

Analysis And Forecast of The Development of E-Commerce Based on R Language.

Xiaoming Zhang *

Robert Louis Stevenson School, Pebble Beach, Ca, United States

* Corresponding Author Email: szhang25@stevensonschool.org

Abstract. Since E-commerce's very start in the 1960s and through multiple eras of evolution, it has transformed the retail industry by offering a convenient and accessible platform for businesses and customers to conduct transactions without the need for physical shops. The COVID-19 pandemic was a pivotal point in the development of e-commerce, pushing it to the very top of retail. The paper reviews the development of e-commerce retail sales from the start of the 21st century based on two quarterly seasonally adjusted time series, mainly focusing on E-commerce's sales growth and its industry proportion. Subsequent works center around conducting forecasts and data analysis using various methods in R. However, many other factors play a significant role in the development of the industry. For example, no one might be able to predict its surge during the pandemic. The paper focuses on the potential influence of consumer behavior, environmental issues, and the market.

Keywords: E-commerce; R, time series; forecasts; analysis.

1. Introduction

E-commerce is currently an irreplaceable necessity in most people's lives. The history of the industry could be divided into four main eras: The first era was from the 60's to the 80s when merchants used Electronic Data Interchange (EDI) technology to exchange invoices, orders, and other business-type transactions [1,2]. EDI led to remote shopping, which is similar to e-commerce today. EDI Users were able to place orders with grocery stores and local stores, for example. The industry underwent its second period during the 80's and 90's. Namely, in 1982, the French developed The Minitel (consists of a beige computer and a keyboard). It is essentially a videotex terminal and network that provides online services, including a phone book directory, chat room, stock price checking, and reservations making. However, the devices of this period could not compete with what the Internet would offer users. The period of 90's to 2000's marks the third era. The period contains two key events in e-commerce history. In 1990, Tim Berners Lee and his friend Robert Cailliau published a proposal for the so-called "World Wide Web". Five years later, the National Science Foundation (NSF) withdrew the previous restrictions that limited the web's commercial use. These two changes enable the upcoming flourish of the Internet. Notably, Jeff Bezos established Amazon.com as an online bookshop in 1994, setting up the foundation for one of the world's most renowned e-commerce businesses. The dot-com bubble, which occurred in the late 1990s, was characterized by fast expansion and speculation of internet-based companies and e-commerce businesses. The bubble burst in the early 2000s, causing the bankruptcy of countless unproductive e-commerce firms. Despite the dotcom bust, e-commerce thrived in the early 2000s. The era witnessed significant developments in e-commerce platforms — for instance, improved security and simplicity of digital transactions. E-commerce has also extended to include online services, digital commodities, and software distribution.

The fourth era of e-commerce started from 2010 and lasts till now. Consumers, in this period, start making purchases and transactions via mobile applications and mobile-optimized websites. This is resulted from the increased usage of smartphones and mobile devices. Furthermore, social networks like Facebook, Instagram, and Pinterest have integrated e-commerce functions, allowing firms to sell things directly to customers. The COVID-19 epidemic in the 2020s was a pivotal moment for e-commerce. Before the pandemic, e-commerce had been steadily growing, fueled by technical advances, shifting customer preferences, and rising internet usage. However, brick-and-mortar retail

remained to dominate the retail sales. Some consumers were hesitant to adopt online purchasing owing to concerns about product quality, security, and the loss of physical contact. COVID-19 pushed e-commerce to the very top of retail [3,4]. In this special period, physical stores experienced severe issues with lockdowns and social distancing measures in place, and many were forced to close temporarily or permanently. Despite the worldwide recession, e-commerce experienced a sharp increase. The epidemic also spawned new business models, such as direct-to-consumer (D2C) brands, which enabled corporations to offer items directly to customers. Overall, E-commerce grew exponentially in this period [1].

2. Analyzing and Forecasting

2.1. Interpretation of data

The data sets of this report entitled "E-commerce retail sales as a percentage of total sales" and "E-commerce retail sales" come from the U.S. Census Bureau [5,6]. They are published in 2023 and document quarterly sales of e-commerce and its percentage of total sales from October 1999 to January 2023. The authors downloaded the time series from the website FRED ECONOMIC DATA and transformed them into Excel form when importing them into the R-studio.

After importing the time series into R-studio, the following two line graphs entitled "E-commerce Retail Sales as a Percentage of Total Sales" and "E-commerce Retail Sales" could be obtained.

