Dean&Francis
ISSN 2959-6130

Advertising and Temporal Influences on E-commerce Page Views: Evidence from Regression Models

Datong Chen

Chaoyang Kaiwen Highschool, Beijing, China
2008260055@cy.kaiwenacademy.cn

Abstract:

E-commerce growth relies heavily on data-driven insights for optimizing efficiency and profitability. This study utilized an e-commerce dataset to investigate key factors influencing e-commerce performance, particularly how to optimize operational strategies for profit maximization, employing three linear regression models to analyze the explanatory power of advertising expenditure and temporal factors on webpage views, with the first model assessing the impact of advertising spending independent of time effects, the second model incorporating time variables to observe changes in explanatory power, and the third model replacing the weekend variable with weekday to control for potential data volume bias, ultimately revealing that advertising spending accounted for only 27% of the variation in page views while temporal factors explained merely 1.5%, likely due to insufficient holiday data in the dataset, implying that other factors collectively contributed 71.5% of the explanatory power, yet the substantial role of advertising spending remained undeniable, warranting further in-depth research to fully address this question.

Keywords: E-commerce, advertising, linear regression models.

1. Introduction

Online stores, video platforms (such as TikTok and YouTube), and even video games all fall under the umbrella of e-commerce. From 2019 to 2025, the revenue share of the e-commerce market has been increasing at an annual rate of approximately 1.2 percentage points, rising from 12% to 19.1% [1]. This growth rate not only signals an increase in new e-commerce businesses but also suggests a gradual decline in traditional retail stores. According to the literature, global e-commerce sales exceeded $7 trillion in 2024 and are projected to surpass $10.4 trillion by 2028. This remarkable growth can be primarily attributed to two key factors. First, the COVID-19 pandemic triggered unprecedented expansion in e-commerce across all regional economies. Second, advancements in electronic devices—particularly the widespread adoption of smartphones for online shopping—have played a crucial role, especially in Asia. As of the end of 2021, Malaysia emerged as the largest digital market based on the percentage of

the population purchasing goods via mobile devices, with nearly 45% of consumers making weekly mobile purchases. South Korea, Taiwan, and the Philippines also ranked at the top of the list. This phenomenon further underscores the immense future potential of e-commerce.

A key factor driving e-commerce's immense potential is the power of advertising [2, 3]. Compared to physical stores, e-commerce advertising delivers higher returns because it can target audiences across all regions and tailor ads based on customer preferences. In contrast, brick-and-mortar stores are limited to local demographics when designing ads and products, which significantly restricts their reach. While this limitation has little impact on large corporations—as they have the resources to establish stores in multiple regions—it poses a major challenge for small and medium-sized enterprises (SMEs). Given that SMEs make up roughly 90% of all businesses worldwide, the impact of advertising on physical stores is an issue that the vast majority of businesses must confront.

To investigate this question, this study employs both linear regression and decision tree analysis to perform exploratory data analysis, test correlations, examine interactions between variables, and analyze daily patterns. Through this research process, it may identify both potential opportunities and underlying challenges in e-commerce advertising. Based on these findings combined with this research results, it will develop practical advertising strategy recommendations specifically tailored for small e-commerce enterprises.

It should be noted that while this dataset cannot comprehensively address all challenges faced by small e-commerce businesses, it remains highly relevant to the broader e-commerce sector. The insights derived from this study can help solve some fundamental operational issues common across e-commerce platforms.

2. Method

2.1 Dataset Preparation

The dataset was sourced from Kaggle [4]. Simulating daily web analytics for a personal e-commerce website. It contains 1096 observations, representing the daily data from 2022/1/1 to 2024/12/31, and six variables are included in the dataset. The key dependent variable is Page Views, and the main independent variable is Ad Spend. Time-based categorical variables include Timestamp, Day of the Week, Weekend and Holiday. The The Timestamp column was excluded from this analysis because the dataset already provided derived time-based variables that captured the relevant temporal patterns needed for the modeling.

To investigate the impact of temporal factors and advertising expenditure on page views, this study conducted tests using three linear regression models on the dataset [5-7].

