# 📱 Project Name: Mobile Sales Analysis

## 📂 Project Type: *Exploratory Data Analysis (EDA)*

## 👥 Contribution: *Group Project*

## 👨‍🤝‍👨 Team Members: Ashish Guragain, Avishek Guragain

---

# 📄 Project Summary

## 🔍 Overview

This project presents an end-to-end exploratory data analysis (EDA) of a mobile sales dataset consisting of **3,800+ transactions** across various Indian cities. The objective is to uncover **sales patterns**, **customer behaviors**, and **product performance trends** that can support strategic decisions in:

- 📦 Inventory Planning
- 🎯 Targeted Marketing
- 💰 Pricing Strategies

---

## 🛠️ Tools & Technologies Used

- **Programming Language:** Python
- **Libraries:** pandas, matplotlib, seaborn, numpy
- **Environment:** Jupyter Notebook
- **Data Source:** Excel-based sales dataset containing:
  - Transactional data
  - Demographic information
  - Product attributes

---

## 📊 Analysis Structure

### 🔹 1. Univariate Analysis

Focuses on understanding the **distribution** and **key metrics** of individual variables such as:

- Total revenue
- Customer age
- Product ratings
- City-wise sales

- Payment methods

### 🔹 2. Bivariate Analysis

Explores the **relationships between variables** to extract deeper business insights:

- 📈 Sales trends over time (monthly/daily)
- 💵 Price vs. Units Sold
- 👴 Customer Age vs. Spending
- 🌆 City-wise brand preferences
- 💳 Payment Method vs. Revenue patterns

---

# 🔍 Problem Statement

The goal of this project is to analyze **mobile phone sales data** to uncover **customer behavior patterns**, **product performance**, and **revenue trends** across various Indian cities.
The insights derived will help support strategic decisions in:

- 🎯 Marketing and promotional campaigns
- 📦 Inventory and stock planning
- 💰 Pricing and discount strategies

---

# 📊 Univariate Analysis Questions

## 🛍️ Sales & Revenue

- What is the **total revenue** generated?
- What is the **total number of units** sold?
- What is the **monthly revenue trend**?
- Which **brands have the highest total revenue**?
- Which **mobile models sold the most units**?
- What is the **distribution of mobile phone prices**?

## 👤 Customer Demographics

- What is the **age distribution** of customers?
- What is the **distribution of customer ratings**?
- What are the **most common payment methods**?

## 🌆 Geographic Distribution

- Which **cities have the highest number of sales**?
- Which **cities contribute the most to total revenue**?

# 🔗 Bivariate Analysis Questions

## 🗓️ Time vs Sales

- How does **revenue change over months**?
- Are **weekend sales higher than weekday sales**?
- Which **days of the week** have the most transactions?

## 📦 Product vs Sales

- Is there a relationship between **price and units sold**?
- Which **brands receive the highest average ratings**?

## 👫 Customer vs Product

- Which **age people buys the most mobile phones**?
- **Which age customer spend more?**?

## 🌆 City vs Product

- Which **brands/models are most popular in specific cities**?
- Do certain cities prefer **budget vs premium phones**?

## 💳 Payment Method vs Sales

- Which **payment method generates the most revenue**?
- Is there a preferred payment method for **higher-priced phones**?

---


## Importing the necessary libraries

## Load Mobile Sales Dataset

| | Transaction ID | Day | Month | Year | Day Name | Brand | Units Sold | Price Per Unit | Customer Name | Customer Age | City |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 1 | 9 | 10 | 2021 | Sat | Xiaomi | 6 | 10174.70 | Lalita Ahuja | 38 | Ludhian |
| **1** | 2 | 9 | 10 | 2021 | Saturday | Vivo | 6 | 10565.19 | Sneha Sharma | 37 | Delh |
| **2** | 3 | 9 | 10 | 2021 | Saturday | Vivo | 8 | 58527.58 | Radha Srivastava | 40 | Mumba |
| **3** | 4 | 10 | 10 | 2021 | Sunday | Xiaomi | 5 | 25563.98 | Bhavana Arora | 21 | Mumba |
| **4** | 5 | 10 | 10 | 2021 | Sunday | OnePlus | 3 | 48168.02 | Sneha Mehta | 38 | Gorakhpu |

