

The Evolution and Future Trend Analysis of Consumers Preference for Mobile Phone Hardware

Yukuan Jerry Li

Guanghua Cambridge International School, Shanghai, 201204, China

1812010507@stu.hrbust.edu.cn

Abstract. The mobile phone market has experienced continuous expansion, making the price of mobile phones an increasingly crucial factor to discuss. This paper explores the impact of mobile phone hardware on pricing, with a focus on changes over time. Using multiple linear regression models, the study analyses two datasets from Kaggle: one detailing 2023 phone prices and another predicting future prices. Key hardware factors examined include storage, RAM, screen size, and battery capacity. Comparisons between models are made, and reveal that while storage, RAM, and screen size positively correlate with prices, battery capacity shows a negative correlation in 2023. In contrast, predicted future prices show positive correlations for all factors except screen size, which becomes statistically insignificant. The study concludes that the influence of hardware on phone prices is expected to decrease, and consumers are likely to prioritise storage and battery capacity more among hardware. For future research, more relevant data could be taken to make the result more accurate.

Keywords: Multiple linear regression; consumer preference; mobile phone hardware; mobile phone price.

1. Introduction

The market of mobile phone has been expanding constantly since the invention of phones. Between 2000 and 2005, the number of mobile phone users nearly doubled, and in 2005, more than two-thirds of adults in the USA had mobile phones [1]. In Saudi Arabia now, every 100 citizens own approximately 180 smartphones [2]. Under such circumstance, the price of phones becomes an increasingly important topic to discuss as it's the most important factor that affect consumer buying behaviour [3].

There are already some researches on the factors that affect the mobile phone prices. For example, hedonic price function model and multiple linear regression model have been applied to research the determinants of the mobile phone prices [4, 5]. While taking the fact that mobile phone technology evolves rapidly into consideration, this essay believes that previous models cannot be sufficiently predictive for current mobile phone price as new technologies change people's preferences and affects the weight of each factor related to phone prices. Therefore, the essay is interested in exploring the leverage of mobile phone hardware on price and compare their importance over time.

This essay refers to two sets of data on phone prices from Kaggle. One explained the newest mobile phone price in 2023, and the other is the predicted price of future mobile phones. The reliability of both datasets has been certified as they are compatible with multiple sources of information and other price prediction machine learning algorithms models, including the decision tree method, and M5P models [6-8]. Feature reduction method is also used to help predict the phone price more accurately [9].

In this paper, multiple linear regression model is used on both datasets to quantify the influence of each factor on price [10]. This essay mainly chooses hardware of phones to be the research focus. Thus, the independent variables in the regression are storage, random-access memory (RAM), screen size, and battery capacity, and the dependent variable is the phone price. The coefficients of independent variables in the two regressions based on those two datasets are believed to be different, which indicates consumer preferences change throughout the decades.

Overall, this article will use the multiple linear regression model to find the correlation between mobile phone prices and storage, RAM, screen size, and battery capacity of the mobile phone. The paper will also compare the weight changes of each factor and analyse the change of consumers' preference for different hardware of mobile phones over time.

2. Methods

2.1. Data Source

The two datasets used in this essay are from Kaggle website. The first dataset includes information on the prices and hardware details like storage of smartphones from different brands in 2023. The other dataset predicts the future mobile phone price and also includes details like storage and RAM of the phones. There are 407 sets of data in the first dataset, and 807 sets of data in the second. Another hedonic price function model researched in 2013 is also taken to make comparison [4].

2.2. Variable Selection

Among all potential variables on phone hardware, this paper focuses on storage, RAM, screen size, and battery capacity to investigate their influence on prices, due to the accessibility and the accuracy of those variables. The units of the same variables in the two datasets are unified. Price is in US dollars; Storage and RAM are in GB; Battery capacity is in mAh; Screen size is measured in inches. Detailed information is shown in Table 1.

Table 1. Different types of variables

Items	Dataset	Range	Mean value	Std. Deviation
Storage (GB)	1	32~512	123.047	64.963
	2	2~256	64.39	53.869
RAM (GB)	1	2~16	5.838	2.432
	2	0~12	5.942	2.056
Screen Size (inches)	1	4.5~10.7	6.481	0.383
	2	2~44	5.621	3.957
Battery Capacity (mAh)	1	1821~7000	4676.477	797.194
	2	1020~6000	3278.86	937.629
Price (\$)	1	99~1999	402.196	292.183
	2	6~1828	170.533	275.949

In this table 1, dataset 1 represents that this dataset contains information in 2023, and dataset 2 represents that it is the prediction of mobile phone price.

