



**INNOMATICS<sup>®</sup>**  
RESEARCH LABS

**INNOVATION. AUTOMATION. ANALYTICS**

## **PROJECT ON**

**Title:** Web Scraping and Exploratory Data Analysis With Flipkart Mobile Data

PRESENTED BY  
SENAPATHI BHANUTEJA



# Preface:

- **INTRODUCTION**
- **BUSINESS PROBLEM**
- **OBJECTIVE**
- **TOOLS USED**
- **WEB SCRAPPING**
- **DATA CLEANING STEPS**
- **DATA VISUALIZATION**
- **CONCLUSION**

# Introduction:



- Flipkart is a major player in India's online retail ecosystem, serving millions of customers across diverse product categories such as smartphones, electronics, fashion, and household goods.
- Within the smartphone category, Flipkart offers comprehensive product details including pricing, specifications, customer ratings, reviews, and promotional discounts, enabling users to compare devices effectively before purchase.
- This extensive availability of structured data makes the platform a valuable source for analysis, allowing insights into consumer preferences, brand performance, pricing patterns, and overall market dynamics in the Indian smartphone industry.

# Business Problem:

- Flipkart hosts a diverse range of smartphones across multiple brands, price segments, and hardware configurations, making it difficult to clearly determine the key factors influencing customer purchase decisions and satisfaction
- Understanding the impact of variables such as pricing, discounts, brand perception, and technical specifications (RAM, storage, processor) on customer ratings and user engagement is essential for effective decision-making
- Analyzing large-scale product and customer interaction data is therefore necessary to refine pricing strategies, enhance product discoverability, and highlight devices that offer the best value for money in a competitive marketplace.

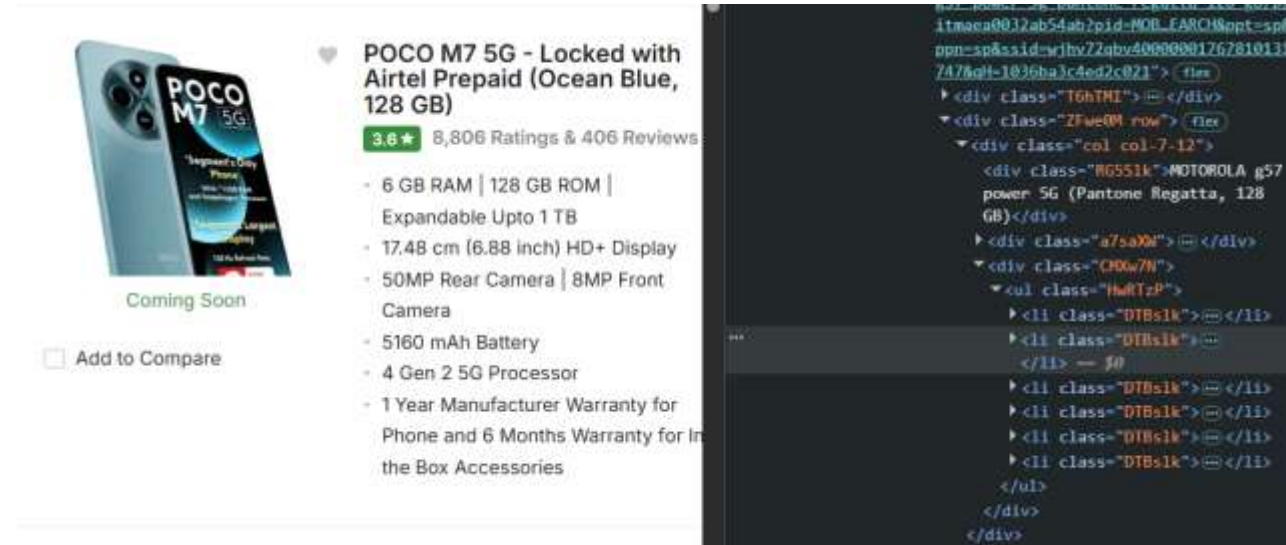


# Objectives:

- 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 Parse the data
- Sent **HTTP requests** to fetch web pages.
- Performed **Exploratory Data Analysis (EDA)** to identify trends.