# Data Exploration and Data Cleaning

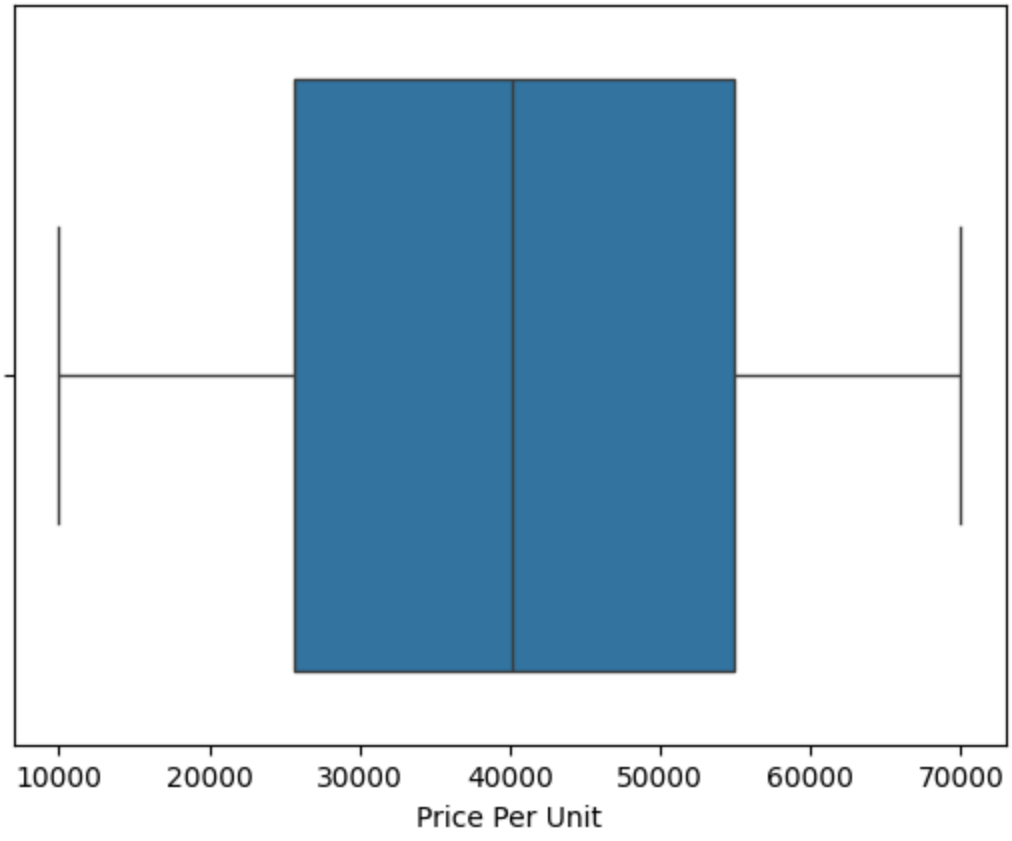
```
Transaction ID      0
Day                 0
Month               0
Year                0
Day Name            0
Brand               0
Units Sold          0
Price Per Unit      0
Customer Name       0
Customer Age        0
City                0
Payment Method      0
Customer Ratings    0
Mobile Model        0
dtype: int64
```

```
  Day Month  Year       Date
0   9    10  2021 2021-10-09
1   9    10  2021 2021-10-09
2   9    10  2021 2021-10-09
3  10    10  2021 2021-10-10
4  10    10  2021 2021-10-10
```

# Describing the data

| | Transaction ID | Units Sold | Price Per Unit | Customer Age | Customer Ratings | Date | Total_rev |
|---|---|---|---|---|---|---|---|
| **count** | 3835.000000 | 3835.000000 | 3835.000000 | 3835.000000 | 3835.000000 | 3835 | 3835.00 |
| **mean** | 1918.000000 | 4.993481 | 40114.036816 | 38.098305 | 3.693090 | 2023-04-07 13:31:25.893090048 | 200574.96 |
| **min** | 1.000000 | 1.000000 | 10011.110000 | 18.000000 | 1.000000 | 2021-10-09 00:00:00 | 10115.91 |
| **25%** | 959.500000 | 3.000000 | 25617.405000 | 27.500000 | 3.000000 | 2022-07-10 00:00:00 | 83930.94 |
| **50%** | 1918.000000 | 5.000000 | 40136.490000 | 38.000000 | 4.000000 | 2023-04-07 00:00:00 | 163855.84 |
| **75%** | 2876.500000 | 7.000000 | 54973.475000 | 49.000000 | 5.000000 | 2024-01-05 00:00:00 | 289246.93 |
| **max** | 3835.000000 | 9.000000 | 69984.480000 | 59.000000 | 5.000000 | 2024-10-08 00:00:00 | 628391.79 |
| **std** | 1107.213469 | 2.581122 | 17310.488427 | 12.110622 | 1.325069 | NaN | 142402.09 |

