

A Report On: -

Mobile Price Classification

|  |  |
| --- | --- |
| **Submitted By: -**  **Rajan Ghimire** | **Submitted To: -**  **Victoria Shtern** |

# Introduction

Mobile devices are widely used in everyday life, and the characteristics of these devices are the essential factor that may affect the overall performance and usability. This dataset contains all the features which define the capabilities of the mobile device such as battery, processor speed, memory, camera, network and physical size of the device. The goal of this report is to perform Variable Analysis and Exploratory Data Analysis to find out hidden insights, trends and patterns form data. We want to find out the general characteristics of the distribution and dependencies between these parameters that will help to make the correct decisions in the framework of product development, pricing and marketing.

# Problem Statement

The purpose of this exploratory data analysis is to identify trends within a comprehensive dataset of mobile-device specifications for company managers to effectively develop, price and sell their products within the mobile market niche.

# About the Dataset

Bob has started his own mobile company. He wants to give tough fight to big companies like Apple, Samsung etc. He does not know how to estimate the price of mobiles his company creates. In this competitive mobile phone market, you cannot simply assume things. To solve this problem, he collects sales data of mobile phones of various companies. Bob wants to find out some relation between features of a mobile phone (eg:- RAM,Internal Memory etc.) and its selling price

# 

# Feature Analysis

# A screenshot of a document Description automatically generated

# Exploratory Data Analysis

# Missing Values in Data:

# The dataset contains no such missing values.

|  |  |  |
| --- | --- | --- |
|  | **Column** | **Missing\_Percentage** |
| **0** | battery\_power | 0 |
| **1** | blue | 0 |
| **2** | clock\_speed | 0 |
| **3** | dual\_sim | 0 |
| **4** | fc | 0 |
| **5** | four\_g | 0 |
| **6** | int\_memory | 0 |
| **7** | m\_dep | 0 |
| **8** | mobile\_wt | 0 |
| **9** | n\_cores | 0 |
| **10** | pc | 0 |
| **11** | px\_height | 0 |
| **12** | px\_width | 0 |
| **13** | ram | 0 |
| **14** | sc\_h | 0 |
| **15** | sc\_w | 0 |
| **16** | talk\_time | 0 |
| **17** | three\_g | 0 |
| **18** | touch\_screen | 0 |
| **19** | wifi | 0 |
| **20** | price\_range | 0 |

# Analysis of class variable ‘price range’:

|  |  |
| --- | --- |
| **price\_range** | **Count of price\_range** |
| 3 | 500 |
| 2 | 500 |
| 0 | 500 |
| 1 | 500 |
| **Grand Total** | **2000** |

# The price range is equally distributed among 4 classes. 0 means low cost, 1 means Medium Cost, 2 means high cost and 3 means very high cost.

# A pie chart of class variables.

# Analysis of battery power:

# A graph of blue and red lines

# 

# The graph shows the distribution of battery power. Here, the yellow line represents the minimum values if the distribution, red line depicts the average of the distribution and green line shows the maximum values of the distribution.

# A graph of blue rectangular bars Description automatically generated

# 

# Here, after looking at average battery power per each price range we can interpret that: There is positive co-relation between price range and battery power. As the battery power increases, the price of the mobile also increases.

# Analysis of ram:

# A graph with numbers and a line Description automatically generated with medium confidence

# The histogram shows the distribution of ram. Here, the yellow line represents the minimum values if the distribution, red line depicts the average of the distribution and green line shows the maximum values of the distribution.

A graph of blue squares

Description automatically generated

# Here, after looking at the average ram per each price range we can interpret that: There is positive co-relation between price range and ram. As the size of ram increases, the price of the mobile also increases.

# Analysis of 3G:

# The feature three\_g is a categorical variable with two classes 1 and 0. 1 represents presence of 3G feature in phone and 0 represents absence of meeting.

|  |  |
| --- | --- |
| **three\_g** | **Count of three\_g** |
| 1 | 1523 |
| 0 | 477 |
| **Grand Total** | **2000** |

# A blue circle with numbers and text Description automatically generated

# From the given donut plot, we can understand that most of the phones have 3G.

# A screenshot of a graph Description automatically generated

# Now, when we look at the count of 3G per price category, we can interpret that the price goes up when the phone has 3G.