# Dataset:

	Product_Name	Rating	No_Of_Ratings	Specifications	Price	Original_Price	Discount	Page
0	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
1	Samsung Galaxy F07 (Green, 64 GB)	4.3	2,647 Ratings & 136 Reviews	4 GB RAM   64 GB ROM   Expandable Upto 2 TB17....	₹7,849	₹9,999	21% off	1
2	MOTOROLA g35 5G (Midnight Black, 128 GB)	4.2	1,17,212 Ratings & 7,254 Reviews	4 GB RAM   128 GB ROM   Expandable Upto 1 TB17...	₹9,999	₹12,499	20% off	1
3	POCO C71 (Cool Blue, 64 GB)	3.9	54,664 Ratings & 2,462 Reviews	4 GB RAM   64 GB ROM   Expandable Upto 2 TB17....	₹6,799	₹8,999	24% off	1
4	Ai+ Pulse (Blue, 64 GB)	4.3	40,425 Ratings & 2,483 Reviews	4 GB RAM   64 GB ROM   Expandable Upto 1 TB17....	₹6,499	₹7,999	18% off	1
5	MOTOROLA g57 power 5G (Pantone Fluidity, 128 GB)	4.4	6,095 Ratings & 368 Reviews	8 GB RAM   128 GB ROM17.07 cm (6.72 inch) Full...	₹14,999	₹17,999	16% off	1
6	MOTOROLA g05 (Forest Green, 64 GB)	4.2	75,469 Ratings & 4,605 Reviews	4 GB RAM   64 GB ROM   Expandable Upto 1 TB16....	₹7,299	₹9,999	27% off	1
7	POCO M7 5G (Satin Black, 128 GB)	4.1	70,470 Ratings & 3,269 Reviews	6 GB RAM   128 GB ROM   Expandable Upto 1 TB17...	₹9,999	₹12,999	23% off	1
8	realme C61 (Marble Black, 64 GB)	4.3	40,868 Ratings & 1,318 Reviews	4 GB RAM   64 GB ROM   Expandable Upto 2 TB17....	₹7,249	₹8,999	19% off	1
9	POCO C71 (Desert Gold, 64 GB)	3.9	54,664 Ratings & 2,462 Reviews	4 GB RAM   64 GB ROM   Expandable Upto 2 TB17....	₹6,799	₹8,999	24% off	1

# Tools Used:

BeautifulSoup

 pandas

•[RegEx]\*

matplotlib

 NumPy

 seaborn



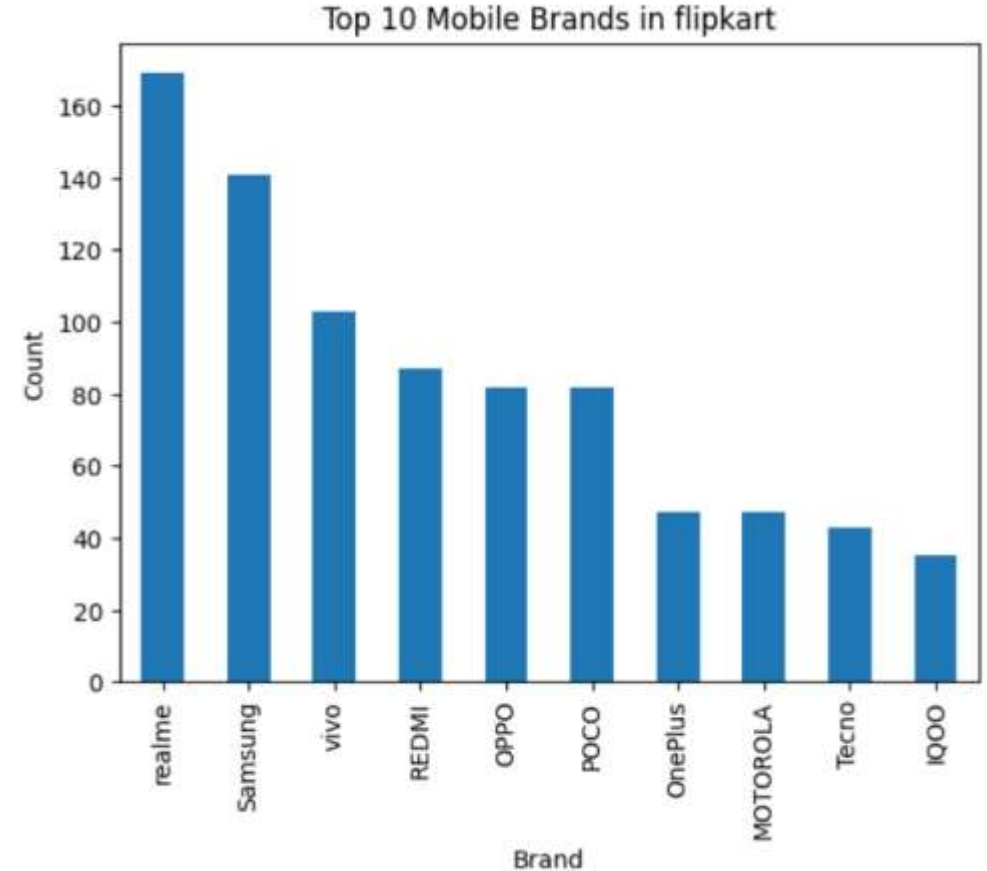
# Data Cleaning Steps:

- Handled missing values by replacing NaN entries with mode in columns like price,
- Extracted Brand of mobile from the product\_name column
- Using regex extracted RAM, Storage, Battery\_Capacity from the specifications column
- Split the No\_Of\_Ratings column into Reviews and Ratings
- Dropped unwanted columns
- Converted the every column into their respective datatype

	Rating	Price	Original_Price	Discount	Page	Brand	RAM	Storage	Battery_Capacity	number_of_ratings	number_of_reviews
0	4.1	7499	9999	25%	1	POCO	6	128	5200 mAh	30602	1450
1	4.3	7849	9999	21%	1	Samsung	4	64	5000 mAh	2647	136
2	4.2	9999	12499	20%	1	MOTOROLA	4	128	5000 mAh	117212	7254
3	3.9	6799	8999	24%	1	POCO	4	64	5200 mAh	54664	2462
4	4.3	6499	7999	18%	1	Ai	4	64	5000 mAh	40425	2483
5	4.4	14999	17999	16%	1	MOTOROLA	8	128	7000 mAh	6095	368
6	4.2	7299	9999	27%	1	MOTOROLA	4	64	5100 mAh	75469	4605
7	4.1	9999	12999	23%	1	POCO	6	128	5160 mAh	70470	3269
8	4.3	7249	8999	19%	1	realme	4	64	5000 mAh	40868	1318
9	3.9	6799	8999	24%	1	POCO	4	64	5200 mAh	54664	2462

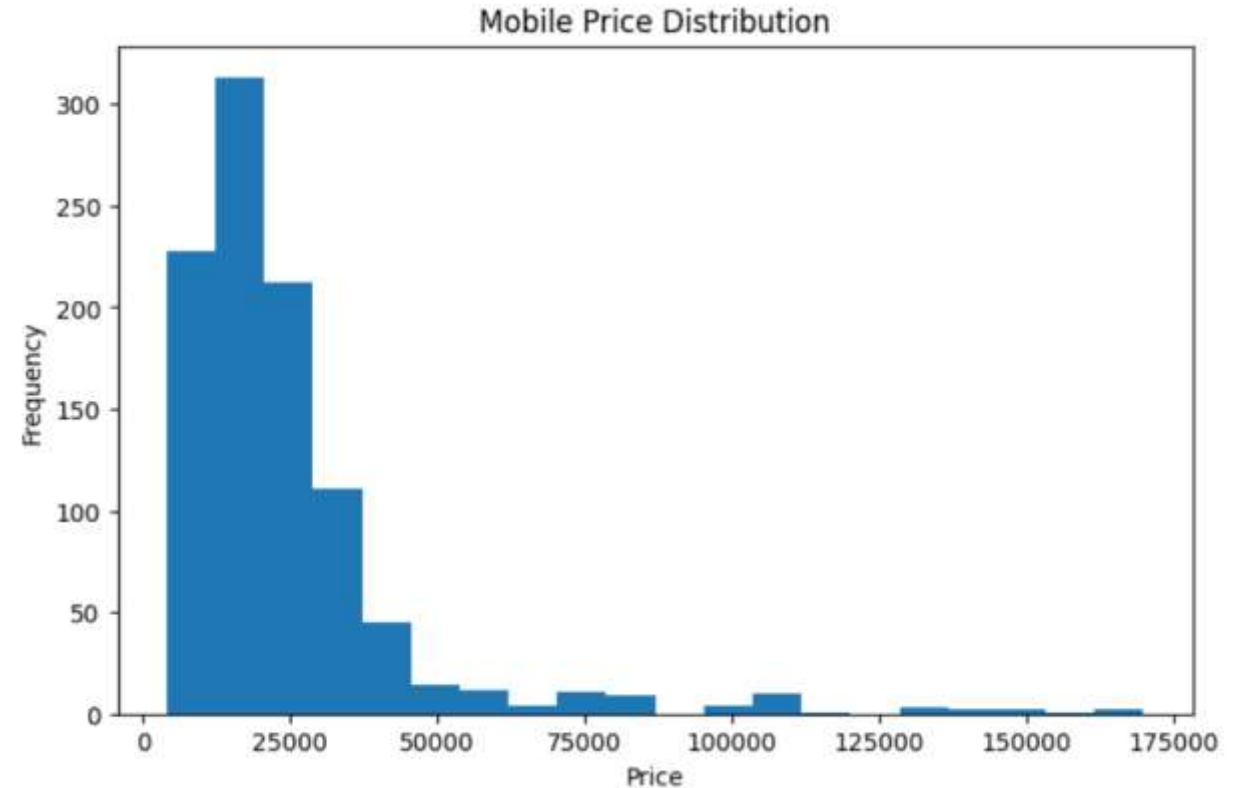
# Data Visualization:

- Realme, Samsung, Vivo, Redmi, POCO dominate listings.
- These brands focus heavily on:
- Online sales
- Multiple variants (RAM/Storage)
- Shows strong competition in the budget & mid-range segment.



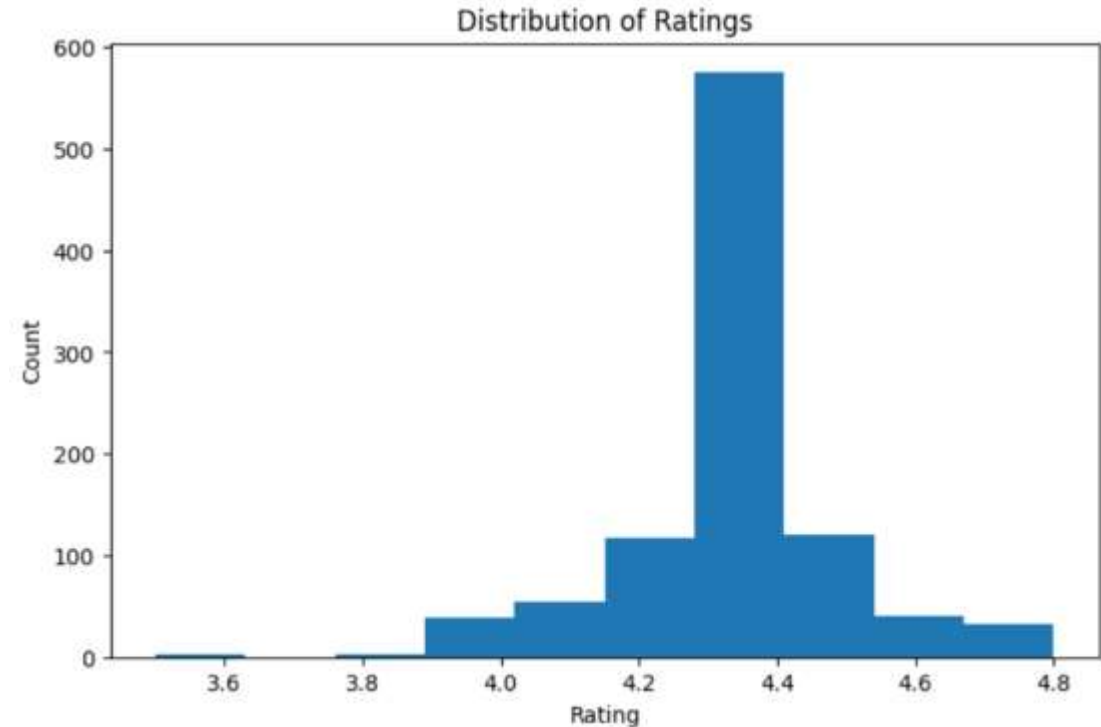
# Data Visualization:

- Most mobiles are priced between ₹5,000 – ₹30,000.
- The distribution is right-skewed, indicating:
- Flipkart has many budget & mid-range phones
- Few premium flagships above ₹80,000
- This suggests Flipkart targets mass-market consumers more than luxury buyers..



