



INNOVATION. AUTOMATION. ANALYTICS

## PROJECT ON

# ISHOPSMART - Analysing the Best iPhones on Flipkart

Team Member(s)

M. Pramodha Varshini

K. Praveen



# About us

- M. Pramodha Varshini [ B.Tech ]
- K. Praveen [B.Tech]

# Agenda

- **Objective of the Project**
- **Data Collection**
- **Raw Data**
- **Data Cleaning**
- **Data Visualization**
- **Uni-Variate Analysis**
- **Bi-Variate Analysis**
- **Conclusion**

**Objective:** The objective of this project is to analyse iPhone listings on Flipkart, leveraging data scraped from the platform.

**URL :** [https://www.flipkart.com/search?sid=tyy%2C4io&otracker=CLP\\_Filters&p%5B%5D=facets.brand%255B%255D%3DAPPLE](https://www.flipkart.com/search?sid=tyy%2C4io&otracker=CLP_Filters&p%5B%5D=facets.brand%255B%255D%3DAPPLE)



### APPLE iPhone 11 (Black, 128 GB)

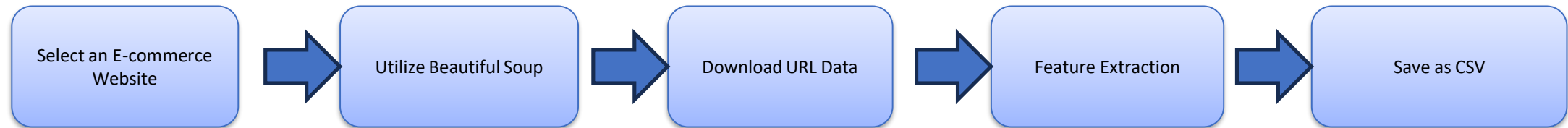
4.6 ★

2,00,475 Ratings & 11,412 Reviews

- 128 GB ROM
- 15.49 cm (6.1 inch) Liquid Retina HD Display
- 12MP + 12MP | 12MP Front Camera
- A13 Bionic Chip Processor
- Brand Warranty of 1 Year

# Data Collection

For real time analysis data is scrapped from <https://www.flipkart.com/> website



1. **Select an E-commerce Website** : Choose a suitable e-commerce website, such as Flipkart, to explore and collect data on different iPhone models.
2. **Utilize BeautifulSoup** : Implement web scraping techniques using the BeautifulSoup library to extract information from the selected website's HTML structure.
3. **Download URL Data** : Retrieve the relevant data from the identified URLs that contain information about iPhone models, ratings, specifications, and pricing.
4. **Feature Extraction** : Extract the desired features such as model name, rating, ROM, display size, chip processor, number of ratings, number of reviews, and price from the downloaded data.
5. **Save as CSV** : Organize the extracted data into a structured format and save it to a .csv file for further analysis and reference.

# Raw Data

	Model Name	Rating	ROM	Display (cm)	Chip Processor	Number of Ratings	Number of Reviews	Price(₹)
0	APPLE iPhone 11 (Black, 128 GB)	4.6	128	15.49	A13	200360	11406.0	40999
1	APPLE iPhone 11 (White, 128 GB)	4.6	128	15.49	A13	200360	11406.0	40999
2	APPLE iPhone 13 (Midnight, 128 GB)	4.7	128	15.49	A15	255205	12592.0	56999
3	APPLE iPhone 13 (Green, 128 GB)	4.7	128	15.49	A15	255205	12592.0	56999
4	APPLE iPhone 11 (Black, 64 GB)	4.6	64	15.49	A13	200360	11406.0	37999
5	APPLE iPhone 11 (White, 64 GB)	4.6	64	15.49	A13	200360	11406.0	37999
6	APPLE iPhone 14 (Midnight, 128 GB)	4.6	128	15.49	A15	48180	1833.0	67999
7	APPLE iPhone 13 (Pink, 128 GB)	4.7	128	15.49	A15	255205	12592.0	56999
8	APPLE iPhone 13 (Starlight, 128 GB)	4.7	128	15.49	A15	255205	12592.0	56999
9	APPLE iPhone 13 (Blue, 128 GB)	4.7	128	15.49	A15	255205	12592.0	56999
10	APPLE iPhone 14 (Purple, 128 GB)	4.6	128	15.49	A15	48180	1833.0	67999
11	APPLE iPhone 14 (Starlight, 128 GB)	4.6	128	15.49	A15	48180	1833.0	67999
12	APPLE iPhone 14 (Blue, 128 GB)	4.6	128	15.49	A15	48180	1833.0	67999
13	APPLE iPhone SE 3rd Gen (Product (Red), 256 GB)	4.4	256	11.94	A15	900	NaN	46599
14	APPLE iPhone 13 ((PRODUCT)RED, 128 GB)	4.7	128	15.49	A15	255205	12592.0	56999
15	APPLE iPhone 14 Plus (Starlight, 128 GB)	4.6	128	17.02	A15	15369	929.0	76999

# Data Cleaning

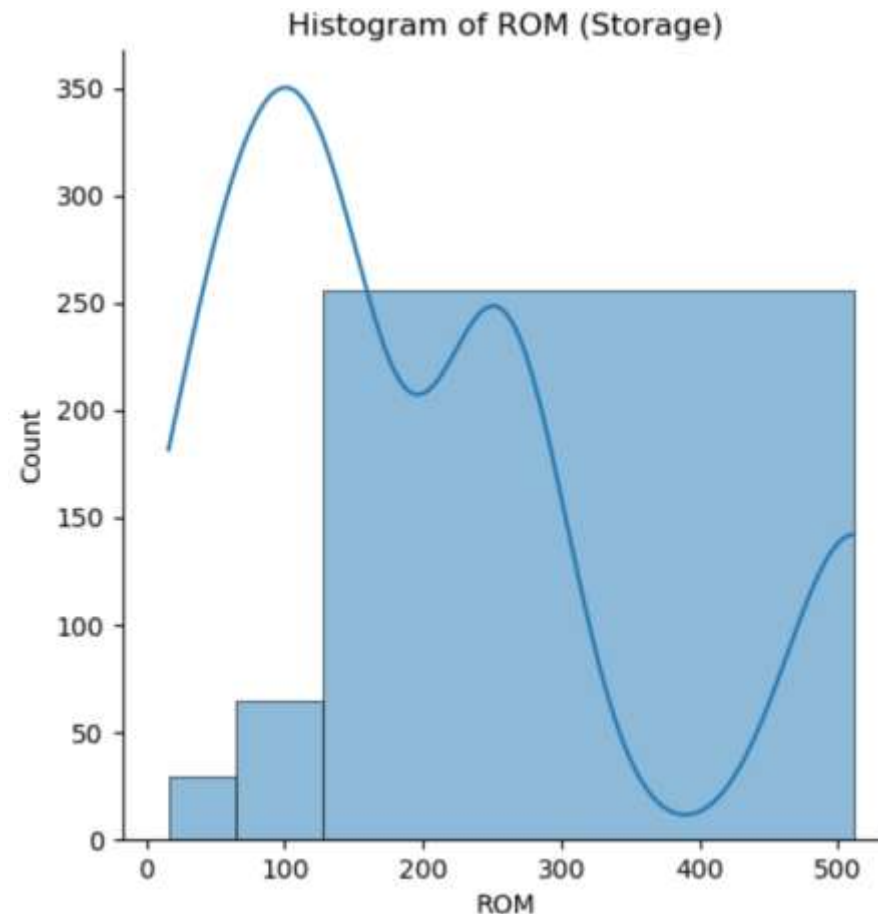
- 1. Remove Duplicates:** Eliminate duplicate records from the dataset.
- 2. Fill Null Values:** Address missing data by filling null values in the "Chip Processor," "Number of Ratings," and "Number of Reviews" columns.
- 3. Check for Invalid Values:** Identify and address any invalid or inconsistent data entries.
- 4. Remove Outliers:** Use the IQR (Interquartile Range) method to identify and remove outliers in the "Display (cm)" column.

