INNOMATICS
RESEARCH LABS

INNOVATION. AUTOMATION. ANALYTICS

PROJECT ON Flipkart

Title: Web Scraping and Exploratory Data Analysis of Flipkart Mobile Data

PRESENTED BY
SHAIK HUSSAIN SIDDIK
YENNAM RAM BRAMHA REDDY
KAYALA DURGA NAGENDRA

INNOMATICS
RESEARCH LABS

Preface:

INNOMATICS
RESEARCH LABS

Introduction:

- Flipkart is one of India's leading e-commerce platforms, offering a wide range of products including electronics, fashion, and home essentials.

- In the mobile segment, Flipkart provides detailed information such as price, discounts, ratings, reviews, and specifications, helping customers make informed decisions.

- The platform generates rich data that can be analyzed to understand customer behavior, market trends, and pricing strategies.

Business Problem:

- Flipkart offers a wide variety of mobile phones with different prices, brands, and specifications, making it challenging to understand what truly drives customer purchases and satisfaction.

- The platform needs to identify how pricing, discounts, brand reputation, and hardware features (RAM, storage) impact customer ratings and engagement.

- There is a need to analyze large volumes of product data to optimize pricing strategies, improve product visibility, and promote value-for-money devices.

Objectives:

- To analyze mobile phone data from Flipkart to understand overall pricing, ratings, and product distribution.

- To identify how hardware features such as RAM and internal storage influence mobile pricing.

- To study the relationship between price, discounts, and customer satisfaction (ratings).

- To evaluate brand performance based on price range, ratings, and customer engagement.

- To uncover insights that help identify value-for-money smartphones and support better business decisions.

Web Scraping:

- Flipkart Mobile data was selected as the data source

- Used **browser developer tools** (Inspect Element) to locate relevant HTML tags Data.

- Used **BeautifulSoup and Requests** to extract data from Flipkart.

- Sent **HTTP requests** to fetch web pages.

- Performed **Exploratory Data Analysis (EDA)** to identify trends.
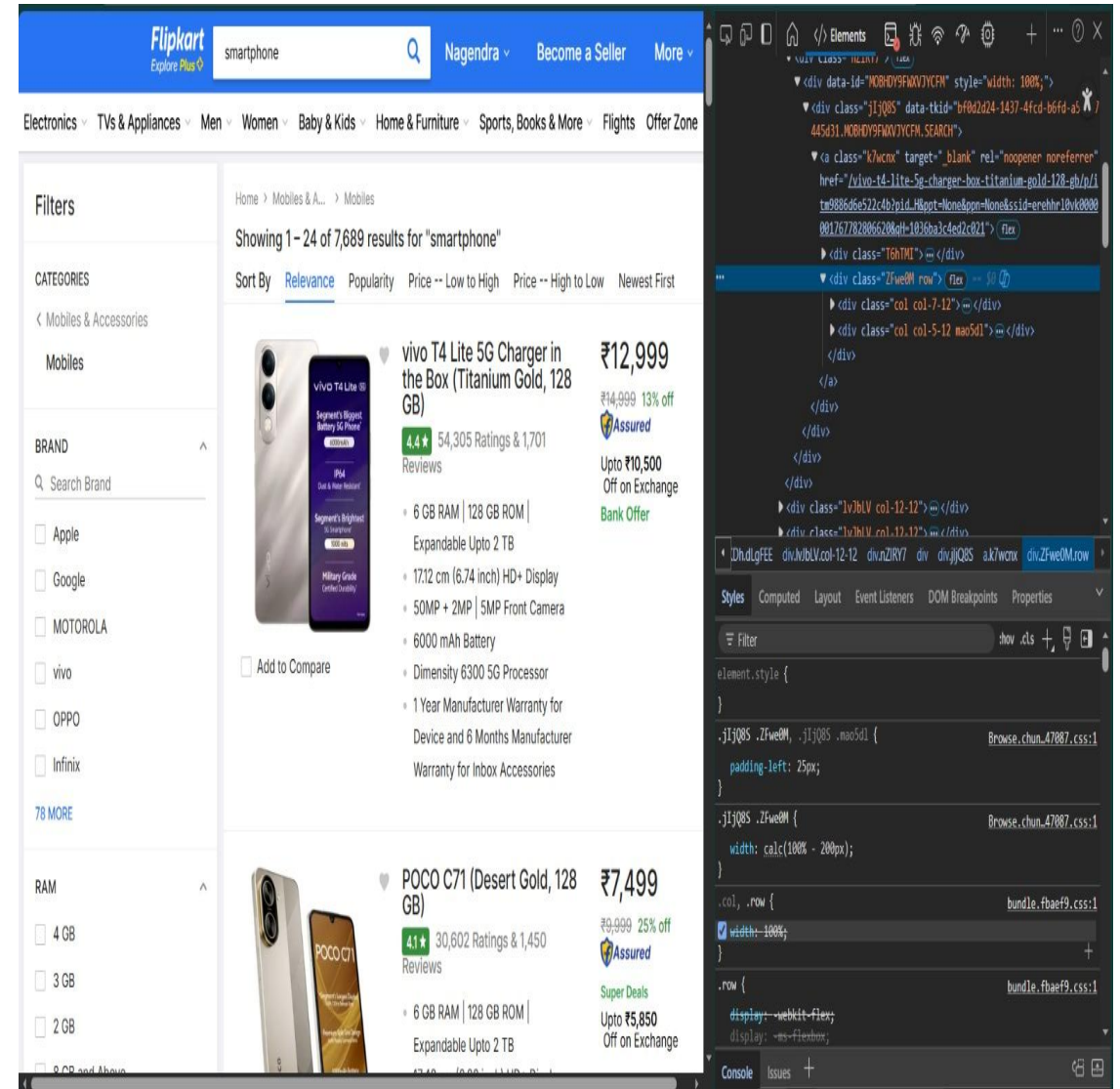
Dataset:

	Unnamed: 0	Product_Name	Rating	No_Of_Ratings	Specifications	Price	Original_Price	Discount	Page
0	0	Google Pixel 9A (Porcelain, 256 GB)	4.4	4,870 Ratings & 334 Reviews	8 GB RAM | 256 GB ROM15.96 cm (6.285 inch) Ful...	₹39,999	₹49,999	20% off	1
1	1	Google Pixel 9A (Obsidian, 256 GB)	4.4	4,870 Ratings & 334 Reviews	8 GB RAM | 256 GB ROM15.96 cm (6.285 inch) Ful...	₹39,999	₹49,999	20% off	1
2	2	POCO C71 (Desert Gold, 128 GB)	4.1	30,602 Ratings & 1,450 Reviews	6 GB RAM | 128 GB ROM | Expandable Upto 2 TB17...	₹7,499	₹9,999	25% off	1
3	3	Samsung Galaxy F07 (Green, 64 GB)	4.3	2,647 Ratings & 136 Reviews	4 GB RAM | 64 GB ROM | Expandable Upto 2 TB17....	₹6,999	₹9,999	30% off	1
4	4	Google Pixel 9A (Iris, 256 GB)	4.4	4,870 Ratings & 334 Reviews	8 GB RAM | 256 GB ROM15.96 cm (6.285 inch) Ful...	₹39,999	₹49,999	20% off	1
...	...	...	...	...	...	...	...	...	...
955	955	Samsung Galaxy M35 5G (Moonlight Blue, 256 GB)	4.3	5,761 Ratings & 363 Reviews	8 GB RAM | 256 GB ROM16.76 cm (6.6 inch) Displ...	₹24,000	₹27,499	12% off	41
956	956	Samsung Galaxy A25 5G (Yellow, 128 GB)	4.2	410 Ratings & 19 Reviews	8 GB RAM | 128 GB ROM | Expandable Upto 1 TB16...	₹18,749	₹28,499	34% off	41
957	957	Samsung Galaxy A17 5G (Gray, 128 GB)	4.3	261 Ratings & 12 Reviews	6 GB RAM | 128 GB ROM | Expandable Upto 2 TB17...	₹19,499	₹21,999	11% off	41
958	958	REDMI Note-14 Pro 5G (Phantom Purple, 128 GB)	4.3	525 Ratings & 24 Reviews	8 GB RAM | 128 GB ROM16.94 cm (6.67 inch) Disp...	₹23,490	₹28,999	18% off	41
959	959	Motorola Edge 50 Pro 5G with 68W Charger (Blac...	4.4	34,594 Ratings & 2,987 Reviews	8 GB RAM | 256 GB ROM17.02 cm (6.7 inch) Displ...	₹27,999	₹36,999	24% off	41

960 rows × 9 columns

Tools Used:

Data Cleaning Steps:

- **Removed unwanted characters** such as currency symbols, quotes and commas from text columns (e.g., product names).

- **Split and reformatted product names** for clean, readable entries.

- **Converted string-based prices** into numeric float values for analysis.

- **Standardized ratings and review counts** by removing text and characters like *R* and *&* by keeping numerical values.

- **Handled missing values** by replacing NaN entries with mode in columns like price, internal_storage.

|	Rating	Price	Original_Price	Discount	Page	Brand	Model	number_of_ratings	number_of_reviews	RAM	Internal_storage
0	4.4	39999	49999	20%	1	Google	Pixel 9A	4870	334	8	256
1	4.4	39999	49999	20%	1	Google	Pixel 9A	4870	334	8	256
2	4.1	7499	9999	25%	1	POCO	128 GB	30602	1450	6	128
3	4.3	6999	9999	30%	1	Samsung	Galaxy F07	2647	136	4	64
4	4.4	39999	49999	20%	1	Google	Pixel 9A	4870	334	8	256

Data Visualization:

- Ratings are heavily concentrated toward higher values, indicating overall positive customer satisfaction.