# Data Visualization:

- Majority of ratings fall between 4.1 – 4.5.
- Very few phones have ratings below 3.8.
- Indicates:
- Listed mobiles generally have good customer satisfaction
- Poorly rated products may be filtered out or discontinued



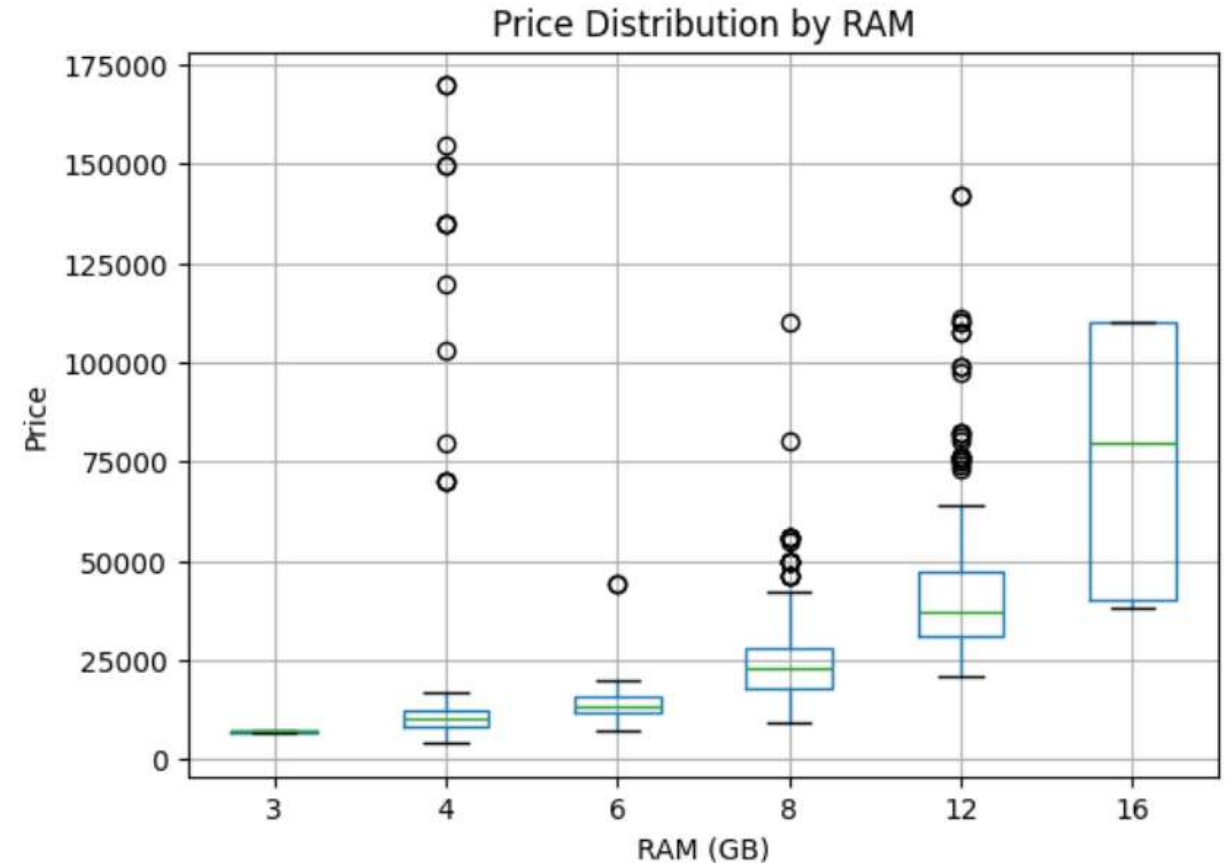
# Data Visualization:

- No strong linear relationship between price and rating.
- Many low-priced phones have high ratings (~4.4).
- Conclusion:
- Expensive phones  $\neq$  better user satisfaction
- Value-for-money devices perform very well