	Model Name	Rating	ROM	Display (cm)	Chip Processor	Number of Ratings	Number of Reviews	Price(₹)
0	APPLE iPhone 11 (Black, 128 GB)	4.6	128	15.49	A13	200360.0	11406.0	40999
1	APPLE iPhone 11 (White, 128 GB)	4.6	128	15.49	A13	200360.0	11406.0	40999
2	APPLE iPhone 13 (Midnight, 128 GB)	4.7	128	15.49	A15	255205.0	12592.0	56999
3	APPLE iPhone 13 (Green, 128 GB)	4.7	128	15.49	A15	255205.0	12592.0	56999
4	APPLE iPhone 11 (Black, 64 GB)	4.6	64	15.49	A13	200360.0	11406.0	37999
...	...	...	...	...	...	...	...	...
344	APPLE iPhone 12 mini (Purple, 64 GB)	4.5	64	13.72	A14	129311.0	10267.0	50999
345	APPLE iPhone 12 Pro Max (Silver, 256 GB)	4.5	256	17.02	A14	1270.0	103.0	129900
346	APPLE iPhone 11 Pro (Space Grey, 64 GB)	4.6	64	14.73	A13	8676.0	629.0	106600
347	APPLE iPhone 12 Pro Max (Pacific Blue, 128 GB)	4.5	128	17.02	A14	1270.0	103.0	95599
348	APPLE iPhone 6s Plus (Space Grey, 16 GB)	4.4	16	13.97	A15	18836.0	2375.0	72000

# Data Visualization

## Uni-Variate Analysis

**1. ROM:** The univariate analysis of ROM revealed the distribution and central tendencies of storage capacities among the iPhone models, highlighting variations in storage options.

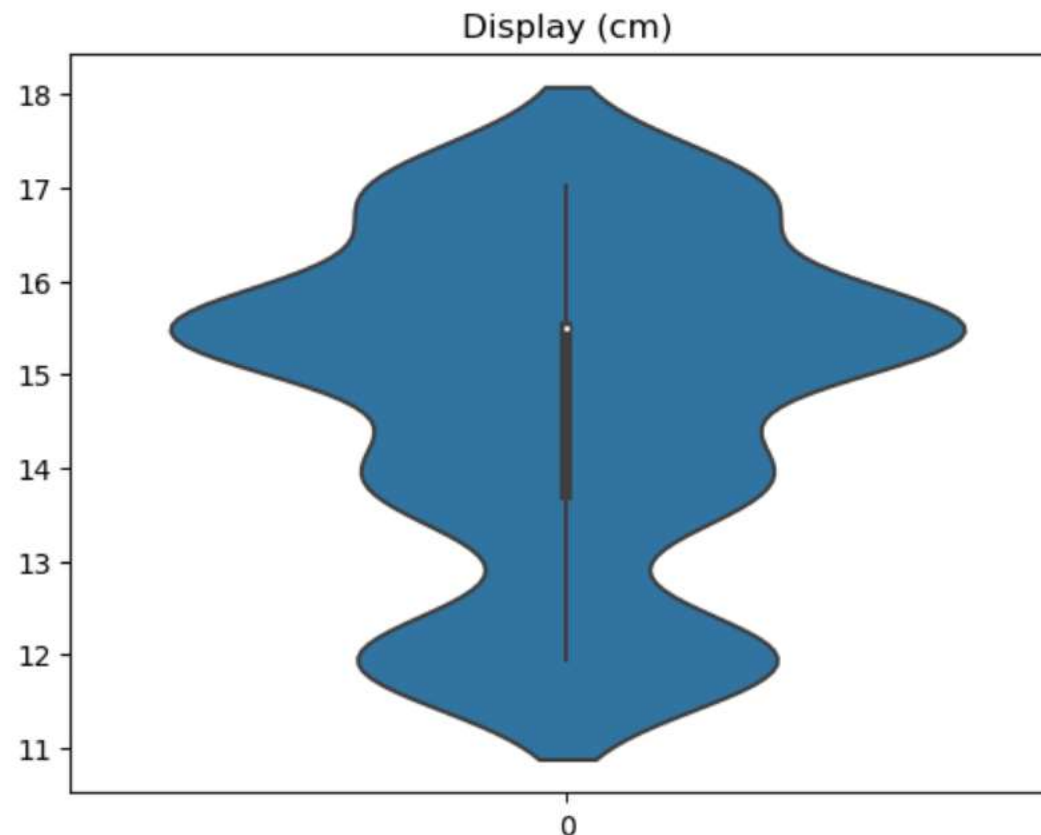




# Data Visualization

## Uni-Variate Analysis

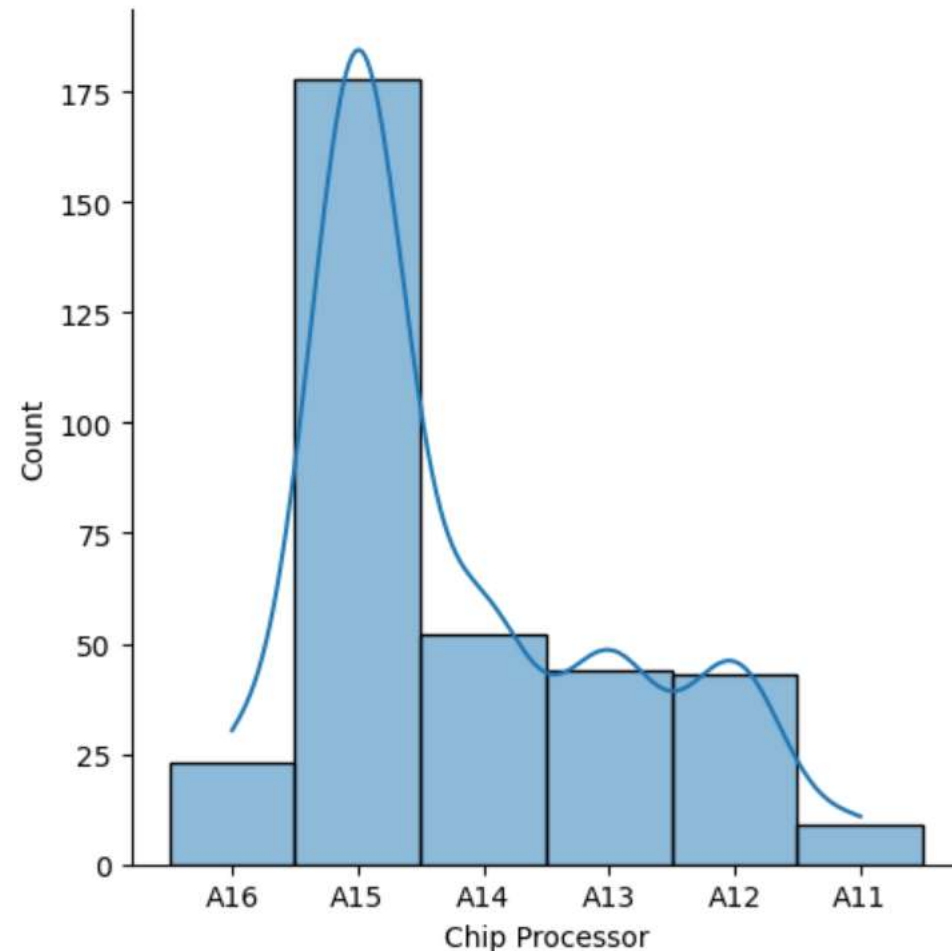
**2. Display (cm):** Univariate analysis of the "Display (cm)" variable provided insights into the distribution of display sizes, indicating the range and common screen dimensions in the iPhone models.