- Very few low-rated mobiles exist, showing quality consistency across products.

- High ratings suggest strong brand trust and user experience.

- Rating stability makes it a reliable metric for comparison.

- Customer feedback is largely favorable across models.


Distribution of Mobile Ratings
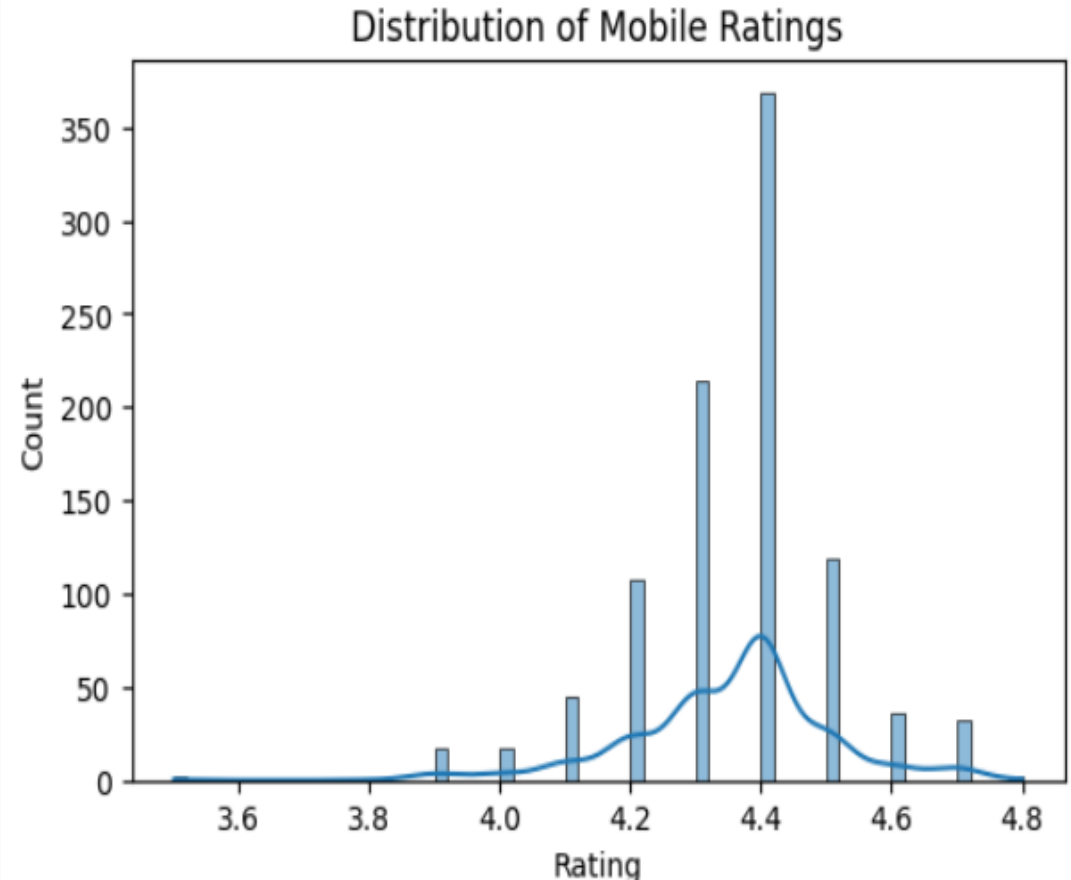
Data Visualization:

- Most mobiles fall within the low-to-mid price range.

- Premium-priced phones form a smaller but distinct segment.

- The distribution is right-skewed, indicating affordability dominance.

- Budget-friendly devices drive market volume.

- Pricing reflects diverse customer segments.