2.3. Method Introduction

Multiple linear regression model is used to identify the quantitative correlation between variables. Phone price is chosen to be the dependent variable, as this paper is determined to research the other factors' influence on price. The rest are chosen to be the regressors. To ensure the validity of the model, there are 1214 sets of data used in total. With the help of SPSS, this paper succeeded in accurately processing these numerous amounts of data [10]. There are two regressions obtained from the datasets. Each represents how the change in the hardware condition affects phone prices through the coefficients of the variables.

$$\text{Price (\$)} = \beta_1 + \beta_2 \times \text{Storage (GB)} + \beta_3 \times \text{RAM (GB)} + \beta_4 \times \text{Screen Size (inches)} + \beta_5 \times \text{Battery Capacity (mAh)} \quad (1)$$

The coefficients in the two regressions will be compared to illustrate the changes of each variable's ability to influence price, which implies the predicted changes of consumers' preferences for mobile phone hardware. Similarly, to investigate the overall trend of change in consumer preferences

throughout the decade, another comparison between the phone price model in 2023 and in 2013 is conducted.

3. Results and Discussion

3.1. Correlation Analysis

Before using the multiple linear regression model, correlation analysis is conducted on the two datasets (figure 1 and 2).

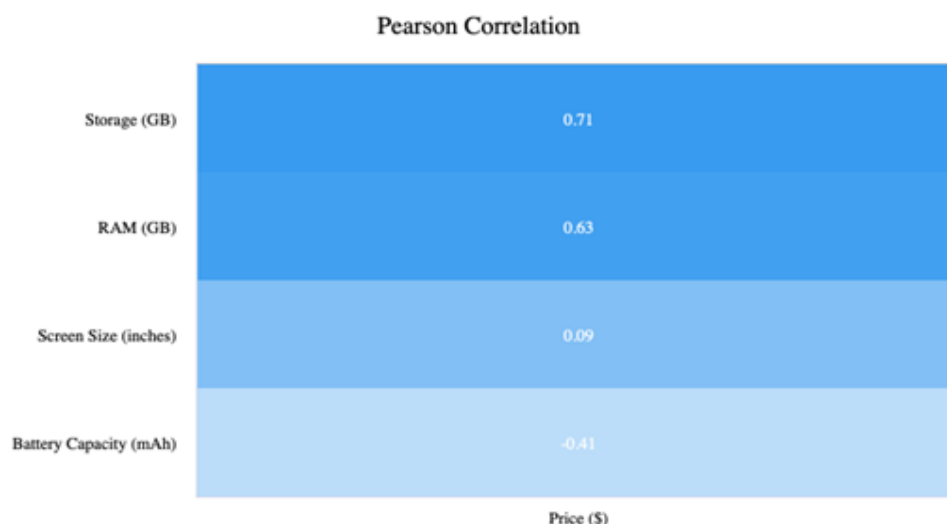


Fig. 1 Pearson correlation result for dataset 1

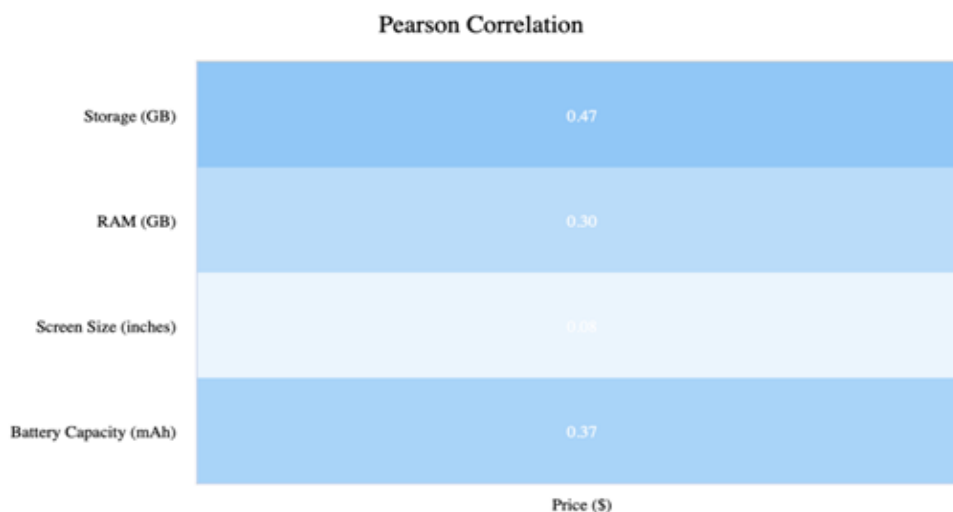


Fig. 2 Pearson correlation result for dataset 2

The preliminary correlation analysis shows that screen size correlate relatively weakly with price in both datasets. Most of the variables are in positive correlation with mobile phone price, except that battery capacity in the first dataset has a negative impact on price.