# Data Visualization

## Uni-Variate Analysis

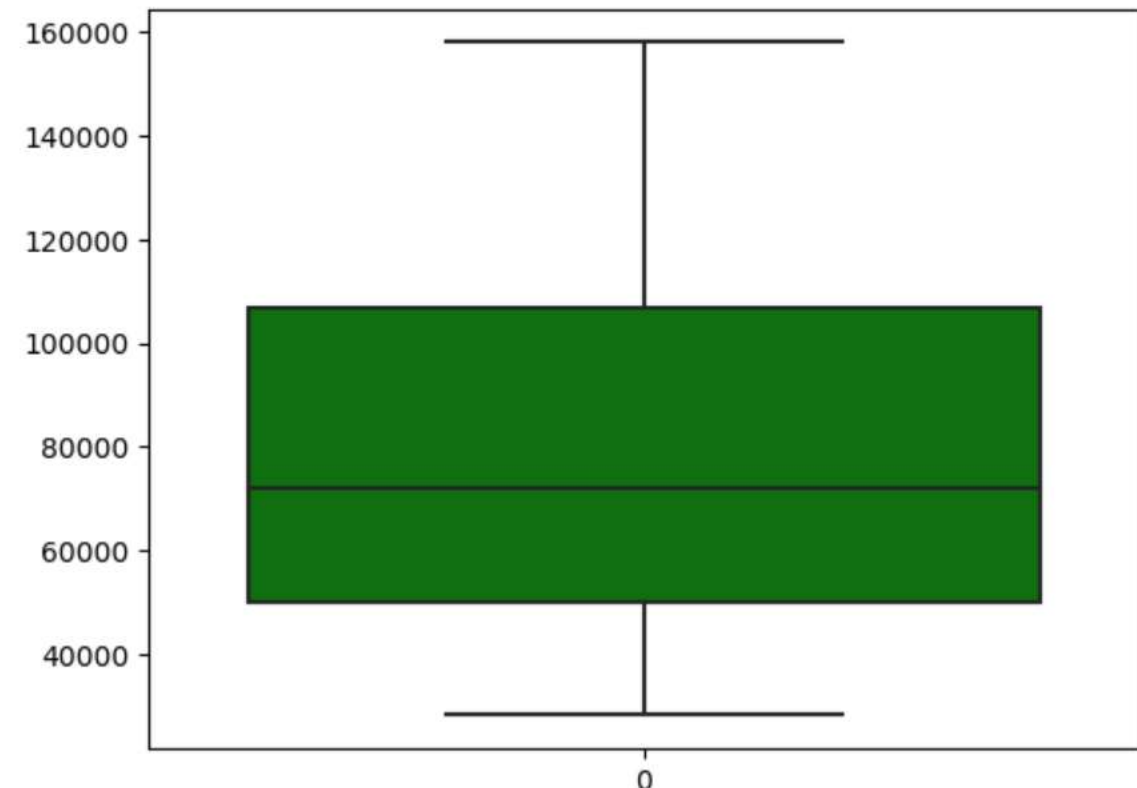
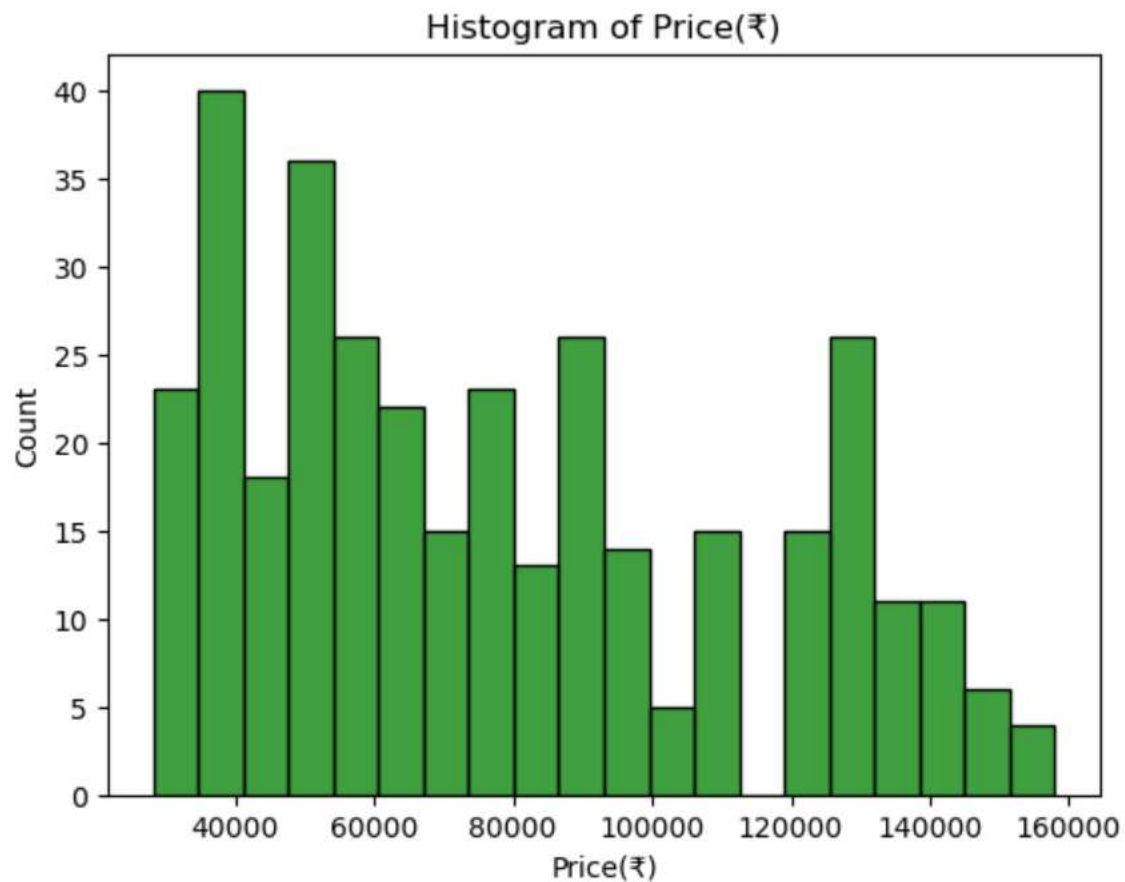
**3. Chip Processor:** In the univariate analysis of the "Chip Processor," we examined the distribution of different processor types used in iPhones, shedding light on the prevalence of specific chips in the dataset.