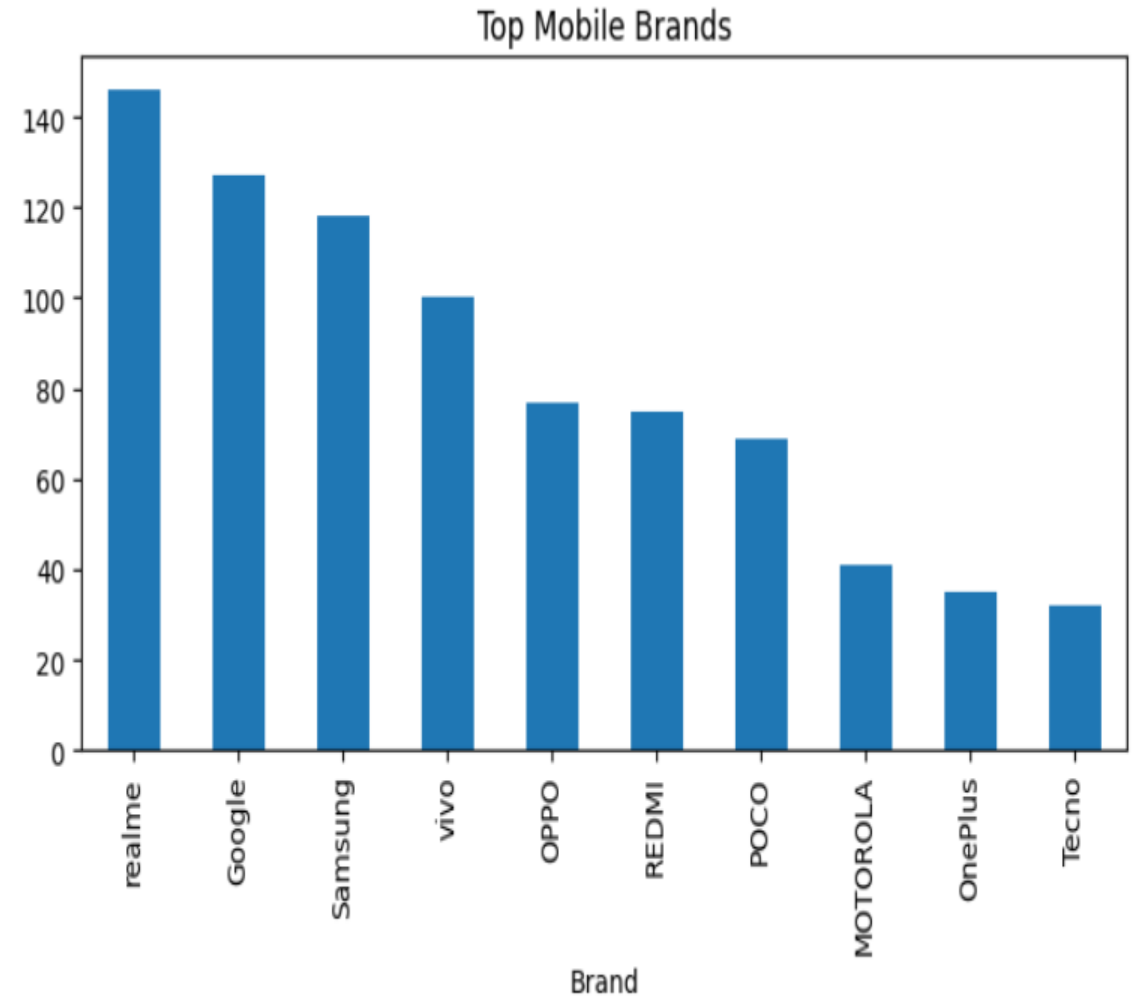Distribution of Mobile Prices

INNOMATICS
RESEARCH LABS
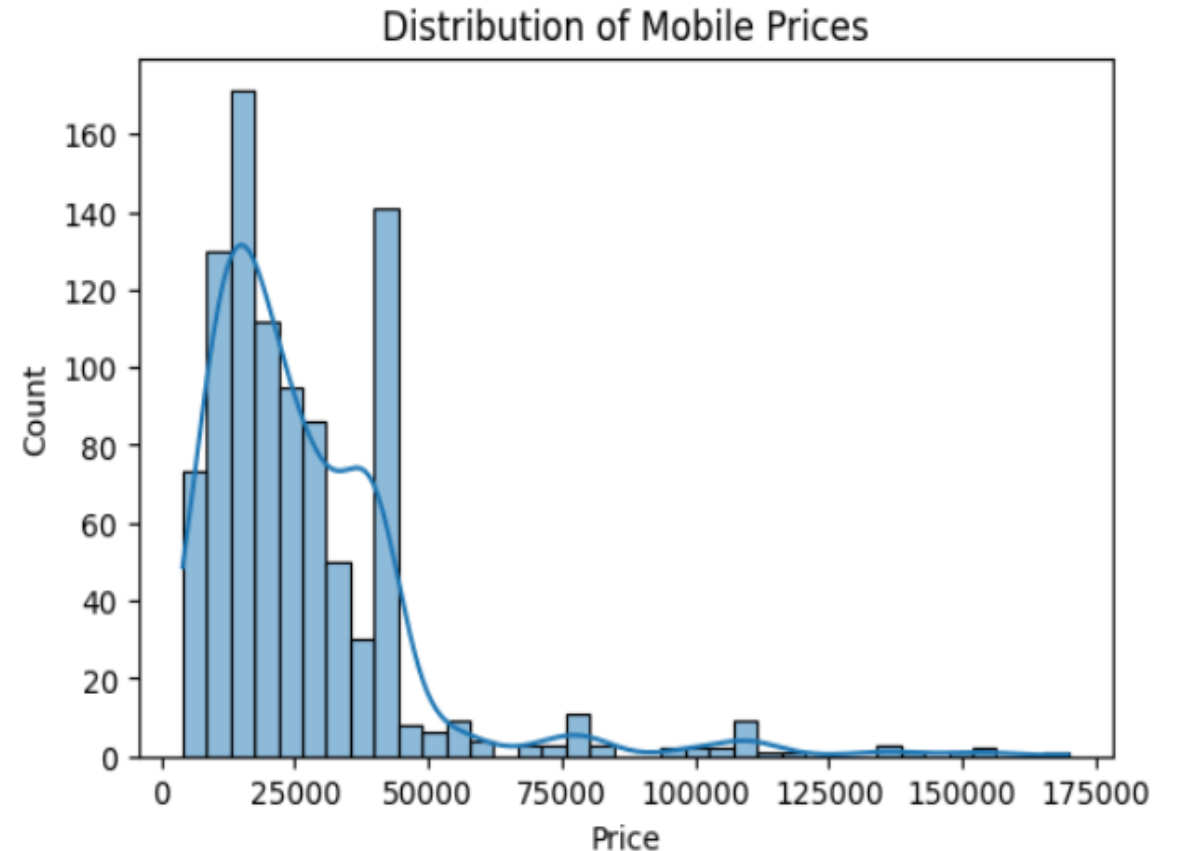
Data Visualization:

- Original prices vary widely across models.

- High original prices indicate premium positioning.

- Helps identify discount strategies used by brands.

- Shows manufacturer-intended pricing.

- Useful for price comparison analysis.

Boxplot of Original Price

INNOMATICS
RESEARCH LABS
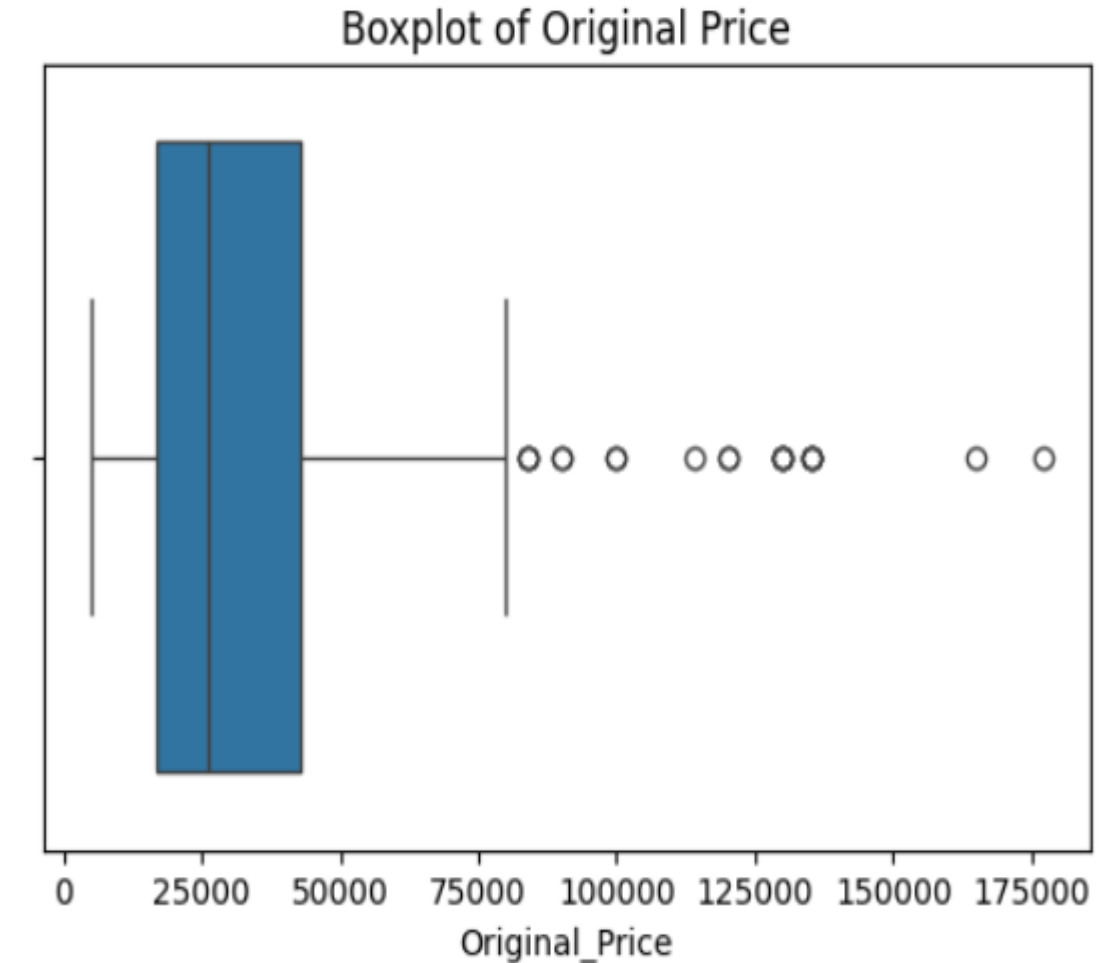
Data Visualization:

- A few brands dominate the dataset.