**Note** - Price per unit column is very important so we have to find big outliers in important column




**Note** - No big outliers detected.

# Univariate Analysis

# Sales & Revenue

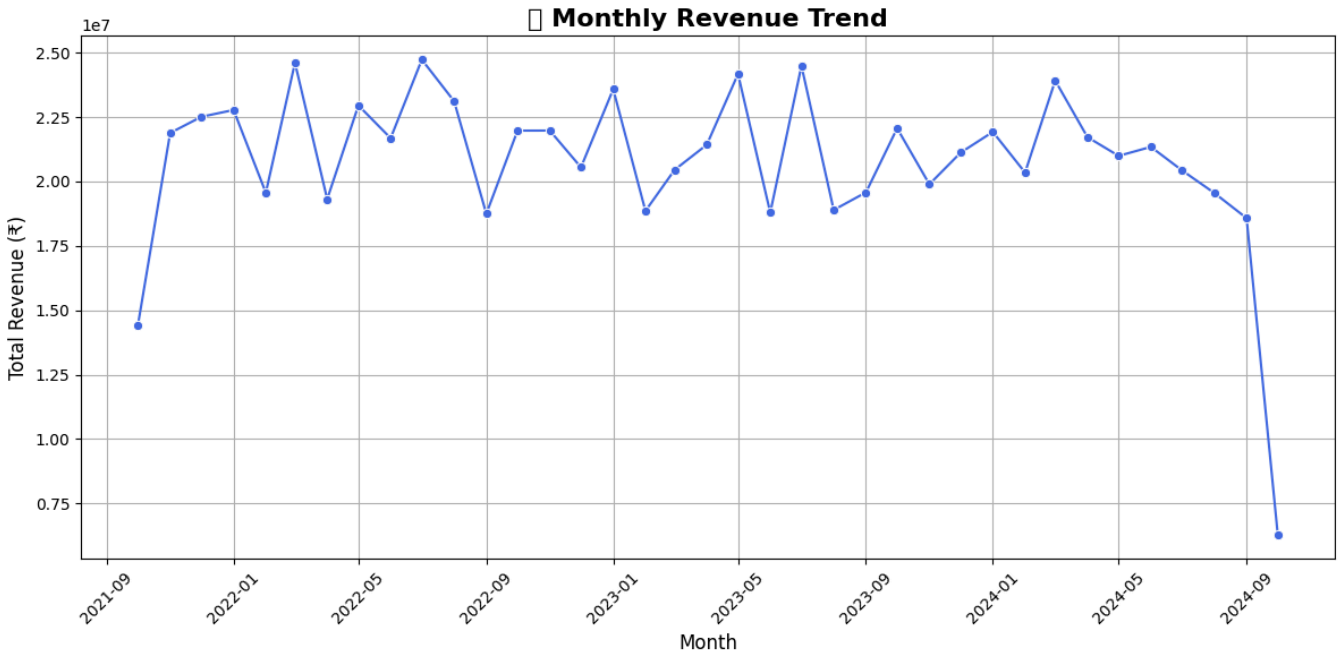
### 1.What is the total revenue generated?

```
np.float64(769204987.97)
```

### 2.What is the total number of units sold?

```
np.int64(19150)
```

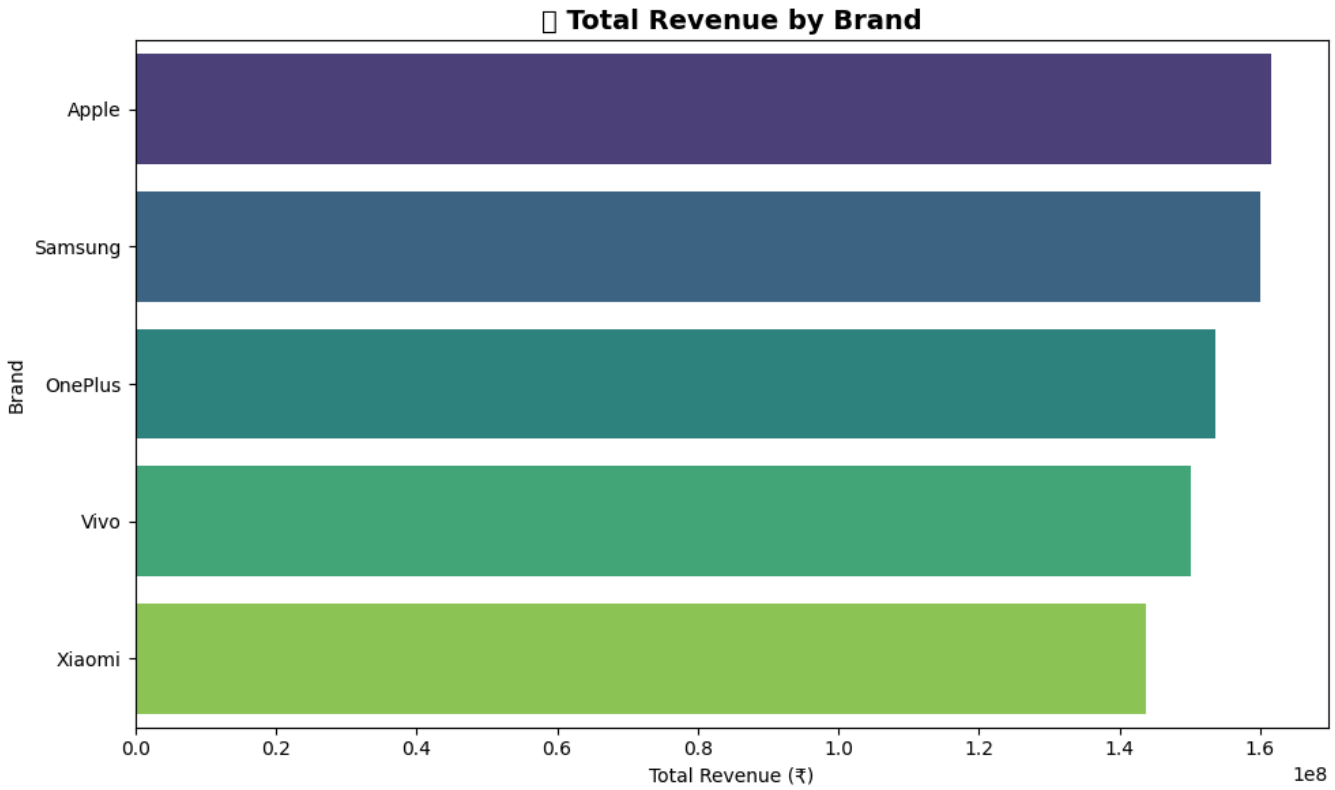
### 3.What is the monthly revenue trend?