# Data Visualization

## Uni-Variate Analysis

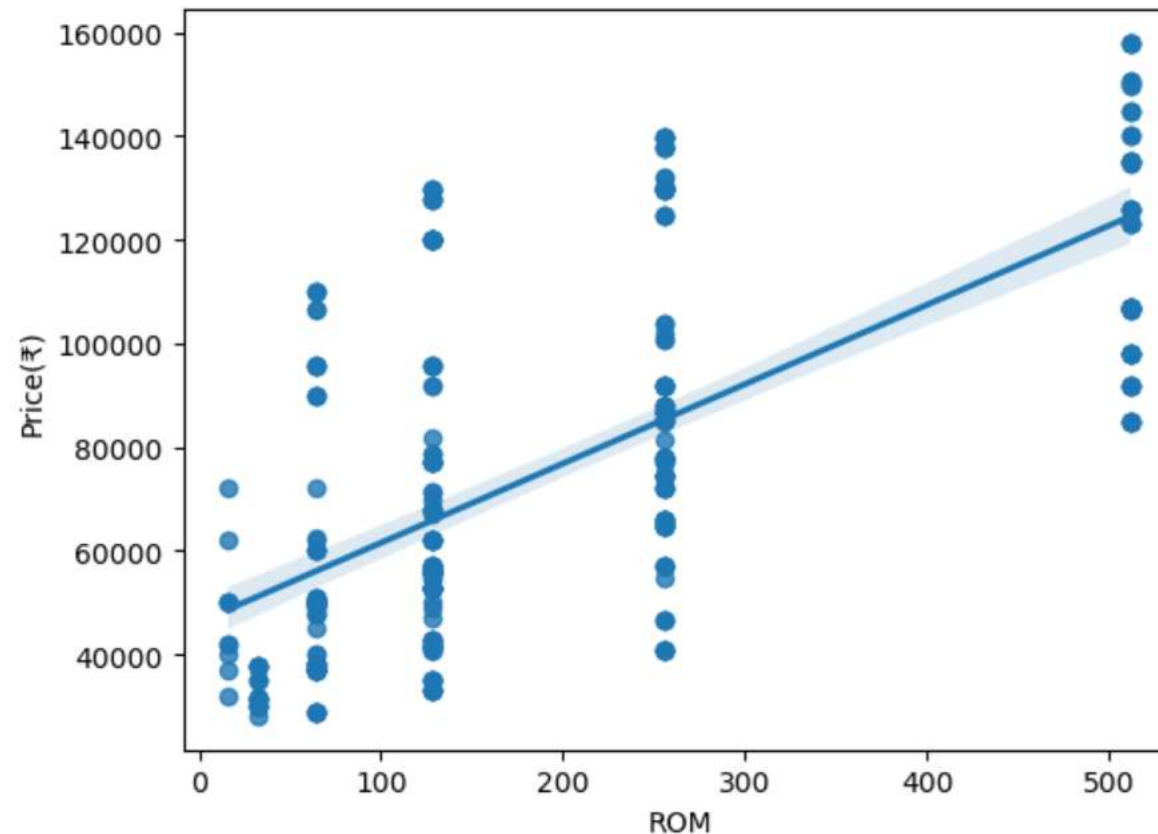
**4. Price:** Univariate analysis of the "Price" variable allowed us to explore the price distribution, central tendencies, and potential outliers among the iPhone models, aiding in understanding pricing patterns.



# Data Visualization

## Bi-Variate Analysis (Numerical vs Numerical)

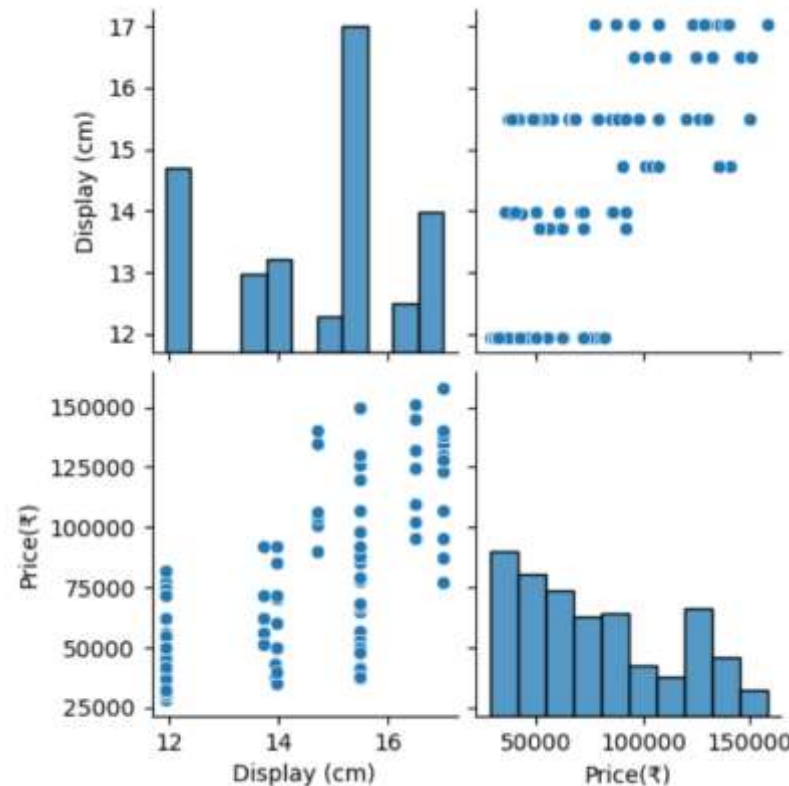
A bivariate analysis between ROM and Price revealed a positive correlation, indicating that as the ROM (storage capacity) of iPhones increased, so did their prices, suggesting that higher storage options typically come at a premium cost.



# Data Visualization

## Bi-Variate Analysis (Numerical vs Numerical)

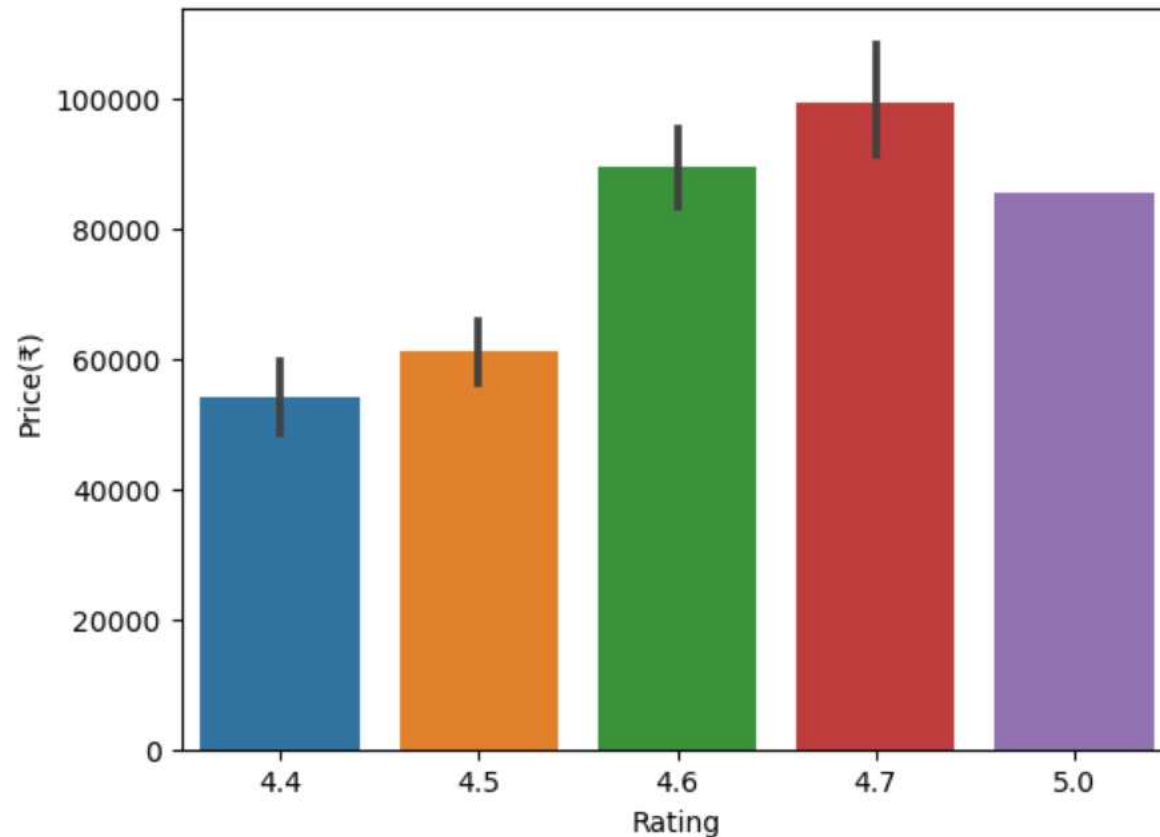
The bivariate analysis between "Display (cm)" and "Price" revealed a moderate positive correlation, indicating that, on average, iPhones with larger displays tended to have higher prices; however, this relationship also exhibited some variability, with occasional exceptions where smaller displays had higher prices, suggesting other factors may influence pricing decisions.