- Strong brand presence reflects customer trust.

- Popular brands list more models.

- Brand loyalty plays a major role.

- Market competition is brand-centric.

Top Mobile Brands — bar chart of model counts by brand: realme (~146), Google (~127), Samsung (~118), vivo (~100), OPPO (~77), REDMI (~75), POCO (~69), MOTOROLA (~41), OnePlus (~35), Tecno (~32)

Data Visualization:

- No strong relationship between price and rating.

- Expensive phones do not guarantee higher ratings.

- Affordable phones often receive high ratings.

- Satisfaction depends more on features than price.

- Value-for-money products perform well.

Price vs Rating

INNOMATICS
RESEARCH LABS

Data Visualization:

- Strong linear relationship observed.
- Discounts reduce prices proportionally.
- Premium phones remain expensive despite discounts.
- Pricing strategy is consistent.
- Indicates transparent discounting.

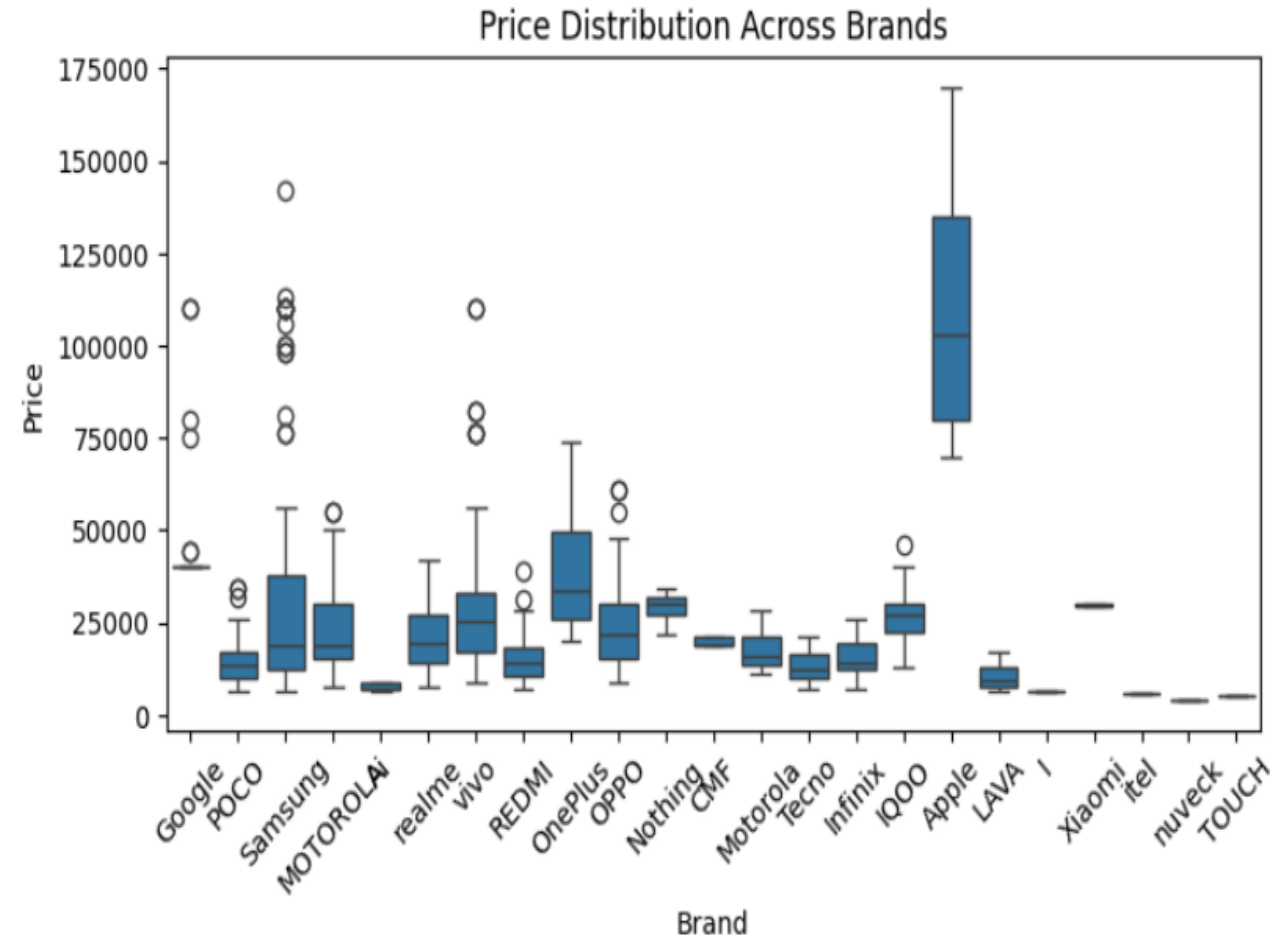Original Price vs Selling Price

Data Visualization:

- Brands follow different RAM strategies.