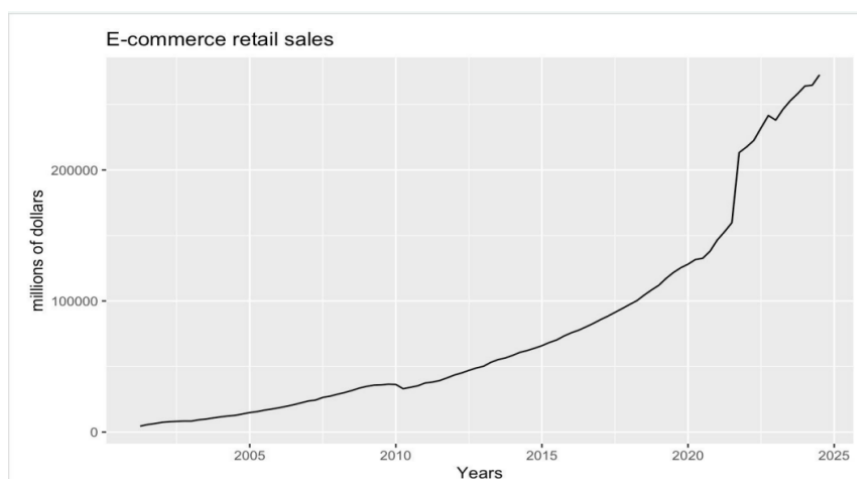


Fig. 1 Diagram of E-commerce retail sales

As figure 1 above indicates, there is an overall upward trend in the image, with total E-commerce retail sales showing a relatively slow upward trend from 2005 to 2020. However, from 2020 onwards, the overall slope of the image increases abruptly, before gradually returning to its original slope in 2022. This sudden increase may be caused by COVID-19, where people are staying at home longer and the demand for spending is shifting from offline brick-and-mortar stores to online e-commerce. This is likely to be one of the reasons for the dramatic change in data in 2020.

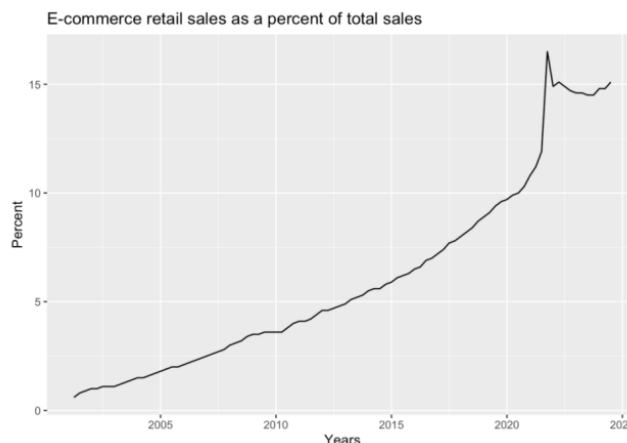


Fig. 2 Diagram of E-commerce retail sales as a percent of total sales

As figure 2 above illustrates, the diagram shows an upward trend overall but takes a turn in the second quarter of 2020. After that, it shows a slow downward trend. The graph also shows a dramatic increase in early 2020. This may be closely related to COVID-19, and the downward trend shown in the second quarter of 2020 may indicate that people are starting to shop offline again as they get used to COVID-19.

2.2. Examine the ACF

The report analyzes the ACF for "e-commerce retail sales as a percentage of total sales" and "e-commerce retail sales" using R, as shown in figure 3 and figure 4. As portrayed by the graphs, the two-time series are not white noises.