# Data Visualization

## Bi-Variate Analysis (Numerical vs Numerical)

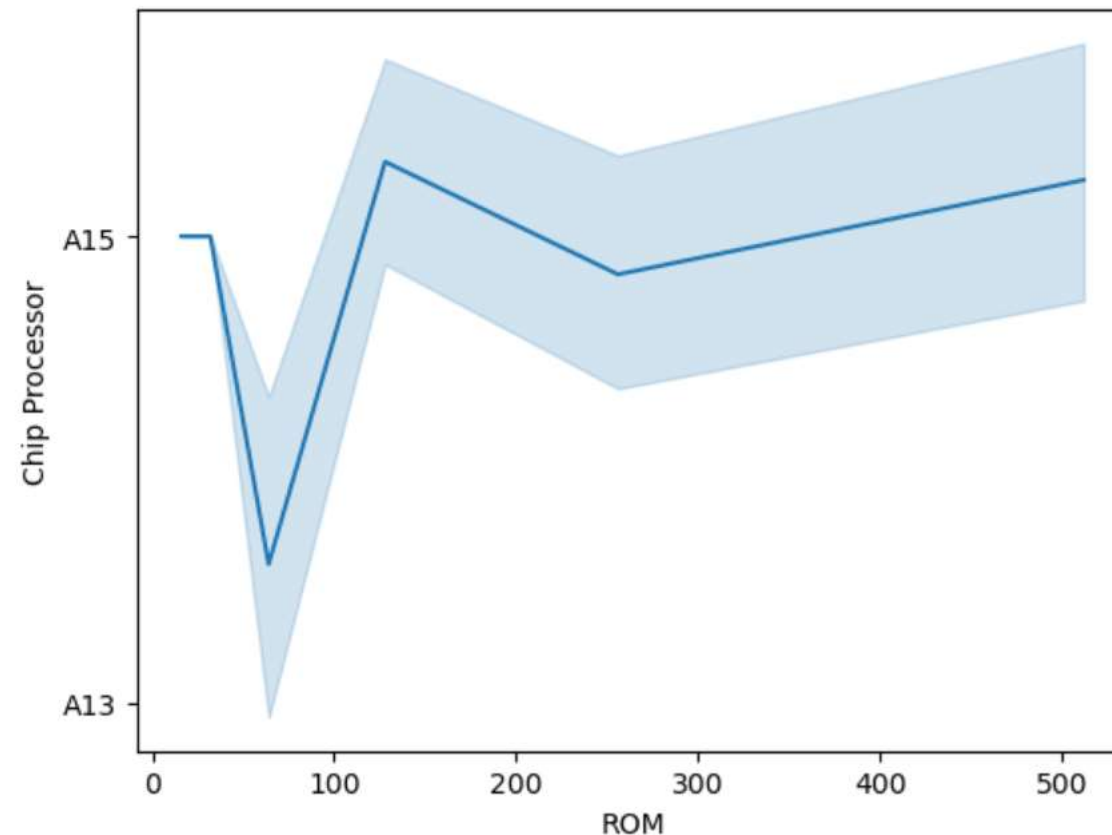
The bivariate analysis between "Rating" and "Price" revealed a nuanced relationship, indicating that higher-priced iPhones tended to have slightly higher average ratings, suggesting a potential correlation between price and customer satisfaction, but with some exceptions.



# Data Visualization

## Bi-Variate Analysis (Categorical vs Numerical)

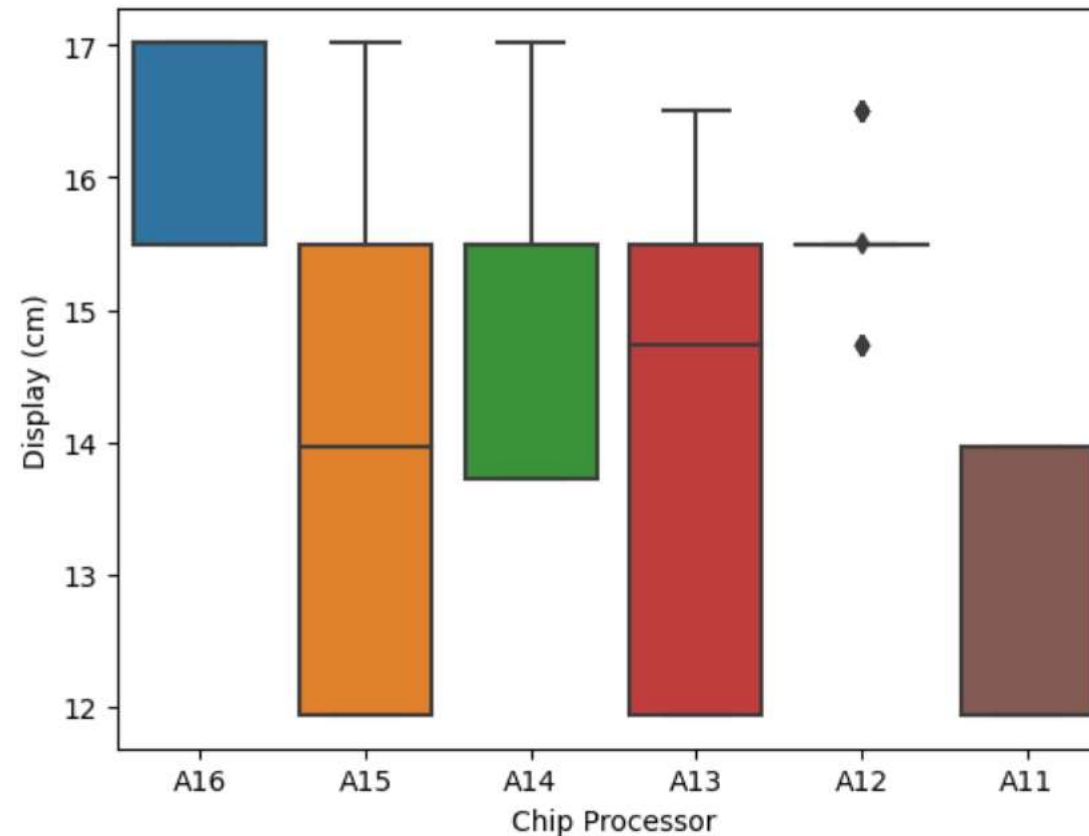
The bivariate analysis between ROM and Chip Processor revealed potential associations between storage capacity (ROM) and the type of chip processor used in iPhones, indicating that certain chip processors are more commonly paired with specific ROM capacities.



# Data Visualization

## Bi-Variate Analysis (Categorical vs Numerical)

In the bivariate analysis between "Display (cm)" and "Chip Processor," we investigated the relationship between the size of iPhone displays and the types of chip processors used, revealing potential correlations between these two key hardware features.