- Budget brands focus on lower RAM.

- Premium brands offer higher RAM.

- RAM choice aligns with brand positioning.

- Shows market segmentation.


Brand vs RAM Distribution
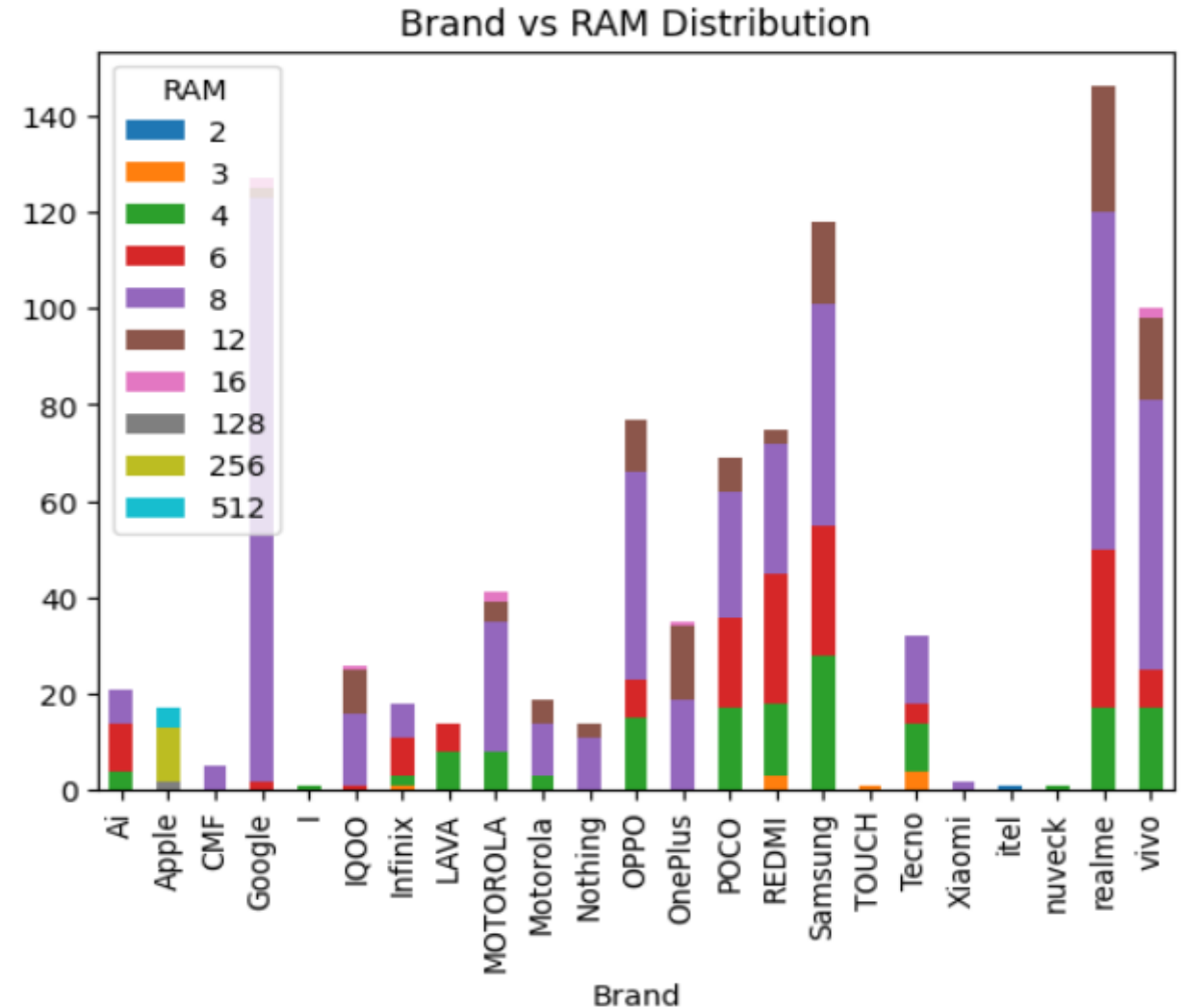
Data Visualization:

- Price distribution varies significantly across brands.

- Some brands target premium customers.

- Budget brands maintain affordable pricing.

- Outliers represent flagship models.

- Brand strongly influences price.

Price Distribution Across Brands

Data Visualization:

- Strong correlation between original and selling price.