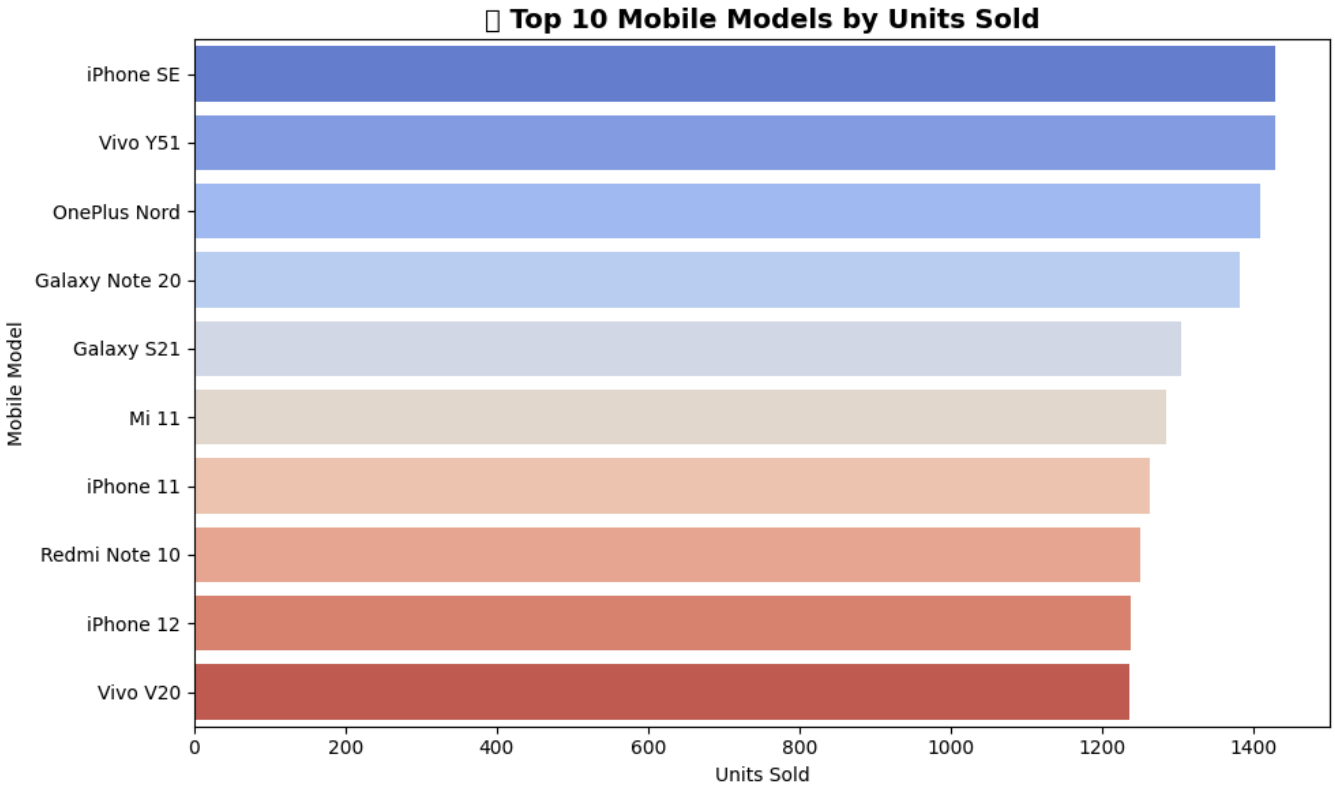

### 4.Which brand have the highest total revenue?

```
'Apple'
```

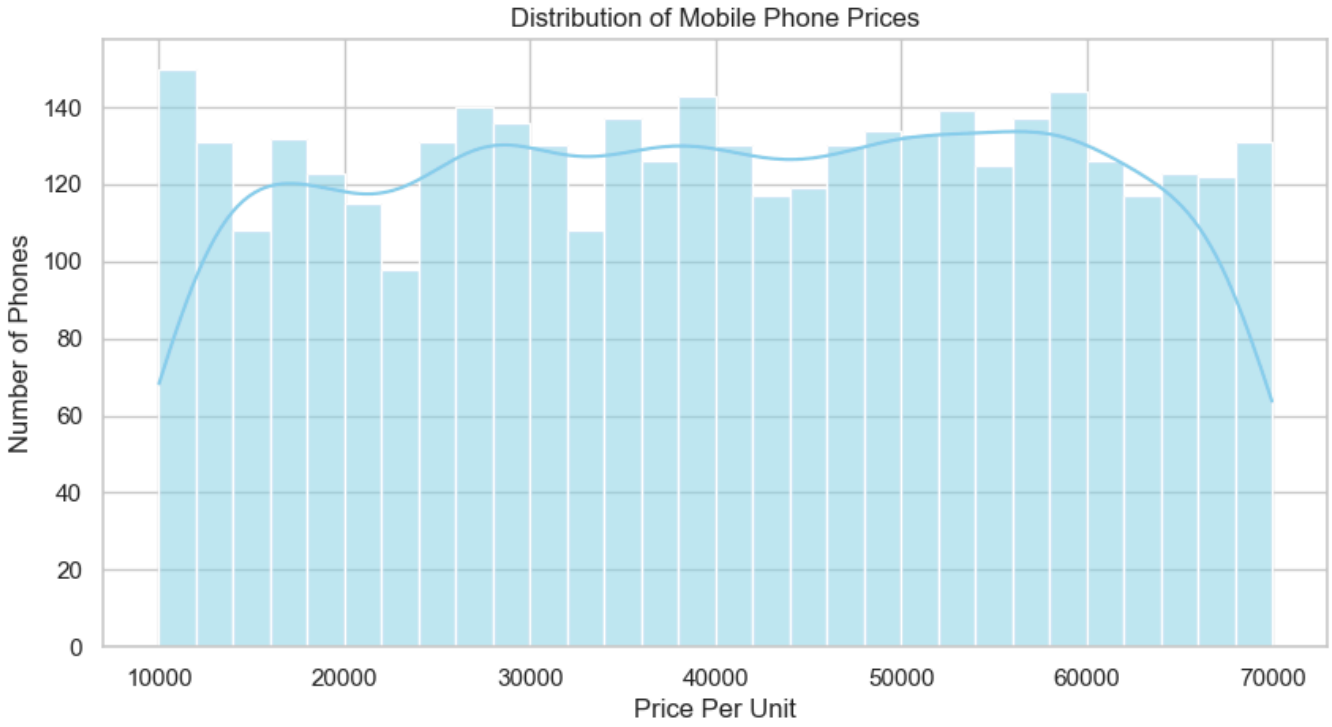
### 5. Distribution of revenue across the brands




