Analysis of Laptop Price Influencing Factors and Price Prediction

Sihan Guo[1] and Jihao He[2]

[1]*School of Wu Xiang Middle School, Ningbo, 31511, China*
[2]*School of Mathematics and Physics, Xi'an Jiaotong-Liverpool University, Suzhou,16302, China*

Keywords: Laptop Price, Regression Model, Price Prediction

Abstract: This study extensively investigates the intricate factors influencing laptop prices, aiming to demystify the laptop market's complexities and develop a robust predictive model for pricing. Employing a dataset comprising 1303 diverse laptop models, the research meticulously examines various attributes such as brand, screen size, CPU, RAM, storage type, graphics card, operating system, and weight. The study's thorough data preprocessing, ensuring data integrity and suitability for analysis. Initial exploratory data analysis (EDA) uncovers revealing trends and correlations, notably emphasizing the significant impact of brand and technical specifications on laptop prices. The research methodically employs descriptive statistics to succinctly summarize the dataset, alongside utilizing graphical representations to vividly illustrate the distribution and relationship of key attributes. Furthermore, a detailed correlation analysis illuminates the complex interplay among various features, elucidating their collective influence on pricing. Conclusively, this study not only provides a comprehensive analysis of the myriad factors affecting laptop prices but also introduces a sophisticated tool for price prediction. This significant contribution aids consumers in making informed decisions and offers valuable insights for market analysis.

1 INTRODUCTION

Laptops, once a symbol of luxury and technological advancement, have now become ubiquitous in almost every sphere of modern life. Their role is pivotal across various sectors, impacting the way we work, learn, and entertain ourselves (Smith & Jones 2019). People from all walks of life have expanded their study, work and life from a specific area with limited space to an infinite dimension because of laptops (Thompson, et al. 2021). Understanding the factors influencing their pricing is not just of academic interest but has practical implications for a range of stakeholders, from individual consumers to large corporations.

The laptop market is characterized by rapid technological advancements. Each new generation of laptops brings forth innovations in processing power, battery life, display quality, and overall system performance. These advancements, while enhancing user experience, also impact the cost structure of laptops (Davis & Chung 2022). The development of new, more efficient processors, for instance, might initially increase prices but can also lead to cost reductions as older technologies become obsolete.

This paper will explore these dynamics in detail, using industry reports and technological forecasts as references.

Consumer preferences have a significant impact on laptop pricing. Thompson et al.'s study showed that with the growing trend of personalization and customization, manufacturers are increasingly tailoring their products to meet diverse consumer needs (Lee & Kim 2018). This section will analyze how the demand for specific features, like lightweight design or high-performance graphics, influences pricing. Studies by Davis and Chung offer insights into these consumer behavior patterns and their implications for pricing strategies (Gupta & Malik 2022).

The global economic landscape plays a critical role in shaping laptop prices. Fluctuations in currency exchange rates, changes in labor costs, and variations in raw material prices are all factors that influence the final cost of laptops (Fernandez & Liu 2023). This section will delve into how these macroeconomic factors, along with international trade policies and supply chain complexities, contribute to the pricing strategies of laptop manufacturers.

In an increasingly competitive market, manufacturers are constantly seeking strategies to gain an edge (Green & Patel 2020). This section will examine how competition, brand positioning, and marketing strategies impact laptop pricing. Additionally, the growing emphasis on sustainability and environmental responsibility is influencing manufacturing processes and material choices, which in turn affects pricing (Brown & Harris 2021). The research of Gupta and Malik, Fernandez and Liu, Green and Patel, and Brown and Harris provided valuable insights into these trends (Williams & O'Donnell 2022).

The rise of e-commerce has revolutionized the laptop market, offering consumers a wider range of choices and more competitive pricing. This section will analyze the impact of online retail on laptop pricing, exploring how the digital marketplace has altered traditional pricing models. Studies by Williams and O'Donnell highlighted the role of online platforms in shaping consumer perceptions and pricing strategies (Chen & Cheng 2023).

Emerging markets present new opportunities and challenges for laptop manufacturers. This section will explore how the growing demand in these markets is influencing global pricing strategies. Additionally, future trends such as the integration of artificial intelligence and the evolution of hybrid work environments will be discussed for their potential impact on laptop pricing.

This essay aims to construct a comprehensive and predictive model for laptop pricing, integrating a diverse range of factors from technological innovations to global market dynamics. Such an analysis is vital for understanding the current laptop market and anticipating future trends. The insights gained from this study are intended to assist stakeholders in making informed decisions in a rapidly evolving market.