# Data Visualization:

- Higher RAM → Higher price
- 4GB–6GB phones dominate the budget segment
- 12GB–16GB RAM phones show:
- Wider price range
- Presence of premium and gaming phones



# Data Visualization:

- Strong positive correlation
- Selling price is consistently lower than original price
- Indicates:
- Heavy use of discount strategies
- Discounts increase with higher original prices



# Data Visualization:

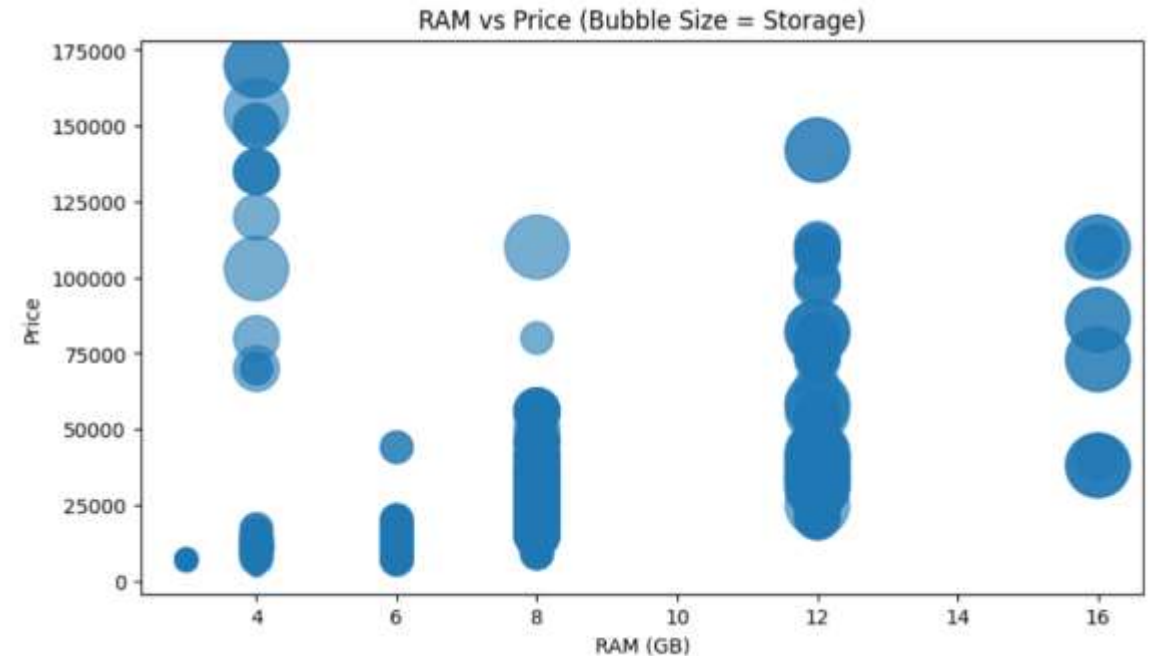
- Ratings are consistently high across all brands. Most phones, regardless of brand, have ratings between 4.2 & 4.6.
- Higher price does not guarantee a higher rating.  
Expensive phones and budget phones often receive similar ratings.
- Premium phones are few but well-rated.  
High-priced phones (₹80,000+) have good ratings, but they are limited in number.