## 6.Distribution of mobile models across sold units

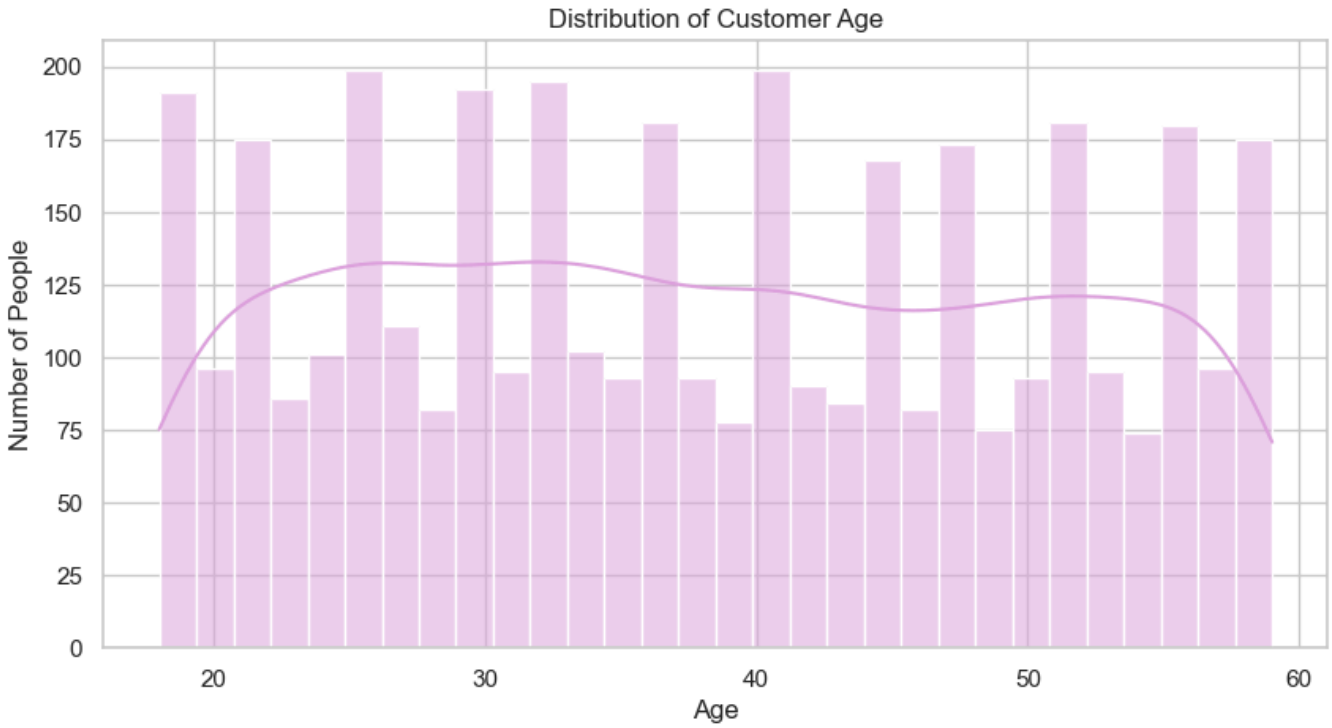



## 7.What is the distribution of mobile phone prices?

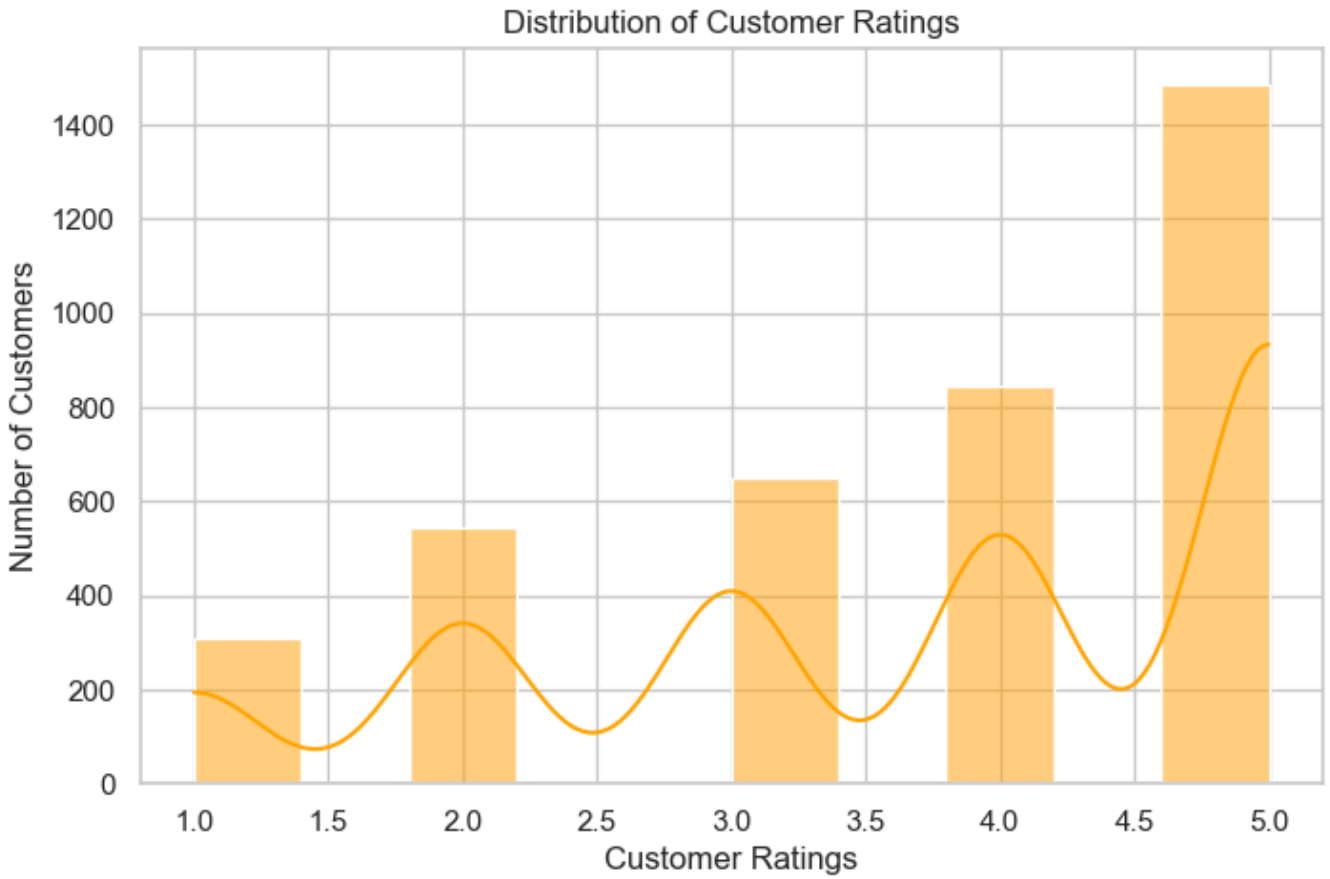