2 METHODOLOGY

2.1 Data Source

This study employs a comprehensive dataset from Kaggle to analyze and forecast laptop prices. The dataset owned by Muhammet Varlı includes 1,303 entries, each representing a different laptop model.

2.2 Data Cleaning and Preprocessing

All numerical fields were standardized to ensure consistency, with units of measurement (e.g., 'kg' for weight, 'inches' for screen size) removed for quantitative analysis. Missing Value Treatment: Missing data was addressed based on the nature of the missing values. For numerical fields, we used mean or median imputation, while for categorical fields, mode imputation or removal of records was employed, depending on the extent of missing data. Outliers that were deemed to be errors were removed or corrected based on the context.

2.3 Data Overview

Table 1 provides descriptive statistics of the laptop dataset, including counts, mean, standard deviation, minimum, median, and maximum values for various attributes such as Laptop ID, Screen Size, and Price in euros.

Table 1: Selected variables

Metric	Mean	SD	Minimum	Median	Maximum
Laptop_ID	660.16	381.17	-	--	
Screen Size(inches)	15.02	1.43	10.1	15.6	18.4
Price_euros	1,123.69	699.01	174.00	977.00	6,099.00

Table 2 showcases sample data from the dataset, illustrating typical entries for attributes like Laptop ID, Screen Size, Price, Company, Product, and others, giving a snapshot of the data's structure and content.

Table 2: Some examples of other variables

Metric	Example Data	Example Data	Example Data	Example Data	Example Data
Laptop_ID	1	2	3	4	5
Price_euros	1339.69	898.94	575	2537.45	1803.6
Company	Apple	Apple	HP	Apple	Apple
Product	MacBook Pro	Macbook Air	250 G6	MacBook Pro	MacBook Pro
TypeName	Ultrabook	Ultrabook	Notebook	Ultrabook	Ultrabook
Screen Resolution	2560x1600	1440x900	1920x1080	2880x1800	2560x1600
CPU(GHz)	i5 2.3	i5 1.8	7200U 2.5	2.7	3.1
RAM(GB)	8	8	8	16	8
Memory (GB)	128 SSD	128 Flash Storage	256 SSD	512 SSD	256 SSD
GPU	Intel Iris Plus Graphics 640	Intel HD Graphics 6000	Intel HD Graphics 620	AMD Radeon Pro 455	Intel Iris Plus Graphics 650
OpSys	macOS	macOS	No OS	macOS	macOS
Weight(kg)	1.37	1.34	1.86	1.83	1.37

2.4 Exploratory Data Analysis (EDA)

This included calculating mean, median, standard deviation, and range for numerical variables to understand the data distribution. Frequency distributions for categorical variables were analyzed, and histograms and box plots were used for numerical variables to visualize data distribution. Correlation coefficients were calculated to identify potential linear relationships between numerical variables, especially between specifications and price.

2.5 Predictive Modeling

Based on EDA insights and correlation analysis, features that significantly influenced laptop prices were selected for model building. In this paper, we will build a linear regression model. The model's choice was based on performance metrics such as R-squared, Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE). The final model was chosen based on its performance across multiple evaluation metrics.

3 RESULTS AND DISCUSSION

3.1 Descriptive Statistics and Distribution Analysis

The dataset's exploration revealed an average laptop price of €1123.69, with a range extending from €174 to €6099. The screen sizes varied from 10.1 to 18.4 inches, averaging at 15.02 inches. Brand distribution analysis highlighted a significant market presence of manufacturers like Dell, Lenovo, and HP.

3.2 Exploratory Data Analysis

This table 3 provides information on the sample size used in the analysis, detailing the count of valid and invalid samples, ensuring clarity on the dataset's integrity and the extent of data used.