- Ratings show weak correlation with price.

- Number of ratings and reviews are highly related.

- Discount has minimal impact on rating.

- Helps identify influential variables..

Correlation Heatmap of Numeric Features
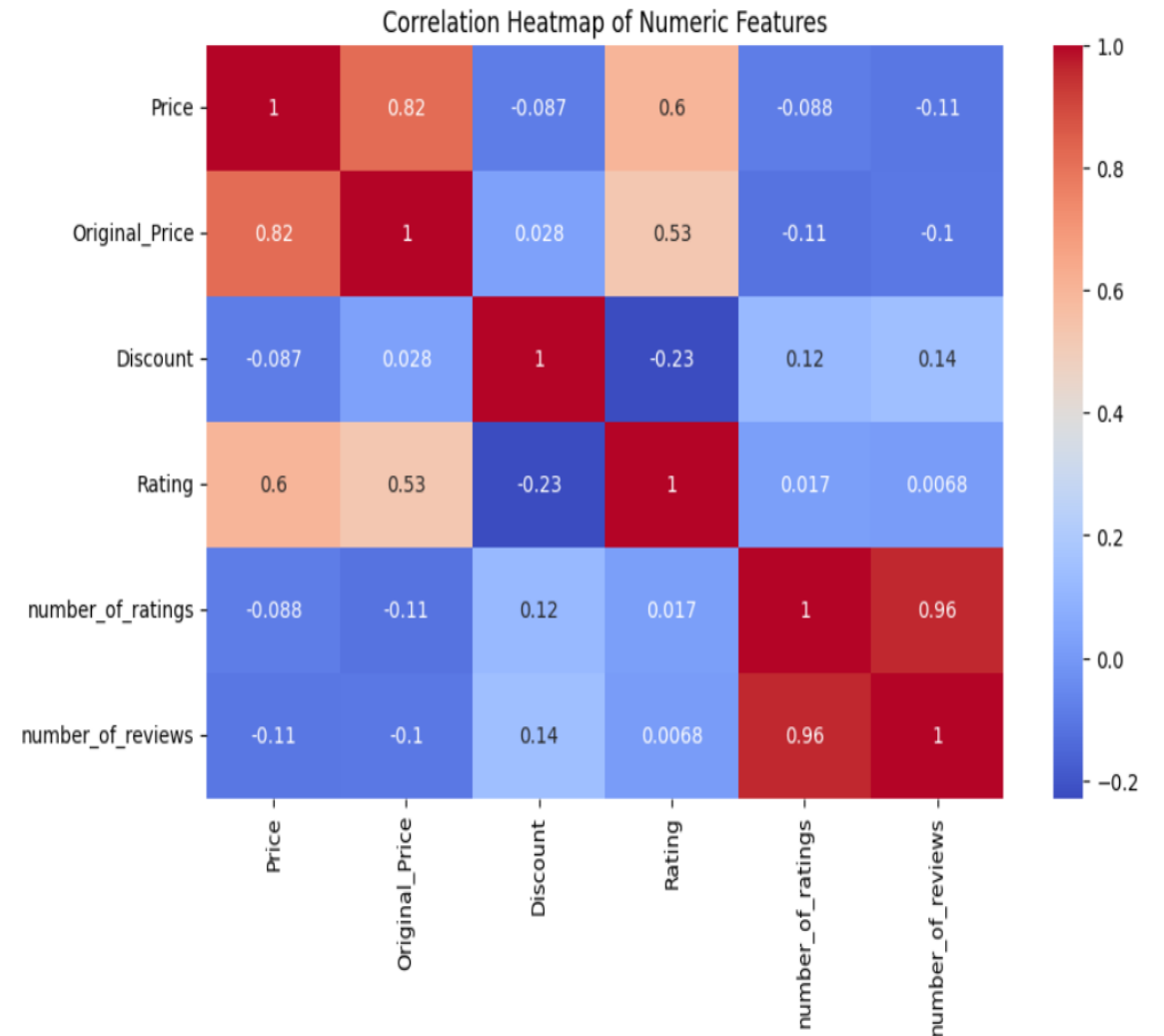
Data Visualization:

- Ratings remain stable across price ranges.

- Brand does not drastically affect ratings.

- Budget brands perform well in ratings.

- Satisfaction is feature-driven.

- Brand trust stabilizes ratings.

Price vs Rating across Brands

Conclusion:

1. The mobile market is mainly driven by budget and mid-range devices, with premium models forming a smaller segment.

2. Price is strongly influenced by hardware features such as RAM, internal storage, and brand positioning.

3. Customer ratings remain consistently high across all price ranges, showing satisfaction is not price-dependent.

4. Popular phones receive more ratings and reviews, indicating a strong link between customer engagement and product visibility.

5. Overall, the market favors value-for-money smartphones that balance performance, price, and user experience.

THANK YOU