# Customer Demographics

**8.What is the age distribution of customers?**

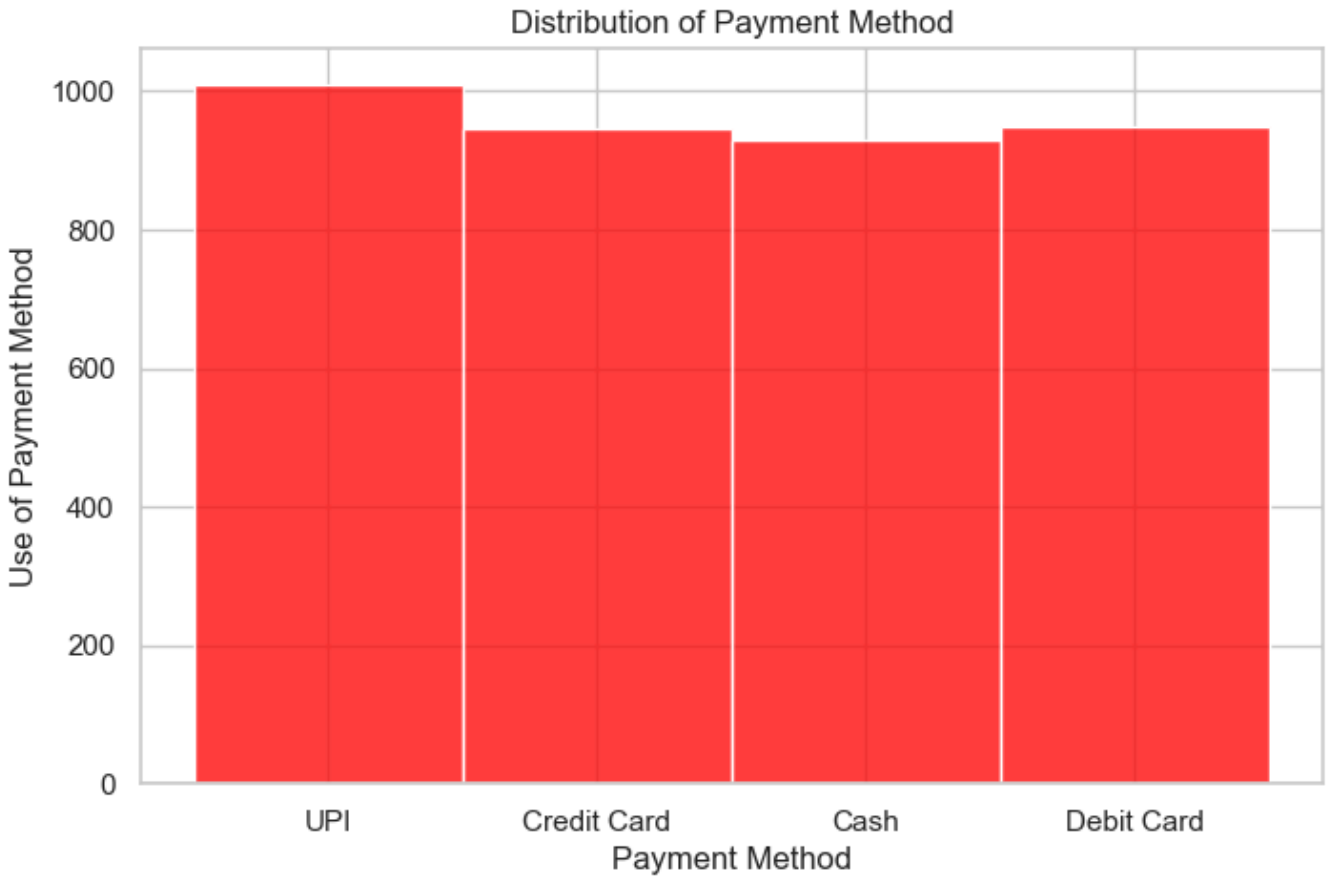



**9.What is the distribution of customer ratings??**




**10.What are the most common payment methods?