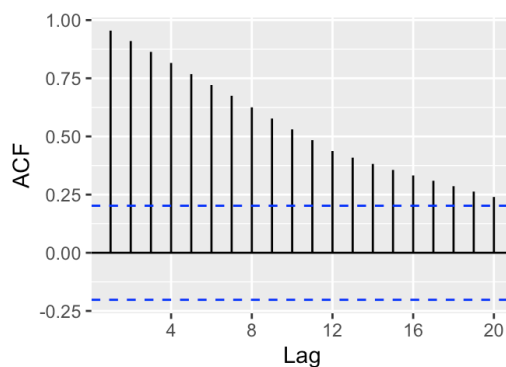


Fig. 3 ACF of e-commerce retail sales

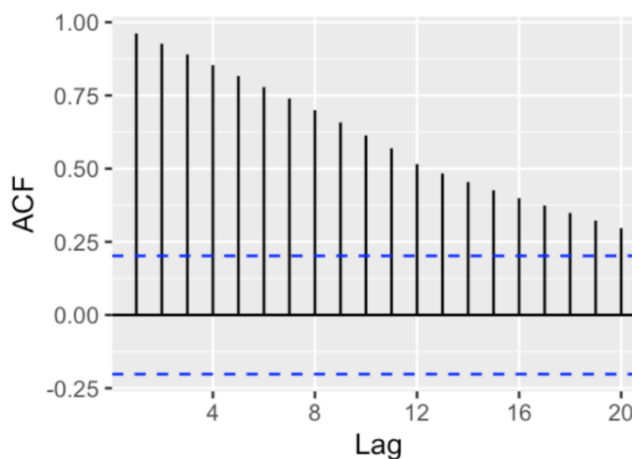


Fig. 4 ACF of e-commerce retail sales as a percentage of total sales

2.3. Residuals Checking

Utilizing the multiplicative decomposition method, data could be slitted into three separate components: trend-cycle component (T_t), seasonal component (S_t), and remainder component (R_t). In the method, it is assumed that the original data Y_t could be decomposed as equaling to $Y_t = T_t * S_t * R_t$. The decomposition of the two-time series is shown in figure 5 and figure 6.

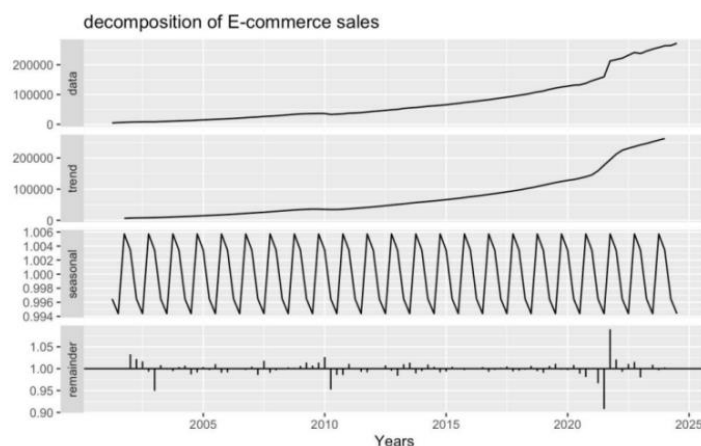


Fig. 5 Diagram of the decomposed E-commerce retails sales

In figure 5, the trend-cycle component demonstrates a mild upward trend. More specifically, the slope of the trend-cycle component approaches zero during the recession around 2008. It also manifests a stronger than-usual trend during the Pandemic period in which offline shopping was less accessible. The online shopping habit does not vanish after the end of lockdown policies. The graph continues to convey steadily growing quarterly sales data, reaching 272.6 billion in the first day of 2023. Although the data is already seasonally adjusted by U.S. Census Bureau, the graph still portrays a minor annual seasonality - maximum at the quarter starting from October and minimum at the quarter starting from July. Deducing this annual cycle, it might be caused by the fact that people generally tend to have more spare time for shopping during fall (summer breaks) while the cold weather, schoolwork, and jobs preoccupy people.

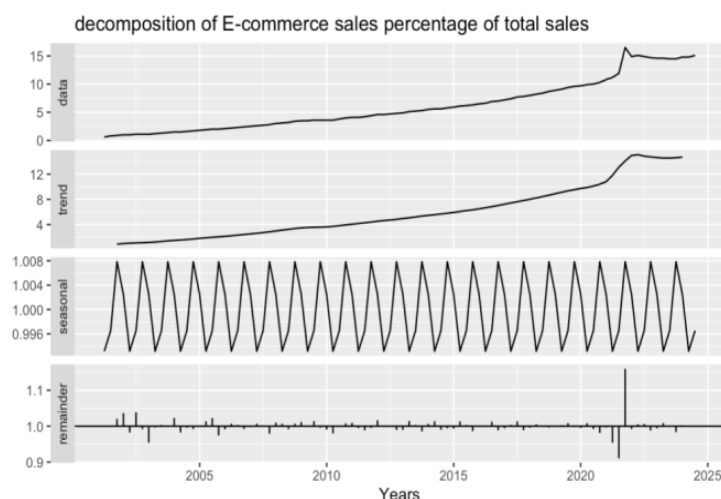


Fig. 6 Diagram of the decomposed E-commerce retails sales percentage

Figure 6 portrays a similar message. The trend-cycle component of the percentage data exhibits parallel ups and downs (for example, the recession in 2008 and the robust increase during the covid-19). Additionally, the seasonal component of the percentage graph shows an alike and minor annual seasonality,

2.4. Forecasting of The Time Series

- Drift: The forecast of future values is based on the average change seen in the previous data

Figure 7 demonstrates the application of the four forecasting methods. An accuracy test is applied to compare and contrast the different methods.