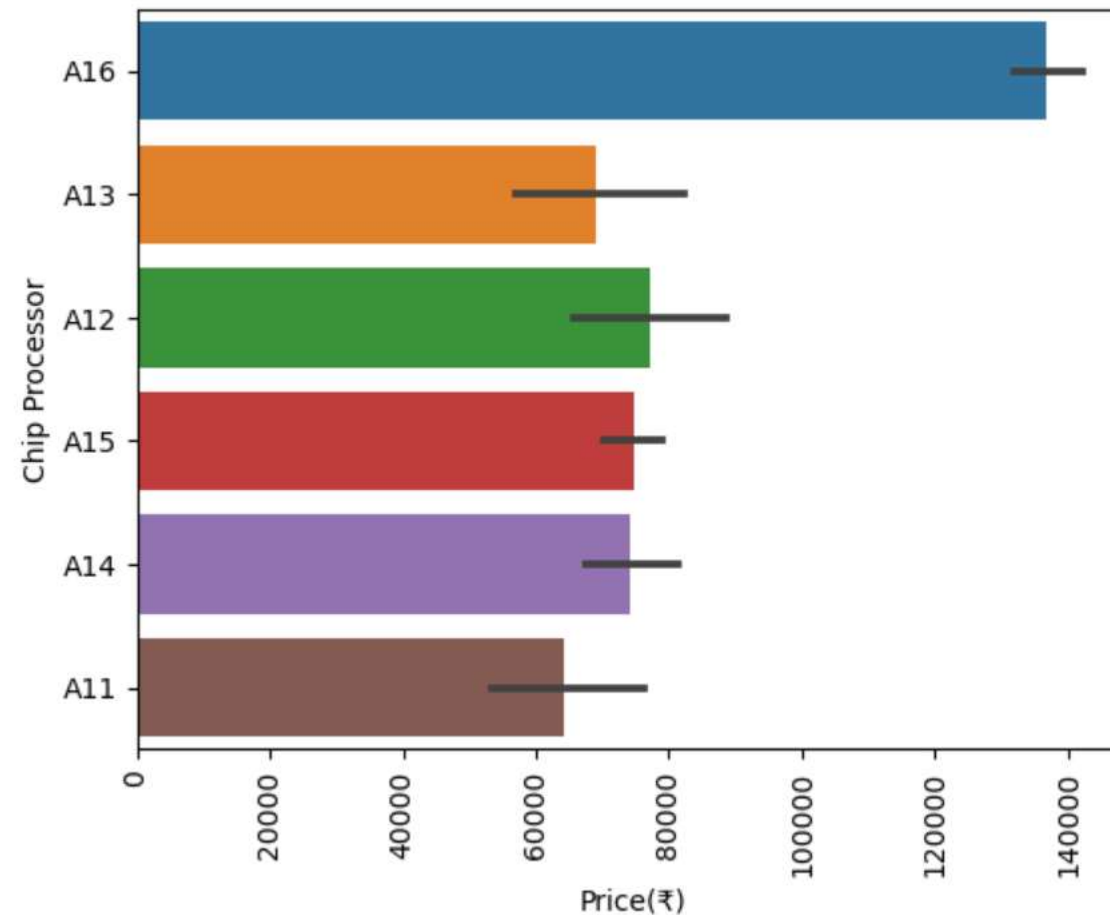




# Data Visualization

## Bi-Variate Analysis (Categorical vs Numerical)

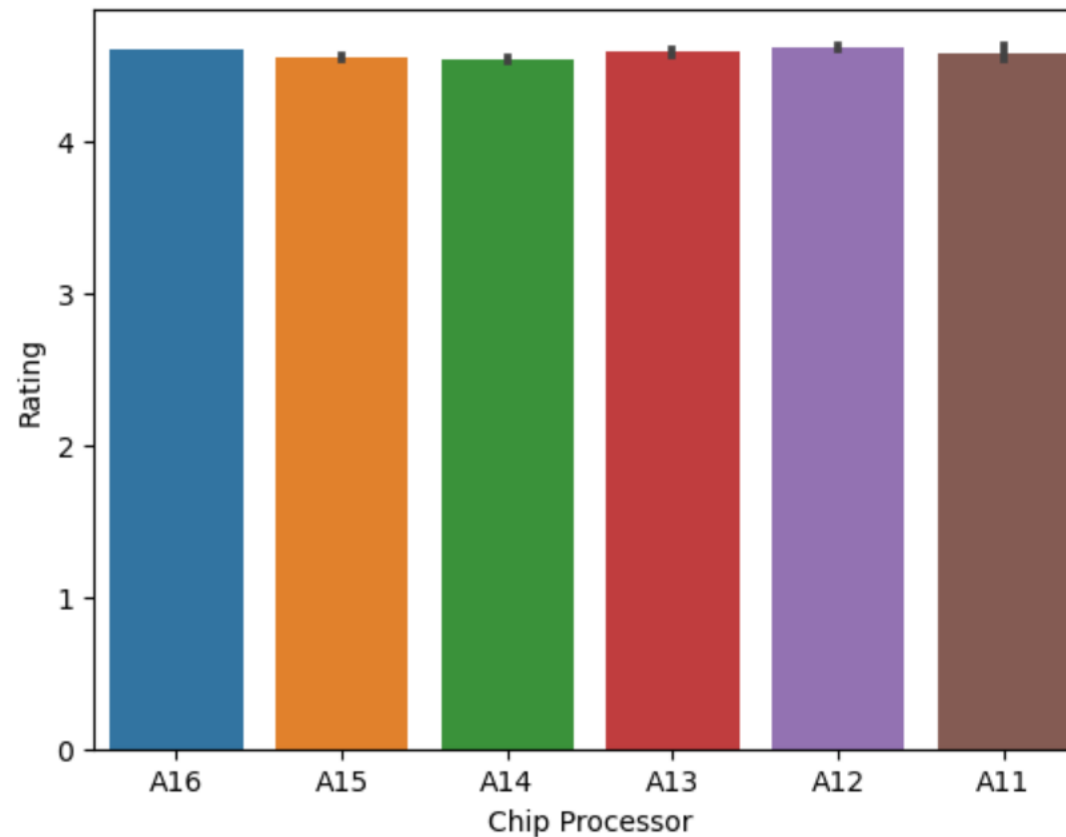
The bivariate analysis between Price and Chip Processor revealed that there is a discernible relationship between the two variables, with higher-priced iPhones typically equipped with more advanced and powerful chip processors, demonstrating a positive correlation between price and processor capabilities.



# Data Visualization

## Bi-Variate Analysis (Categorical vs Numerical)

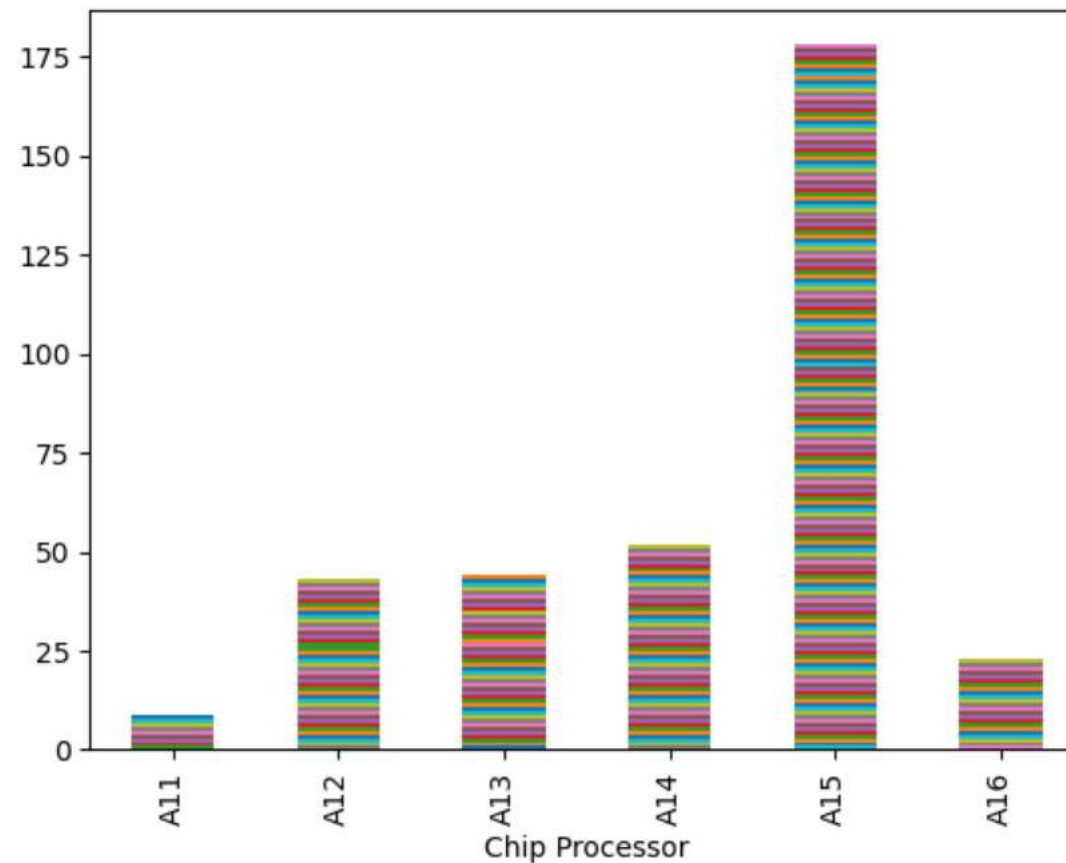
The bivariate analysis between "Rating" and "Chip Processor" revealed that specific chip processor types were associated with variations in user ratings, with some processors consistently yielding higher ratings than others, indicating a potential relationship between processor choice and customer satisfaction.