3.2. Multiple Linear Regression

For the first dataset, the model is that $\text{Price (\$)} = 190.723 + 2.144 \times \text{Storage (GB)} + 32.047 \times \text{RAM (GB)} + 72.037 \times \text{Screen Size (inches)} - 0.151 \times \text{Battery Capacity (mAh)}$.

The analysis shows that storage, RAM, and screen size are all positively correlated with mobile phone prices, while there is a negative relationship between the battery capacity and the phone price. The value of R^2 is 0.680, and the F test is performed ($F=211.058$, $p=0.000$). VIF value for each

factor is smaller than five, and D-W value is 1.972. P value of every factor is less than 0.010. This confirms the validity and accuracy of the regression (Table 2).

Table 2. Multiple linear regression result for Dataset 1

	Coefficient	Std. Error	t-value	p-value	VIF
Constant	190.723	145.192	1.314	0.19	-
Storage (GB)	2.144	0.18	11.943	0.000**	1.957
RAM (GB)	32.047	4.772	6.716	0.000**	1.912
Screen Size (inches)	72.037	26.645	2.704	0.007**	1.518
Battery Capacity (mAh)	-0.151	0.012	-12.274	0.000**	1.403
R ²	0.68				
Adj R ²	0.676				
F	F (4,399) = 211.508, p = 0.000				
D-W value	1.972				

The model resulted from the second dataset $\text{Price (\$)} = -311.416 + 1.620 \times \text{Storage (GB)} + 24.906 \times \text{RAM (GB)} + 1.354 \times \text{Screen Size (inches)} + 0.068 \times \text{Battery Capacity (mAh)}$.

The value of R² is 0.292, and F test is also used (F=82.747, p=0.000). Likewise, VIF value of each independent variable is smaller than 5, and D-W value is around two. This confirms that the coefficients are jointly significant in the regression. Table 3 shows the detailed results:

Table 3. Multiple linear regression result for Dataset 2

	Coefficient	Std. Error	t-value	p-value	VIF
Constant	-311.416	39.983	-7.789	0.000**	-
Storage (GB)	1.62	0.179	9.064	0.000**	1.379
RAM (GB)	24.906	4.26	5.847	0.000**	1.142
Screen Size (inches)	1.354	2.093	0.647	0.518	1.021
Battery Capacity (mAh)	0.068	0.01	6.973	0.000**	1.234
R ²	0.292				
Adj R ²	0.289				
F	F (4,802) =82.747, p=0.000				
D-W value	1.815				

However, the p value of screen size is larger than 0.05, which indicates that the influence of screen size on price in this set of data cannot be significantly quantified by the coefficient. Therefore, the model excludes screen size and only discuss the influence of storage, RAM, and battery capacity. The new model is written downwards, with each factor being positively related with the phone prices.

$$\text{Price} = -305.903 + 1.627\text{Storage} + 24.859\text{RAM} + 0.068\text{Battery Capacity} \quad (2)$$

The new result will be shown in Table 4:

Table 4. New multiple linear regression result for Dataset 2

	Coefficient	Std. Error	t-value	p-value	VIF
Constant	-305.903	39.05	-7.834	0.000**	-
Storage (GB)	1.627	0.178	9.122	0.000**	1.374
RAM (GB)	24.859	4.258	5.839	0.000**	1.141
Battery Capacity (mAh)	0.068	0.01	7.067	0.000**	1.223
R ²	0.292				
Adj R ²	0.289				
F	F (3,803) =110.270, p=0.000				
D-W value	1.816				

3.3. Model Comparison

The first model is based on mobile phone price in 2023, and the second model is based on predicted prices (Figure 3 and 4).

$$\text{Price (\$)} = 190.723 + 2.144 \times \text{Storage (GB)} + 32.047 \times \text{RAM (GB)} + 72.037 \times \text{Screen Size (inches)} - 0.151 \times \text{Battery Capacity (mAh)} \quad (3)$$

$$\text{Price (\$)} = -305.903 + 1.627 \times \text{Storage (GB)} + 24.859 \times \text{RAM (GB)} + 0.068 \times \text{Battery Capacity (mAh)} \quad (4)$$