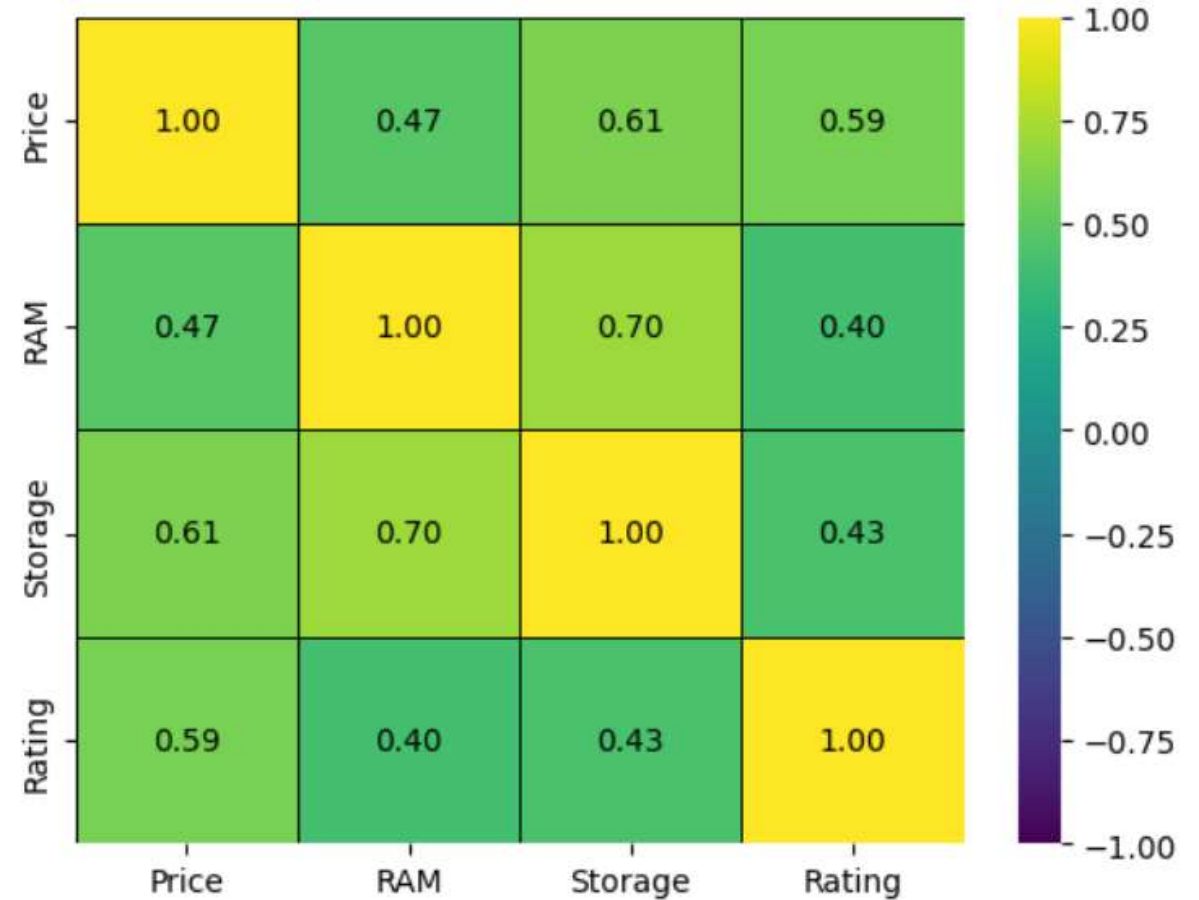
# Data Visualization:

- Phones with same RAM but higher storage cost more.
- Example:
- 8GB RAM with 256GB storage priced higher than 8GB/128GB
- Confirms:
- Storage is a key pricing factor, not just RAM
- users are willing to pay more for higher storage



# Data Visualization:

- Price increases with better hardware
- Storage affects price more than RAM
- RAM and storage usually come together
- Higher price does not guarantee higher rating
- Customer satisfaction depends on value, not just specs
- Prices are driven by RAM and storage, but ratings depend more on value-for-money than on hardware alone.



# Conclusion:

1. The Flipkart mobile market is dominated by budget and mid-range smartphones, with most devices priced below ₹30,000.
2. Customer ratings remain consistently high across all price ranges, showing that higher prices do not necessarily guarantee better user satisfaction.
3. Premium smartphones are fewer in number, have good ratings, but attract lower customer engagement due to higher prices.
4. Brands compete heavily in the same price–rating space, indicating strong competition and feature parity in the market.
5. Overall, customer satisfaction depends more on overall experience and pricing fairness than on high-end hardware specifications.

# Experiences & Challenges:

- Performing univariate, bivariate, and multivariate analysis provided a clear understanding of market trends, pricing behavior, and customer preferences.
- Data cleaning was a major challenge, especially converting percentage-based and string values into numeric format for analysis.
- Visual analysis helped easily identify relationships and patterns, but selecting the right plot for each variable required careful consideration.
- Handling skewed data and outliers in price and engagement metrics was challenging but essential for accurate interpretation.
- Combining multiple variables revealed deeper insights, though it increased complexity and required strong analytical reasoning.

# THANK YOU