# Data Visualization

## Bi-Variate Analysis (Categorical vs Categorical )

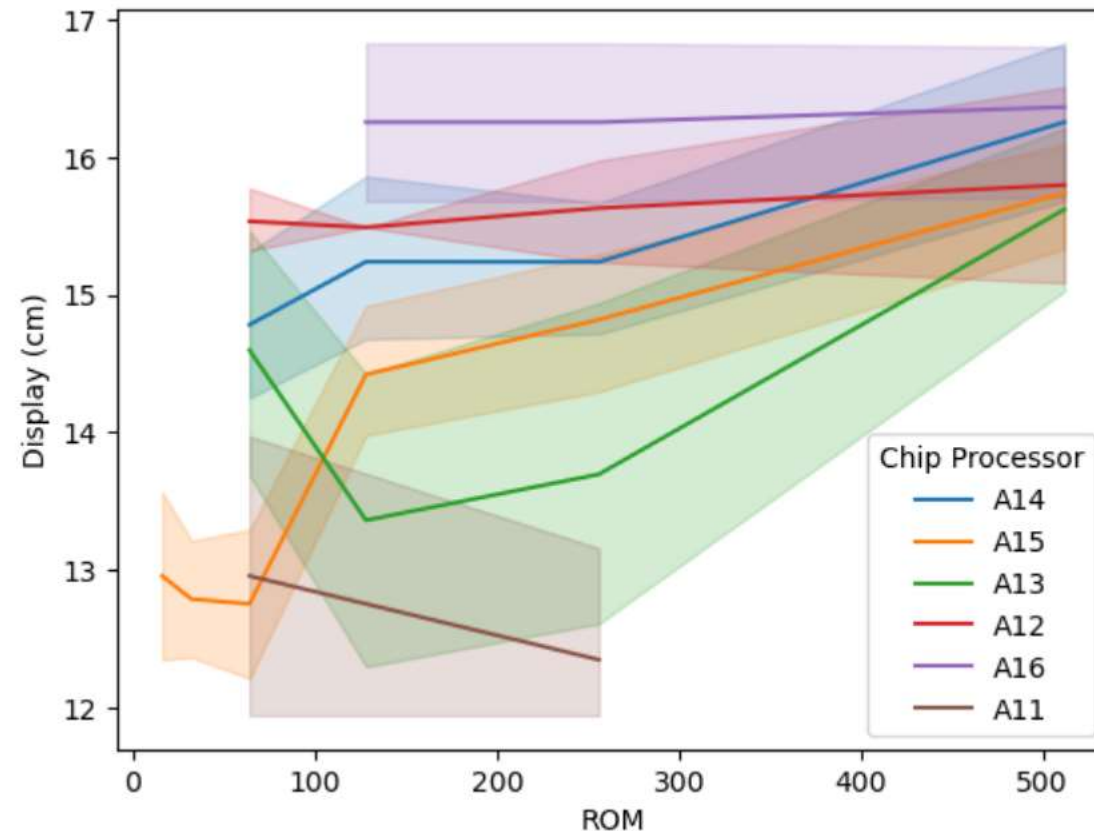
The bivariate analysis between "Model Name" and "Chip Processor" revealed the associations between specific iPhone models and the chip processors they are equipped with, helping identify patterns of chip usage across different iPhone variants.



# Data Visualization

## Multi-Variate Analysis

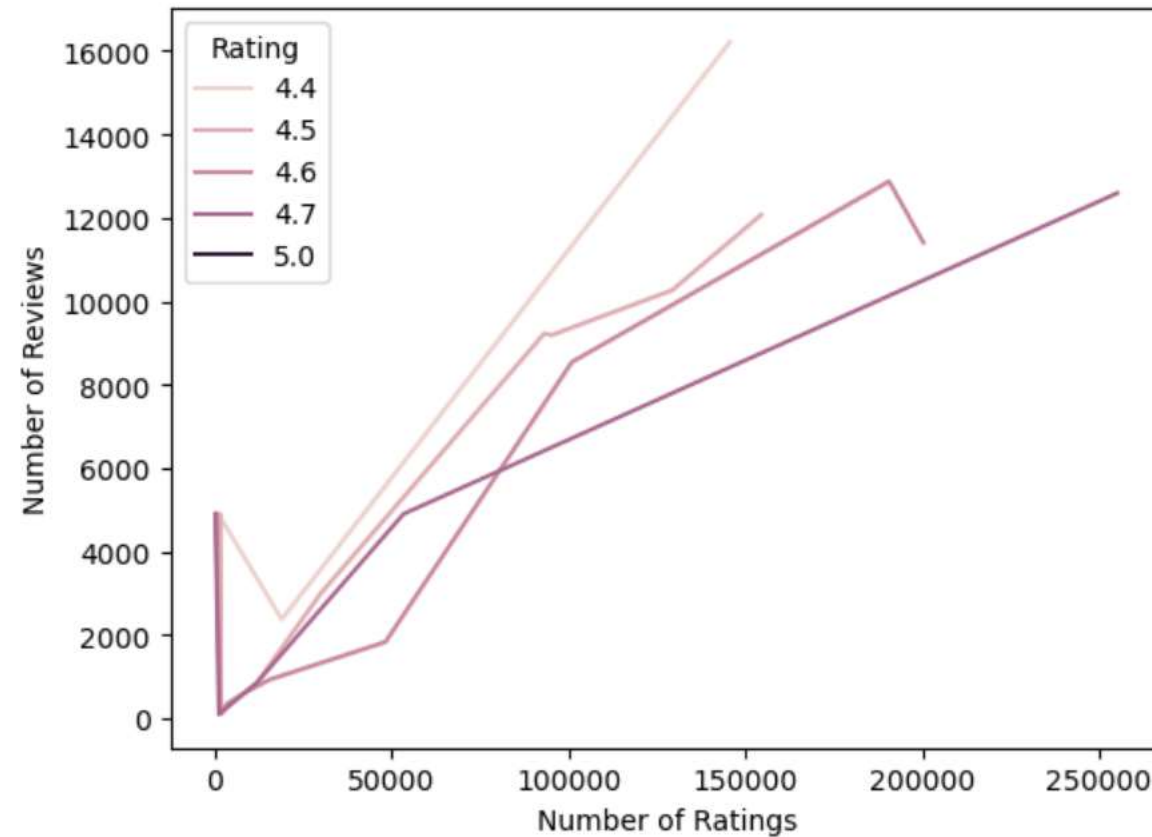
The multivariate analysis among ROM, Display (cm), and Chip Processor unveiled intricate relationships and patterns within iPhone models. It allowed us to explore how storage capacity (ROM), screen size (Display (cm)), and processor type (Chip Processor) interacted, enabling us to identify trends in feature combinations that appealed to consumers and impacted pricing and ratings.