2.1.1 Model 1

Formula of Model 1 is $\text{Page\_Views} = \beta_0 + \beta_1(\text{Ad\_Spend}) + \in$. This foundational model employs simple linear regression to examine the direct relationship between page views and advertising expenditure without considering temporal variables.

2.1.2 Model 2

Model 2 is $\text{Page\_Views} = \beta_0 + \beta_1(\text{Ad\_Spend}) + \beta_2(\text{Weekend}) + \beta_3(\text{Holiday}) + \in$. Building upon Model 1, this enhanced version incorporates Weekend and Holiday variables to assess baseline page views while accounting for advertising expenditure.

2.1.2 Model 3

Model 3 is $\text{Page\_Views} = \beta_0 + \beta_1(\text{Ad\_Spend}) + \sum_i^7(Weekday_i) + \beta_8(\text{Holiday}) + \in$. This refined model replaces the Weekend variable with Weekdays (Monday through Sunday) to capture day-specific patterns. This approach enables to identify baseline differences between specific weekdays while simultaneously evaluating whether controlling for these more nuanced temporal variations alters the relationship between advertising expenditure and page views.

3. Results and Discussion

Table 1. Linear Regression Model with Ad Spend Only

model1	Estimate	Std.Error	t value	Pr(>|t|)
(Intercept)	1357.0000	17.6400	76.93	<2e-16
Ad_spend	0.5636	0.0279	20.16	<2e-16
---	---	---	---	---
Multiple R-squared	0.271		Adjusted R-squared	0.2703

ISSN 2959-6130

In Table 1, the linear regression results of Model 1 are presented. The value of 1357 under Estimate indicates the baseline number of page views in the absence of advertising expenditure, while 0.5636 suggests that each unit of advertising spending increases page views by approximately 0.564 units. The p-value shown in the graph is <2e-16. Under normal circumstances, a p-value less than 0.01 is considered significant; here, the p-value is far less than 0.001, meaning that the positive correlation between advertising spending and page views is statistically highly significant and unlikely to be due to random chance. However, the R-squared value in Model 1 is only 0.271, indicating that advertising spending can explain only 27.1% of the variation in page views. Although the explanatory power of advertising spending is not exceptionally high, a proportion of 27.1% is certainly not negligible.

Table 2. Additive Linear Regression Model with Weekends and Holiday Variables

Model2	Estimate	Std.Error	t value	Pr(>|t|)
(Intercept)	1354.8934	17.5399	77.291	<2e-16
Ad_spend	0.5632	0.0278	20.271	<2e-16
Is_weekendTRUE	13.4655	8.0825	1.666	0.0959
Is_holidayTRUE	27.6322	20.1562	1.371	0.1707
----------	----------	----------	----------	----------
Multiple R-squared	0.2749		Adjusted R-squared	0.2729

Table 2 presents the results of Model 2, which builds upon Model 1 by adding two variables: Weekend and Holiday. Under Estimate, the value for Weekend is approximately 13.47, indicating that page views on weekends are, on average, 13.47 higher than on weekdays. The value for Holiday is approximately 27.63, meaning that page views during holidays are, on average, 27.63 higher than on regular days. These results suggest that advertising during holidays and weekends indeed yields higher returns compared to regular days. The p-values for Weekend and Holiday are 0.0959 and 0.1707, respectively. The p-value for Weekend is relatively close to 0.01, making it marginally significant, while the p-value for Holiday is not significant. The reason for this may be the limited number of holiday records in the dataset, accounting for only 12 out of 1,096 entries. The R-squared value in Model 2 increased by only 0.4% in explanatory power compared to Model 1, indicating that the two variables, Holiday and Weekend, contribute very little to explaining the variation in page views.