		ME	RMSE	MAE	MPE	MAPE	MASE
Training set	0.00000000000000001026791	3.366819	2.716327	-75.13533	103.64415	4.24426	
Test set	9.7290476190476198325996	9.731329	9.729048	65.90805	65.90805	15.20164	
ACF1 Theil's U							
Training set	0.9247298	NA					
Test set	0.4243243	55.04049					
		ME	RMSE	MAE	MPE	MAPE	MASE
Training set	0.1722892	0.5630168	0.2108434	3.6767811	3.935534	0.3294428	-0.2081929
Test set	-0.1400000	0.2529822	0.2200000	-0.9689719	1.498773	0.3437500	0.4243243
Theil's U							
Training set	NA						
Test set	1.476233						
		ME	RMSE	MAE	MPE	MAPE	MASE
Training set	0.000000000000000003609348	0.5360078	0.1622006	-2.047173	3.423870	0.2534385	
Test set	-1.08759036144578313809461	1.2175711	1.0931325	-7.392645	7.429348	1.7080196	
ACF1 Theil's U							
Training set	-0.2081929	NA					
Test set	0.6949994	7.32125					
		ME	RMSE	MAE	MPE	MAPE	MASE
Training set	0.64	1.018086	0.64	12.83290	12.83290	1.000000	0.6401021
Test set	1.55	2.703146	2.39	10.39287	16.14656	3.734375	1.1100520
							14.35655
ACF1 Theil's U							
Training set	0.64	1.018086	0.64	12.83290	12.83290	1.000000	0.6401021
Test set	1.55	2.703146	2.39	10.39287	16.14656	3.734375	1.1100520
							14.35655

Based on comparing the RMSE values from the second row of figure 8 of the four test sets, it may be concluded that the drift method fits the model the best with the smallest number among drift, mean, naive, and seasonal naive. Consequently, figure 9 portrays the forecasting of E-commerce retail sales with the drift.



The same four methods are applied to the percentage data as shown in figure 10. The following figure 11 shows the comparison of these methods.



Fig. 11 Comparison of the methods of forecasting E-Commerce sales percentage

531

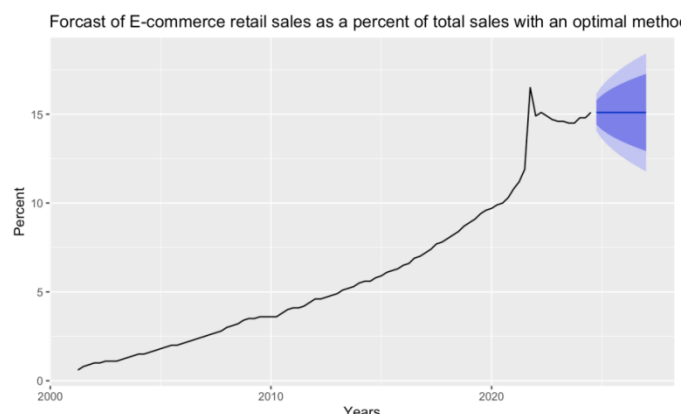


Fig. 12 Forecast of E-commerce retail sales percentage with an optimal method

3. Discussion

3.1. Consumer Behavior and Environmental Issues Might be Difficult to Predict

Consumers' behavior may be irrational, and they may choose to shop at offline supermarkets because of discounts or advertisements. Therefore, because consumer behavior is unpredictable and future popular consumption methods cannot be determined, future trends in e-commerce may be difficult to predict.

Environmental factors can also be unpredictable. As with COVID-19 in 2020, it will take a while for people to get used to it, and this stay-at-home time for consumers may boost e-commerce retail totals due to their inability to go out and shop. In addition to this, if consumers are physically close to the mall, they will prefer offline shopping because they can see the quality of the product in person and get good service faster. Meanwhile, e-commerce might pose significant environmental issues [7,8]. In addition to the amount of plastic, it is claimed that the industry constantly contributes to the generation of waste in packaging materials, including paper and cardboard. The alteration toward an environmentally friendly sustainable e-commerce might pose an unexpected impact on the industry.