# Analysis of 4G:

# The feature four\_g is a categorical variable with two classes 1 and 0. 1 represents presence of 4G feature in phone and 0 represents absence of meeting.

|  |  |
| --- | --- |
| **four\_g** | **Count of four\_g** |
| 1 | 1043 |
| 0 | 957 |
| **Grand Total** | **2000** |

# A blue circle with white text Description automatically generated

# From the given donut plot, we can understand that there is equal distribution between phone with and without 4G.

# A graph with orange and purple bars Description automatically generated

# Now, when we look at the count of 4G per price category, we can interpret that 4G feature does not have that much of difference in price of the mobile phone.

# Analysis of n\_cores:

# The n\_cores features represent the total number of processing cores within the system. The higher the number of processing cores, the better the performance.

|  |  |
| --- | --- |
| **n\_cores** | **Count of n\_cores** |
| 4 | 274 |
| 7 | 259 |
| 8 | 256 |
| 2 | 247 |
| 3 | 246 |
| 5 | 246 |
| 1 | 242 |
| 6 | 230 |
| **Grand Total** | **2000** |

# 

# The histogram shows the distribution of n\_cores. The black dotted line represents overall trend of the distribution.

# Now, when we look at the count of n\_cores per price category, we can interpret that n\_cores feature does not have that much of difference in price of the mobile phone. This means, mobile phones of different price range have same number of cores.

# A graph of different colored bars Description automatically generated

# Analysis of talk\_time vs battery\_power:

# In this analysis, we are trying to find out if mobile phones with longer talk time have longer battery life. This is a confirmatory data analysis where we want to confirm this hypothesis.

|  |  |
| --- | --- |
| **talk\_time** | **Count of talk\_time** |
| 7 | 124 |
| 4 | 123 |
| 16 | 116 |
| 15 | 115 |
| 19 | 113 |
| 6 | 111 |
| 10 | 105 |
| 8 | 104 |
| 11 | 103 |
| 20 | 102 |
| 14 | 101 |
| 9 | 100 |
| 13 | 100 |
| 18 | 100 |
| 12 | 99 |
| 2 | 99 |
| 17 | 98 |
| 3 | 94 |
| 5 | 93 |
| **Grand Total** | **2000** |

# A screenshot of a graph Description automatically generated

Here we have two-line charts that show count of talk time and battery for different bins. The red line is to indicate overall trend in the talk time.

# A screenshot of a graph Description automatically generated

Here If we select lower battery power, the overall talk time is decreasing shown by the red trend line in talk time.

# A graph of a graph Description automatically generated with medium confidence

# Similarly, when we select higher talk time, we can see that overall talk time is increasing shown by the red trend line which proves our hypothesis.

# Conclusion

The exploratory data analysis on the mobile device dataset has yielded valuable insights into the relationships between various features and the price range of the devices. The analysis revealed a positive correlation between battery power, RAM capacity, and the device's price range, suggesting that higher-end devices tend to offer larger batteries and more RAM. Interestingly, the presence of 3G connectivity appeared to influence pricing, while 4G connectivity and the number of processor cores did not significantly impact pricing across different price categories. Additionally, the analysis confirmed the expected behavior that lower battery power corresponds to shorter talk times, and vice versa. These findings can inform product development strategies, pricing decisions, and marketing efforts, enabling manufacturers and retailers to better align their offerings with consumer preferences and market dynamics within the highly competitive mobile device.

# References

# Iabhishekofficial. (n.d.). Mobile Price Classification. Kaggle. Retrieved June 5, 2024, from <https://www.kaggle.com/datasets/iabhishekofficial/mobile-price-classification>

# Dashboard

# A close-up of a chart Description automatically generated

A screenshot of a graph

Description automatically generated