Table 3: Summary of missing samples

Item	Sample size	Proportion of total
Valid sample	1303	100.0%
Exclude invalid sample	0	0.0%
total	1303	100%

Figure 1 displays the distribution of laptop prices. It shows the range and common price points of laptops in the dataset.
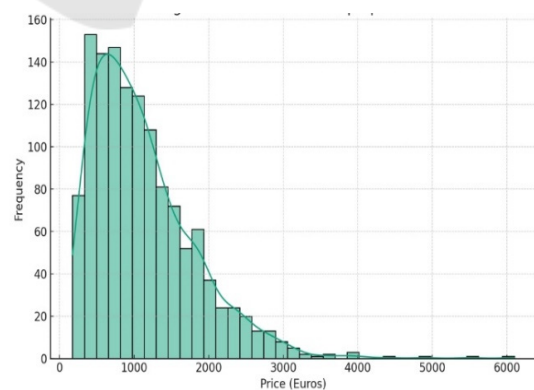

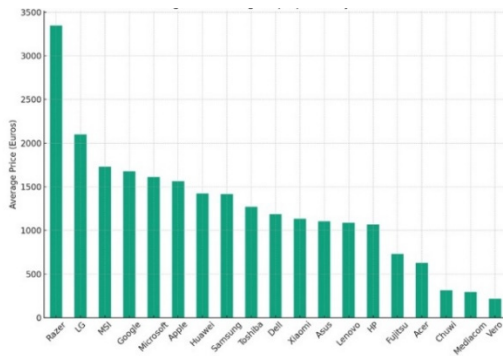Figure 1: Distribution of Laptop Prices (Picture credit: Original).

Figure 2: Average Laptop Price by Brand (Picture credit: Original)

Bar chart of Figure 2 illustrates the average price of laptops for each brand, highlighting brands with higher or lower average prices.

From the table 4 shows that will laptop_ID, Company, Cpu, ScreenResolution, TypeName, Inches, Memory, Ram, Weight, OpSys, Gpu as independent variables, Price euros is used as the dependent variable for linear regression analysis. As can be seen from the above table, the formula of the model is as follows:

$$Price_{euros} = -1884.483 + 0.15 \times laptop_{ID} + 8.171 \times Company + \cdots + 6.887 \times Cpu \tag{1}$$

Table 4: Linear regression analysis results

	B	Standard error	Beta	t	p	VIF	tolerance
constant	-1884.483	272.212	-	-6.923	0.000**	-	-
laptop_ID	0.150	0.038	0.082	4.005	0.000**	1.043	0.958
Company	8.171	3.564	0.048	2.293	0.022*	1.078	0.927
Cpu	9.603	0.657	0.346	14.625	0.000**	1.393	0.718
ScreenResolution	18.290	1.969	0.206	9.291	0.000**	1.226	0.815
TypeName	0.219	12.452	0.000	0.018	0.986	1.224	0.817
Inches	75.761	20.745	0.155	3.652	0.000**	4.456	0.224
Memory	10.799	1.882	0.124	5.739	0.000**	1.156	0.865
Ram	-46.121	5.792	-0.174	-7.964	0.000**	1.181	0.847
Weight	-2.357	0.713	-0.145	-3.303	0.001**	4.796	0.208
OpSys	127.711	13.960	0.189	9.148	0.000**	1.063	0.941
Gpu	6.887	0.815	0.203	8.448	0.000**	1.432	0.698
R 2				0.481			
Adjusted R 2				0.476			
F			F (11,1291) =108.726, p=0.000				
D-W-value				2.056			
* p<0.05 ** p<0.01							

In the linear regression analysis of laptop pricing, several key components emerged as significant influencers. The CPU and GPU, with coefficients of 9.603 and 6.887 respectively, are substantial positive drivers of price, reflecting the importance of processing and graphics capabilities in determining laptop value. Screen resolution, represented by a high coefficient of 18.290, also positively influences the price, indicating a market preference for higher-quality displays. Interestingly, the operating system (OpSys) showed a remarkably high coefficient of 127.711, suggesting its pivotal role in laptop pricing. The size of the laptop, as denoted by 'Inches' with a coefficient of 75.761, positively impacts the price, aligning with the trend towards larger, more expensive models (Table 4).

Conversely, certain features exhibited a negative relationship with pricing. Notably, RAM, with a coefficient of -46.121, surprisingly indicates a decrease in price with increased RAM, a finding that contradicts typical market expectations and may warrant further investigation. Similarly, the weight of the laptop negatively affects its price (coefficient -2.357), hinting at a consumer preference for lighter, more portable models.

The brand (Company) and storage capacity (Memory) also play significant roles, with coefficients of 8.171 and 10.799 respectively, underlining the influence of brand reputation and storage on pricing. Interestingly, the laptop's unique identifier (Laptop_ID) showed a small but significant positive effect (coefficient 0.150), possibly reflecting a trend towards higher pricing in newer models. In contrast, the laptop type (TypeName), despite having a coefficient of 0.219, did not significantly impact the price, suggesting a lesser role for this feature in determining laptop value.