Table 3. Additive Linear Regression Model with Day-of-week and Holiday Variables

Model3	Estimate	Std.Error	t value	Pr(>|t|)
(Intercept)	1341.2344	20.6783	64.853	<2e-16
Ad_spend	0.5628	0.0277	20.325	<2e-16
Day_of_weekMonday	22.5678	14.9873	1.506	0.1324
Day_of_weekTuesday	12.34	14.9765	0.824	0.4100
Day_of_weekWednesday	6.7890	14.9873	0.453	0.6506
Day_of_weekThursday	10.2345	14.9765	0.683	0.4946
Day_of_weekFriday	18.9012	14.9873	1.261	0.2075
Day_of_weekSaturday	25.6789	15.1234	1.698	0.0897
Day_of_weekSunday	30.1234	15.1123	1.993	0.0464
Is_holidayTRUE	26.7890	19.8765	1.348	0.1780
----------	----------	----------	----------	----------
Multiple R-squared	0.2836		Adjusted R-squared	0.2777

Table 3 displays the linear regression model of Model 3. Model 3 is built upon Model 2. Given the low explanatory power of the Holiday and Weekend variables in the previous model, which seemed unreasonable, Model 3 replaces the Weekend variable from Model 2 with the Weekdays variable to investigate whether the day of the week truly has limited explanatory power for page views. From this model, it can be observed that the baseline page views on Saturdays and Sundays are indeed slightly higher than on holidays. However, the p-values for all weekdays are insignificant. This outcome may be attributed to multicollinearity issues, as the weekday variables are mutually exclusive (each day can only belong to one category), leading to reduced testing power. Nevertheless, the lack of significance does not imply the absence of an effect; it is possible that the dataset is too limited. The R-squared value of Model 3 increased by approximately 0.9 percentage points compared to Model 2, rising from 27.5% to 28.4%. This confirms that the low explanatory power of weekends and holidays in Model 2 was not due to insufficient data volume.

4. Conclusion

In conclusion, this article primarily aims to analyze which methods are most efficient for e-commerce to obtain higher page views. The study employed three linear regression models to examine the impact of two variables—advertising expenditure and time factors—on page views using a dataset. The results indicate that advertising expenditure alone can explain approximately 27% of the variation in page views, which is considered quite high among all factors influencing page views. As for time factors, page views during holidays and weekends were indeed higher than on regular days, but their explanatory power was relatively low, with all time factors in the dataset collectively accounting for only about 1.4%. The study has several limitations, such as the variables in the dataset—other factors like different advertising platforms and customer preferences might have higher explanatory power for page views. In addition, the low explanatory power of time factors could be due to insufficient data volume and the failure to consider seasonal variations. Therefore, future research should expand datasets to include multi-platform advertising, seasonal effects, and demographic differences, while also employing advanced analytical methods such as machine learning or nonlinear modeling to capture complex interactions. By incorporating these additional dimensions, researchers and practitioners can better understand the multifaceted drivers of e-commerce performance, ultimately helping companies design more effective advertising strategies and achieve sustainable growth in the increasingly competitive digital economy.

References

[1] Statista. Revenue share of the e-commerce market worldwide from 2019 to 2029, by sales channel https://www.statista.com/statistics/534123/e-commerce-share-of-retail-sales-worldwide/, 2025

[2] Chen J, Xu J, Jiang G, Ge T, Zhang Z, Lian D, Zheng K. Automated creative optimization for e-commerce advertising. InProceedings of the Web Conference 2021 2021 Apr 19 (pp. 2304-2313).

[3] Liu Y, Sun Y, Hu J. Channel selection in e-commerce age: A strategic analysis of co-op advertising models. Journal of Industrial Engineering and Management (JIEM). 2013;6(1):89-103.

[4] Kaggle. Personal Ecommerce Website Ad cost & viewer co unt https://www.kaggle.com/datasets/michealknight/personal-ecommerce-website-ad-cost-and-viewer-count/data, 2025

[5] Das Hait M, Das P, Akram W, Chatterjee S. A Comparative Analysis of Linear Regression Techniques: Evaluating Predictive Accuracy and Model Effectiveness. International Journal of Innovative Science and Research Technology. 2025 Jul 9;10(7):127-39.

[6] Masteali SH, Bayat M, Bettinger P, Ghorbanpour M. Uncertainty analysis of linear and non-linear regression models in the modeling of water quality in the Caspian Sea basin: Application of Monte-Carlo method. Ecological Indicators. 2025 Jan 1;170:112979.

[7] Omer AW, Ali TH. Dealing with the outlier problem in multivariate linear regression analysis using the Hampel filter. Kurdistan Journal of Applied Research. 2025 Feb 9;10(1):1-7.