**

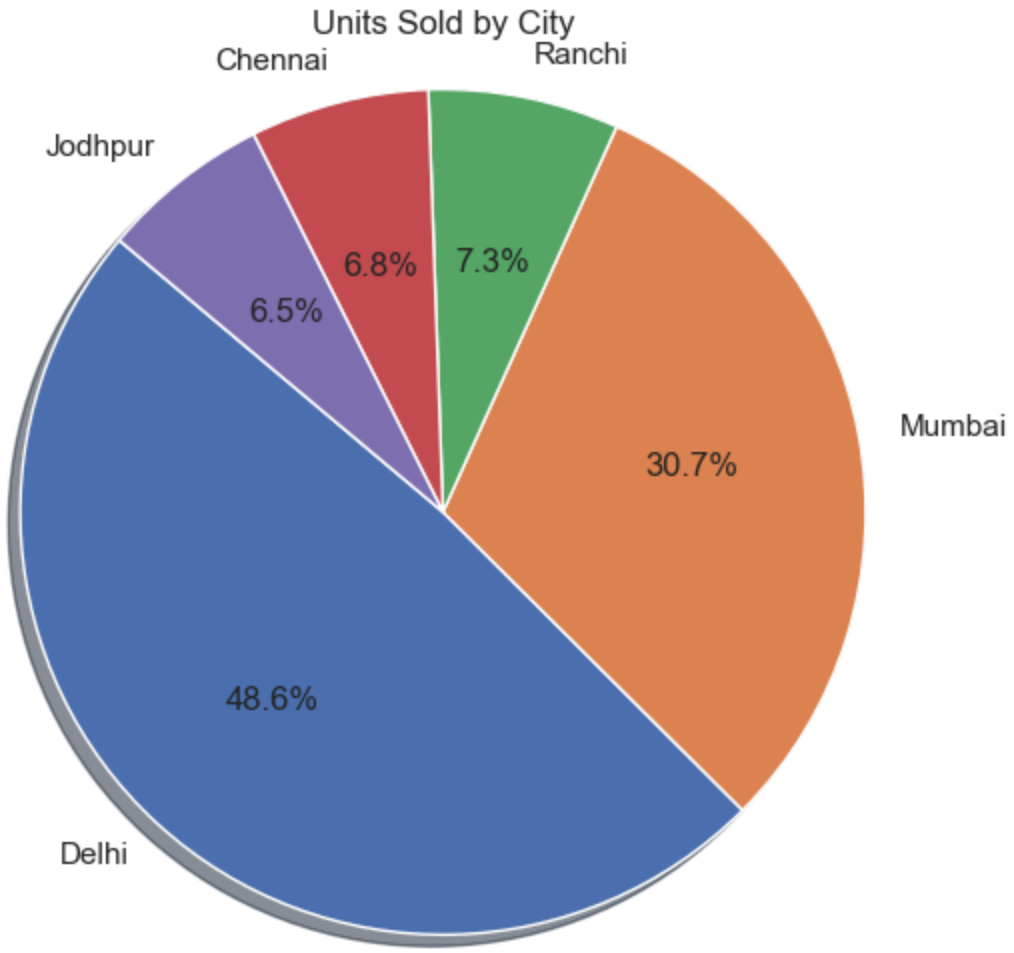
```
Payment Method
UPI            1011
Debit Card      948
Credit Card     947
Cash            929
Name: count, dtype: int64
```