3.2. Markets can be Difficult to Predict

Markets play a vital role in the development of e-commerce [9]. For instance, the economy of any country goes through different phases such as growth, expansion, and recession. The demand for a company's product depends upon which phase is prevalent in the economy. During the growth phase, where demand is robust and consumers have more disposable income, the demand is likely to pick up. During a recession, the economy contracts, the money supply shrinks and so do the demand and supply of a product. Thus, the economic cycle affects the sale of the company's products. Moreover, other factors such as laws and regulations and competitors' Market Positions will also have an impact on the product's selling.

Meanwhile, other factors include season and trends, customer behavior, and market size. Seasons have a significant impact on the e-commerce industry [10]. Trends, needs, and desires all fluctuate with the seasons. Besides, customer behavior is one of the main factors used to forecast demand. As they say, the best predictor of future behavior is past behavior. So, analyzing historical sales and customer behavior trends can offer insight into future demand. What's more, the size and nature of your market will influence your demand forecasting in several ways. In some ways, a smaller market is easier to predict. That said, this isn't necessarily the case if that market is demographically diverse which you'll be dealing with in a multi-channel strategy. A larger market may generally be harder to predict. By capturing and analyzing real-time and historical sales data, you'll be in a much better position to do demand forecasting.

4. Conclusion

Overall, this report yields several forecasting results for e-commerce. The Introduction and the Background provide a brief summarization of the history and development of e-commerce. The Data Description provides an overview of the data and their sources. In the Analyzing part, the author examines the trend of the data graph with both the ACF and residuals of the data output. Four main Forecasting methods were used in the forecasting process: Drift, Mean, Naive, and Seasonal Naive. The knowledge of confidence interval was utilized in the two optimal forecasting methods to determine the range of future trends. The forecasting intervals provide a positive attitude toward the future of the E-commerce with increasing sales and proportion of the industry.

References

- [1] Big Commerce E-commerce. The History and Future of Online Shopping. Sep 6 2023, <https://www.bigcommerce.com/articles/ecommerce/>.
- [2] Miva The History Of E-commerce. How Did It All Begin? Nove 23 2022. Sep 7 2023. <https://blog.miva.com/the-history-of-ecommerce-how-did-it-all-begin>.
- [3] U.S. Census Bureau. Annual Retail Trade Survey Shows Impact of Online Shopping on Retail Sales During COVID-19 Pandemic. April 27 2022, Sep 6 2023. <https://www.census.gov/library/stories/2022/04/ecommerce-sales-surged-during-pandemic.html>.
- [4] Fred Economic Data. The pandemic's boost to online sales: A one-time event or a new normal. Jun 21 2021, Sep 7 2023, <https://fredblog.stlouisfed.org/2021/06/the-pandemics-boost-to-online-sales-a-one-time-event-or-a-new-normal/>.
- [5] U.S. Census Bureau. E-Commerce Retail Sales, retrieved from FRED, Federal Reserve Bank of St. September 2 2023. Louis; <https://fred.stlouisfed.org/series/ECOMSA>.
- [6] U.S. Census Bureau. E-Commerce Retail Sales as a Percent of Total Sales, retrieved from FRED, Federal Reserve Bank of St. September 2, 2023. Louis; <https://fred.stlouisfed.org/series/ECOMPCTSA>.
- [7] Sana Commerce Is e-commerce really sustainable. Understanding its impact on the environment Feburary 17 2021. Sep 6 2023. <https://www.sana-commerce.com/blog/impact-of-ecommerce-on-the-environment/>
- [8] Earth 5R The Environmental Impact Of E-commerce. Apr 24 2023, Sep 6 2023, <https://earth5r.org/convenience-comes-at-a-cost-the-environmental-impact-of-ecommerce-packaging/#:~:text=The%20use%20of%20plastic%20in,the%20amount%20of%20waste%20generated>.
- [9] Brookings The Emerging Markets E-commerce Opportunity. Mar 26 2021, Sep 6 2023, <https://www.brookings.edu/articles/the-emerging-markets-e-commerce-opportunity/>
- [10] Blend Commerce E-commerce Seasonality, Trends, information, and How to Measure. Sep 6 2023. <https://blendcommerce.com/blogs/shopify/e-commerce-seasonality-trends-information-and-how-to-measure#:~:text=E%2Dcommerce%20seasonality%20is%20a,keep%20up%20with%20the%20demand>.