# Data Visualization

## Multi-Variate Analysis

In the multivariate analysis among "Number of Ratings," "Number of Reviews," and "Rating," we explored the relationships and correlations between these variables to understand how customer feedback (ratings and reviews) and engagement (number of ratings and reviews) collectively contribute to the overall rating distribution, providing valuable insights into customer satisfaction and engagement trends.

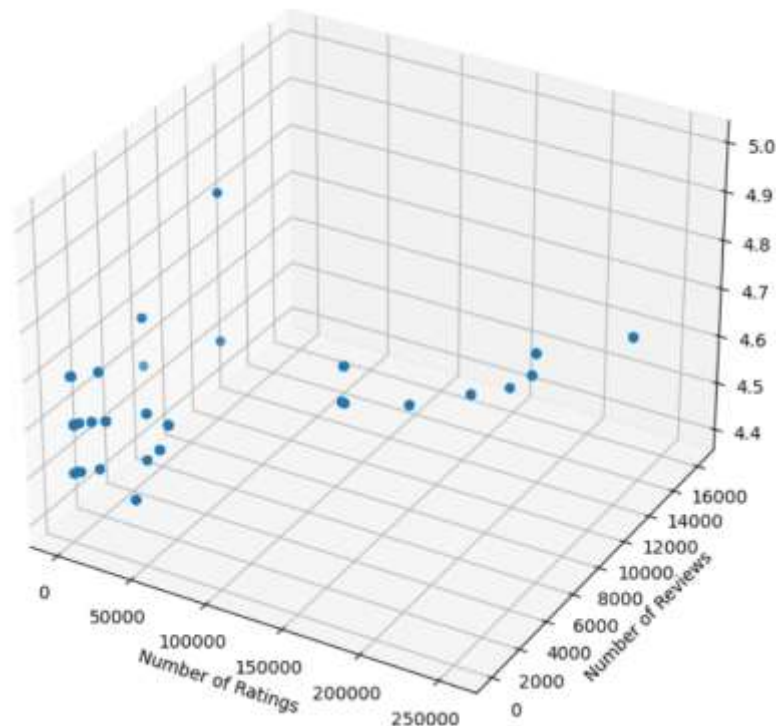


# Data Visualization

## Multi-Variate Analysis

In the multivariate analysis among "Number of Ratings," "Number of Reviews," and "Rating," we explored the relationships and correlations between these variables to understand how customer feedback (ratings and reviews) and engagement (number of ratings and reviews) collectively contribute to the overall rating distribution, providing valuable insights into customer satisfaction and engagement trends.

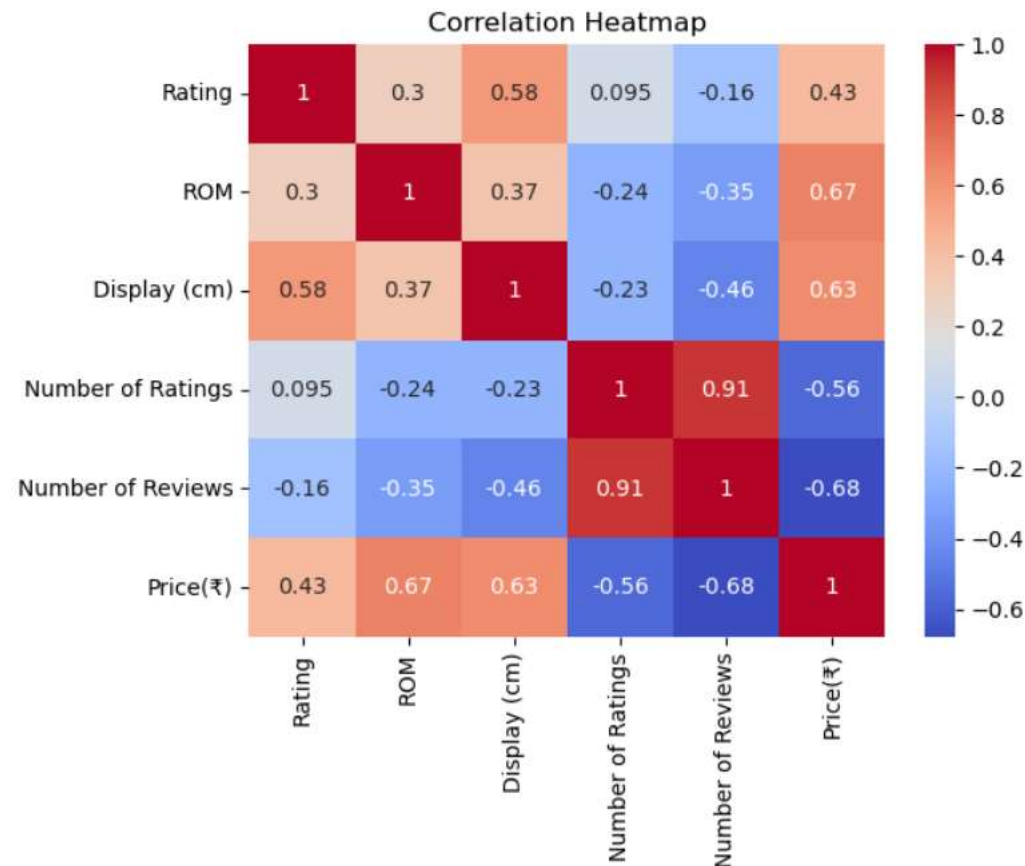
Rating in accordance with number of ratings and reviews



# Data Visualization

## CORRELATION HEATMAP

The multivariate analysis, represented by the correlation heatmap among all eight columns, revealed valuable insights into the relationships and dependencies among the iPhone model features, highlighting potential correlations between variables such as Model Name, Rating, ROM, Display (cm), Chip Processor, customer ratings, customer reviews and Price assisting in the identification of significant factors influencing consumer choices.



# RESULTS

- 1. Optimal Chip Processor - A15:** The A15 processor exhibits extensive availability across various phone models, making it a compelling choice.
- 2. Recommended ROM - 256GB:** Despite having two more options in the 128GB category, the preference leans towards 256GB of ROM (Storage). This choice anticipates the evolving technology landscape, necessitating ample space for software installations and updates.
- 3. Immersive Display - 15.49 cm:** The market offers a plethora of smartphone models with a 15.49 cm display size. This variety provides consumers with a wide range of choices to explore.
- 4. Top Ratings - 4.5 and Above:** Our selection criteria prioritize iPhones that meet all the aforementioned conditions while also boasting a minimum rating of 4.5 or higher.
- 5. Budget-Friendly Pricing - Starting at ₹65,999/-:** Remarkably, you can discover iPhones that meet all the specified criteria, and they start at a competitive price point of ₹65,999/-. This ensures an attractive blend of performance and affordability.

## FINAL ASSUMPTION :

These refined observations guide us towards making an informed choice when selecting the ideal iPhone model that aligns with our preferences and requirements.



# RESULT MODEL :

	Model Name
21	APPLE iPhone 13 ((PRODUCT)RED, 256 GB)
22	APPLE iPhone 13 (Midnight, 256 GB)
25	APPLE iPhone 13 (Pink, 256 GB)
28	APPLE iPhone 13 (Starlight, 256 GB)
30	APPLE iPhone 13 (Blue, 256 GB)
34	APPLE iPhone 13 (Green, 256 GB)



# CONCLUSION

In conclusion, the iPhone 13 excels in all the key features you've scraped from Flipkart. It boasts high ratings, ample storage options, a versatile display, a powerful processor, a substantial user base, and a price point that reflects its premium quality. Therefore, based on this analysis, the iPhone 13 is indeed the best choice for those looking for a top-tier smartphone on Flipkart. However, it's essential to consider your personal preferences and budget before making a final decision.



THANK  
YOU