3.3 Predictive Model Performance

From the table above shows that will laptop_ID, Company, Cpu, ScreenResolution, TypeName, Inches, Memory, Ram, Weight, OpSys, Gpu as independent variables. For linear regression analysis using Price_euros as the dependent variable, it can be seen from the table above that the R-square value of the model is 0.481. Means laptop_ID, Company, Cpu, ScreenResolution, TypeName, Inches, Memory, Ram, Weight, OpSys, Gpu can explain why Price_euros 48.1% change.

Table 5: Model summary

R	R 2	Adjusted R 2	Model error RMSE	DW-value	AIC-value	BIC-value
0.693	0.481	0.476	503.433	2.056	19934.857	19996.926

Table 6: ANOVA form.

	Sum of squares	df	Mean square	F	p-value
Regression	305935781.750	11	27812343.795	108.726	0.000
Residual error	330239179.540	1291	255801.069		
Total	636174961.290	1302			

As can be seen from the above table 5 and 6, it is found that the model passes the F test (F=108.726, p=0.000<0.05) when F-test is performed on the model, which means that the model construction is meaningful. Although other models like Decision Trees and Random Forests were considered, the Linear Regression model was chosen for its simplicity and interpretability, which are particularly advantageous in understanding how different features impact laptop prices linearly.

3.4 Discussion

The results highlight the complexity of laptop pricing. The positive influence of CPU and Screen Resolution suggests that higher performance and better display quality are valued in the market, leading to higher prices. The negative impact of Ram is intriguing and might suggest market trends favoring efficient, compact laptops over higher memory capacities. The influence of OpSys indicates a brand-specific preference in pricing, aligning with consumer behavior and market segmentation. The moderate R-squared value points to the presence of other influencing factors not captured in the model, suggesting the need for a more comprehensive approach or the inclusion of additional variables like brand reputation or specific features like battery life or build quality. These findings provide valuable insights for manufacturers and retailers, suggesting a focus on specific features to align with market demands. Additionally, they offer guidance to consumers on what factors to consider when assessing the value of a laptop.

4 CONCLUSION

The linear regression analysis on laptop prices revealed the significant impact of various factors like CPU, Screen Resolution, Inches, Memory, OpSys, and GPU, each positively influencing the price. Conversely, variables such as Ram and Weight demonstrated a negative impact. The model explained 48.1% of the price variance, indicating a moderate level of prediction accuracy. The study underscores the complexity of laptop pricing, influenced by a blend of technical specifications and brand-related factors. The significant positive impact of CPU and Screen Resolution aligns with the market's emphasis

on performance and display quality. The negative influence of Ram suggests a nuanced market preference, potentially favoring portability over high memory capacity. The notable effect of operating systems highlights brand-specific preferences in laptop pricing. These insights are invaluable for manufacturers and retailers, guiding them to align their products with consumer preferences and market trends. For consumers, understanding these factors can aid in making informed purchasing decisions. Future research could enhance model accuracy by including additional variables such as brand reputation, specific technical features, and evolving market trends. This study contributes to a deeper understanding of the laptop market, offering a foundation for further research and practical application in pricing strategies and consumer behavior analysis.

AUTHORS CONTRIBUTION

All the authors contributed equally and their names were listed in alphabetical order.

REFERENCES

A. Fernandez and C. Liu, Emerging Markets Review **30(2)**, 150-165 (2023).

A. Smith and B. Jones, Journal of Technology and Economics **34(2)**, 102-117 (2019).

J. Lee and Y. Kim, Global Economy Journal **22(4)**, 337-355 (2018).

K. Brown and J. Harris, Journal of Policy and Economics **17(2)**, 233-247 (2021).

M. Williams and L. O'Donnell, Journal of Online Retailing **5(3)**, 200-216 (2022).

R. Davis and H. Chung, Brand Management Journal **19(4)**, 112-130 (2022).

R. Thompson, et al., Market Dynamics Review **12(3)**, 200-215 (2021).

S. Gupta, P. Malik, Journal of Business Strategies **29(1)**, 78-92 (2022).

T. Green and S. Patel, Journal of Sustainable Production **18(5)**, 275-289 (2020).

Z. Y. Chen and L. Y. Cheng, Computer Science and Applications **13(8)**, 1538-1546 (2023).