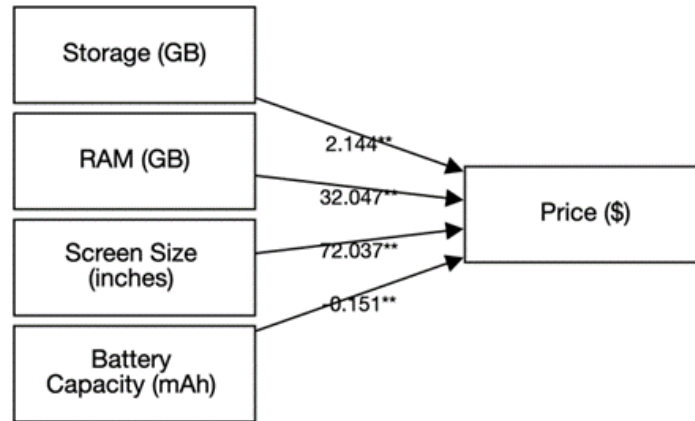


Fig. 3 First model result

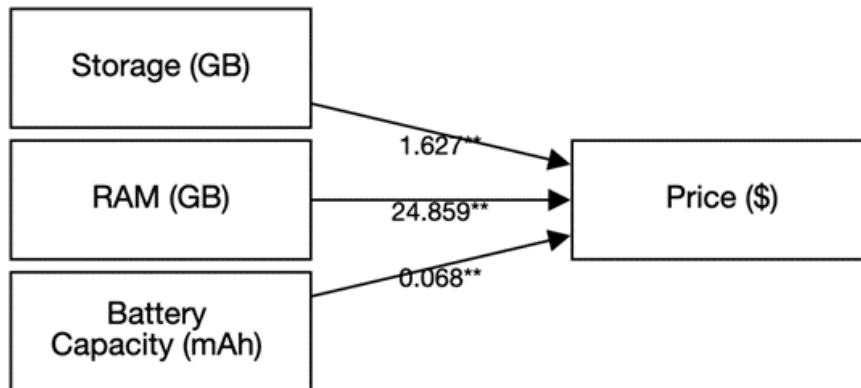


Fig. 4 Second model result

Generally, the absolute values of all coefficients decrease, which can show that influence of mobile phone hardware on price drops. The battery capacity affects the price positively in the second model while its performance in the first model is opposite. This illustrates that consumers might pay more attention to battery capacity in the future. Despite all coefficients experience a fall in numerical value, the influence of each variable on price change differently. The relative impact of the discussed variables on price is shown by dividing the coefficient of one variable with the sum of the absolute values of all coefficients. For example, in 2023, the weight of storage is $\frac{2.144}{2.144+32.047+72.037+|-0.151|} = 0.020$, and the predicted weight of storage is $\frac{1.627}{1.627+24.859+0.068} = 0.061$. This method of calculating the proportion of the coefficient enables the analysis of the change of consumer preferences among the listed hardware conditions. The comparative influence of storage on price is likely to increase significantly. Likewise, RAM goes up in the weight of hardware factors affecting prices as well, from 0.301 to 0.936. Consumers are predicted to concern more about the phone memory. The R^2 value drops significantly from 0.680 to 0.292 in the two different models. This shows that the second model is not accurate enough. It may imply that the hardware factors researched will have much less influence on phone prices. Consumer may prefer factors other than storage, RAM, size and battery capacity in the future.

In the research conducted in 2013, the hedonic price function shows that there is a negative relationship between phone size and prices [4]. The weight of phone memory in influencing the price among size, phone memory, and battery duration is $\frac{7.15 \times 10^{-5}}{7.15 \times 10^{-5} + 4.93 \times 10^{-8} + 0.0007} = 0.953$, and that of battery duration is $\frac{0.0007}{7.15 \times 10^{-5} + 4.93 \times 10^{-8} + 0.0007} = 0.014$ [4]. Compared to the figure in the first model based on the 2023 prices, people consider less about the phone memory and battery capacity, but more on size now.

Overall, consumers place a higher value on screen size now, but consider less about the phone memory and battery capacity. In the future, consumers will pay more attention to the battery according to the prediction data. Accordingly, people will also prefer phones with larger storage, and the influence of hardware on phone prices will decrease.

