung, C.M.K., Lee, Z.W.Y.: The state of online impulse-buying research: a literature analysis. Inf. Manag. **54**(2), 204–217 (2017)

7. Chou, Y.L.: Statistical Analysis with Business and Economic Applications, 2nd edn. Continuum International Publishing Group Ltd., New York (1975)
8. Christiaanse, W.R.: Short-term load forecasting using general exponential smoothing. IEEE Trans. Power Apparatus Syst. **90**(2), 900–911 (1971)
9. Cui, R., Gallino, S., Moreno, A., Zhang, D.J.: The operational value of social media information. Prod. Oper. Manag. **27**(10), 1749–1769 (2017)
10. Currie, C.S., Rowley, I.T.: Consumer behaviour and sales forecast accuracy: what's going on and how should revenue managers respond? J. Revenue Pricing Manag. **9**(4), 374–376 (2010)
11. Everette, G.S.: Exponential smoothing: the state of the art. J. Forcasting **4**(1), 1–28 (1985)
12. Gahan, P., Pattnaik, M.: Optimization in fuzzy economic order quantity (FEOQ) model with promotional effort cost and units lost due to deterioration. LogForum **13**(1), 61–76 (2017)
13. Gallup, published by The Wall Street Journal, June 11, 2014, The myth of social media. http://online.wsj.com/public/resources/documents/sac_report_11_socialmedia_061114.pdf. Accessed 22 Jan 2019
14. Gilliland, M.: Role of the sales force in forecasting. Foresight Int. J. Appl. Forecast. **35**, 8–13 (2014)
15. Gonzalez, R., Hasker, K., Sickles, R.: An analysis of strategic behavior in eBay auctions. Singap. Econ. Rev. **54**(3), 441–472 (2009)
16. Hirano, H., Makota, F.: Just in Time is Flow: Practice and Principles of Lean Manufacturing. PCS Press, Vancouver (2006)
17. Hyndman, R.J., Athanasopoulos, G.: Forcasting: Principles and practice, 3.1 Some simple forecasting methods 2018. https://otexts.com/fpp2/. Accessed 21 Nov 2016
18. Janssen, M., van der Voort, H., Wahyudi, A.: Factors influencing big data decision-making quality. J. Bus. Res. **70**, 338–345 (2017)
19. Kietzmann, J.H., Hermkens, K., McCarthy, I.P., Silvestre, B.S.: Social media? Get serious! Understanding the functional building blocks of social media. Bus. Horiz. **54**(3), 241–251 (2011)
20. Lapide, L.: Are you capturing enough "quick-response" revenue? Supply Chain Manag. Revi. InSights **22**(2), 4–6 (2018)
21. Ma, Q., Zhang, W.: Public mood and consumption choices: evidence from sales of Sony cameras on Taobao. PLoS ONE **10**(4), e0123129 (2015)
22. McElroy, T.: Multivariate seasonal adjustment, economic identities, and seasonal taxonomy. J. Bus. Econ. Stat. **35**, 611–625 (2016)
23. Moon, S., Hicks, C., Simpson, A.: The development of a hierarchical forecasting method for predicting spare parts demand in the South Korean Navy—a case study. Int. J. Prod. Econ. **140**(2), 794–802 (2012)
24. Ramanathan, U., Subramanian, N., Parrott, G.: Role of social media in retail network operations and marketing to enhance customer satisfaction. Int. J. Oper. Prod. Manag. **37**(1), 105–123 (2017)
25. Rao, A., Bergen, M., Davis, S.: How to fight a price war. Harvard Bus. Rev. **78**(2, March/April), 107–116 (2000)
26. Reinmoeller, P.: How to win a price war. MIT Sloan Manag. Rev. **55**(3), 15–17 (2014)
27. Sagaert, Y.R., Aghezzaf, E.H., Kourentzes, N., Desmet, B.: Tactical sales forecasting using a very large set of macroeconomic indicators. Eur. J. Oper. Res. **264**(2), 558–569 (2018)
28. Seaman, B.: Considerations of a retail forecasting practitioner. Int. J. Forecast. **34**(4), 822–829 (2018)
29. Shah, R., Ward, P.T.: Lean manufacturing: context, practice bundles, and performance. J. Oper. Manag. **21**(2), 129–149 (2003)

30. Sillitoe, B.: Retailers urged to change approach to demand forecasting. Comput. Wkly., 15 June 2017. https://www.computerweekly.com/. Accessed 28 May 2019
31. Statistics on Key Figures of E-Commerce, surveyed by ACSI, February 2018, U.S. customer satisfaction with Amazon.com from 2000 to 2017 (index score). https://www.statista.com/statistics/185788/us-customer-satisfaction-with-amazon/. Accessed 21 Jan 2019
32. Vahid, M., Farokhi, M., Ibrahim, O., Nilashi, M.: A user satisfaction model for e-commerce recommender systems. J. Soft Comput. Dec. Support System **3**(3), 42–54 (2016)
33. Wang, G., Gunasekaran, A., Eric Ngai, W.T., Papadopoulos, T.: Big data analytics in logistics and supply chain management: certain investigations for research and applications. Int. J. Prod. Econ. **176**(C), 98–110 (2016)
34. Wild, T.: Best practice in inventory management, 3rd edn. Routledge, New York, NY (2018)
35. Yousef, M.I.: Social media with its role in supporting e-commerce and its challenges. J. Fundam. Appl. Sci. **10**(4S), 336–340 (2018)
36. Zhang, P.G., Qi, M.: Neural network forecasting for seasonal and trend time series. Eur. J. Oper. Res. **160**(2), 501–514 (2005)