3.4. Critical Thinking

In the first model based on the dataset that includes 2023 mobile phone prices, the coefficient of battery capacity is negative, which shows that there is a negative correlation between phone price and battery capacity. Despite its being counter-intuitive, several reasons can be listed to explain this. The first possible explanation is, due to the appearance of the portable charger, consumers become less conscious battery capacity, and thus the positive correlation between battery capacity and price is weakened. Moreover, higher battery capacity is likely to relate with higher weight of the phone. With the clear preference on portability and lightness of phones, and battery capacity might be sacrificed. Furthermore, the same result is acquired by Mohd et al. [5]. This supports that the sample and data used in the model is not biased to affect the result of the model. Meanwhile, R^2 in the second model is 0.292, showing that the regression cannot be representative to the population, which can be resulted from the inadequate sample size.

In addition, in the second model, the variable of screen size is deleted, as the p value indicates its insignificance. This might be a result from the limited sample size. However, it might be the irrelevance of screen size itself that causes the insignificance. Recent research on factors includes battery capacity, camera, screen size, charging speed, weight and model age that affect phone prices, screen size is shown to be irrelevant with the price [5]. Screen size can be irrelevant to phone prices as some smartphones have a folding screen function. The size of screen can be changed. Another possible reason is that in the future, consumers will pay more attention to the functions and software of the smartphones than the hardware. Screen size will have little impact on phones.

3.5. Limitations and Suggestions

In this paper, only two datasets from Kaggle are taken. Although the reliability is verified, more data can be found to make the multiple linear regression results more universal and valid. R^2 value in the second model is too small, and more predicted data should be taken as well. When using the two datasets from Kaggle, data screening is not conducted, which means some outliers are not identified. The validity of the model will be affected. To make the result more accurate, data filtration is necessary. Additionally, this essay only compares the results in 2013. More references related to factors affecting mobile phone price can be taken to make comparison to find a more general trend in consumers' preferences for different mobile phone hardware.

4. Conclusion

This paper analyses two different mobile phone price datasets from Kaggle and obtains two multiple linear regression models with the help of SPSS. The first model based on the 2023 phone prices shows that storage, RAM, and screen size have a positive impact on the phone prices while battery capacity is in negative correlation with the price. The other model which is based on the predicted mobile phone price data indicates storage, RAM, and battery capacity will positively affect the price, and the relationship between screen size and price is not certain.

Compared with the model gained in 2013, the weights of storage, RAM and battery capacity all decrease, but are predicted to increase among the hardware factors. However, R^2 value decreases in the model based on the prediction data, which may indicate that the hardware factors will have less influence on prices in the future. This result shows that among those factors, consumers now concern more about screen size but less about phone memory and battery. In the future, consumers are forecasted to place a lower value on hardware, but among different hardware, they will pay more attention to storage and battery capacity.

From this article, the general changes of consumers' preferences for storage, RAM, screen size and battery capacity are concluded. The predicted future trend in consumer preferences is also included. For future research, more relevant data can be collected to help make the model and the analysis of the changes of consumers' preferences for mobile phone hardware be more accurate and valid. Some different models can also be used to help find the more general trends and influence factors.

References

- [1] Tuckel P, O'Neill H. Ownership and usage patterns of cell phones: 2000-2005. Second international conference on telephone survey methodology, 2006, 13.
- [2] Trabelsi E. Saudi has the world's largest number of mobile phone users: U.N. report. Arabiya, 2024.
- [3] Sata M. Factors affecting consumer buying behavior of mobile phone devices. Mediterranean Journal of Social Sciences, 2013, 4(12): 103-112.
- [4] Mostafavi S M, Roohbakhsh S S, Behname M. HEDONIC PRICE FUNCTION ESTIMATION FOR MOBILE PHONE IN IRAN. International Journal of Economics & Financial Issues, 2013, 3(1): 202-205.
- [5] Mohd Hazhar M H, et al. Determinants of smartphone prices using backward elimination technique in multiple linear regression. Mathematical Sciences and Informatics Journal (MIJ), 2024, 5(1): 72-82.
- [6] Bishnoi A K, Mandal S K. Performance Analysis of Decision Tree Models and M5P Models for Mobile Phone Price Prediction. 2023 International Conference on Advances in Computation, Communication and Information Technology (ICAICCIT), 2024.
- [7] Çetin Mustafa, Yunus Koç. Mobile phone price class prediction using different classification algorithms with feature selection and parameter optimization. 2021 5th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT). IEEE, 2021.
- [8] Muhammad Asim, Zafar Khan, Mobile Price Class prediction using Machine Learning Techniques. International Journal of Computer Applications, 2018, 179.
- [9] Chen M. Mobile Phone Price Prediction with Feature Reduction. Highlights in Science, Engineering and Technology, 2023.
- [10] Zhou Jun, Ma Shipeng. SPSSAU research data analysis methods and applications. Publishing House of Electronics Industry, 2024.