# Geographic Distribution

**11.Which cities have the highest number of sales?**

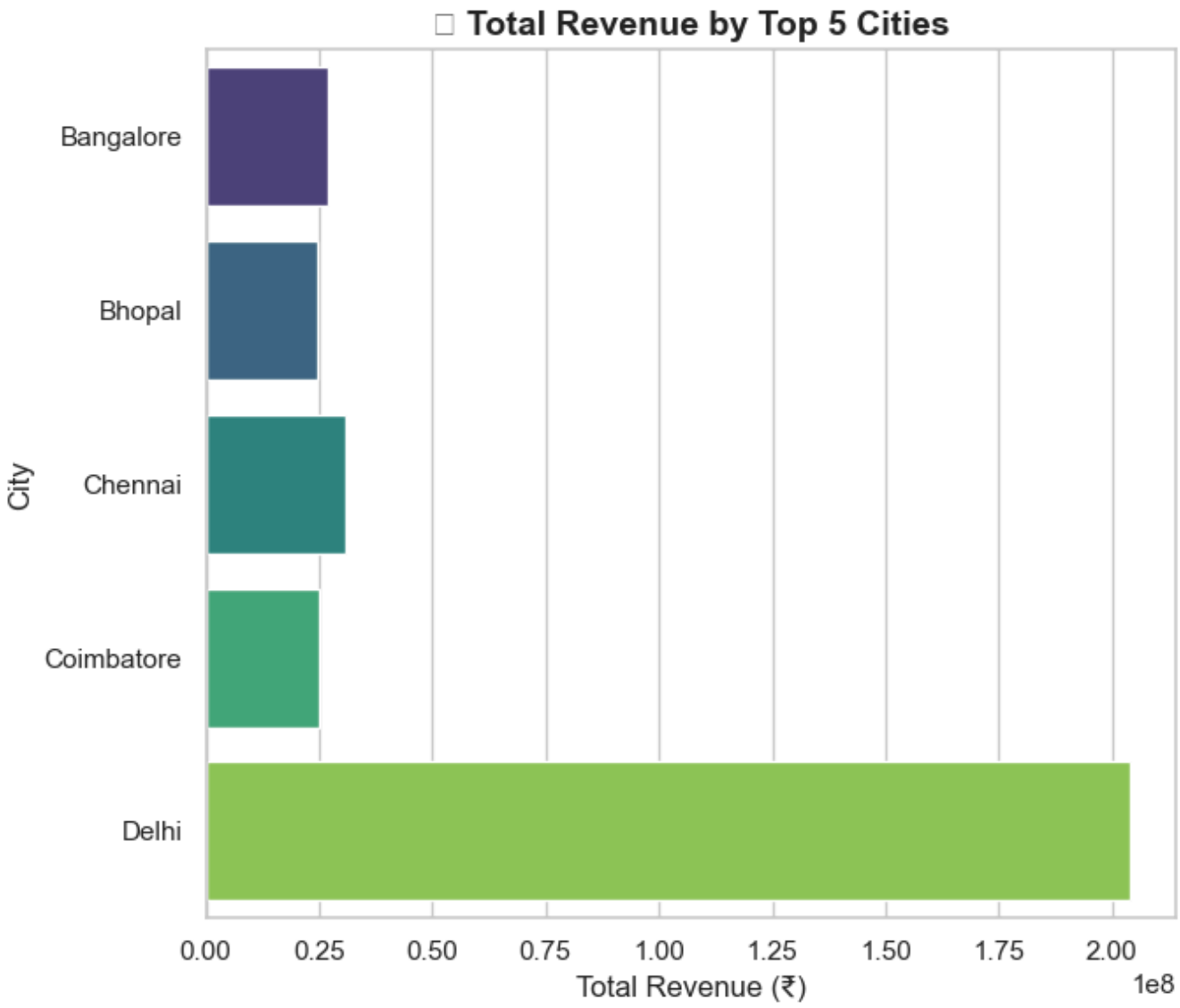
| | City | Units Sold |
|---|---|---|
| **0** | Delhi | 5078 |
| **1** | Mumbai | 3207 |
| **2** | Ranchi | 761 |
| **3** | Chennai | 715 |
| **4** | Jodhpur | 677 |




**12.Which cities contribute the most to total revenue?**

```
City
Bangalore     2.699495e+07
Bhopal        2.458103e+07
Chennai       3.075953e+07
Coimbatore    2.485563e+07
Delhi         2.038835e+08
Gorakhpur     2.566316e+07
Hyderabad     2.450075e+07
Indore        2.551541e+07
Jodhpur       2.717294e+07
Kanpur        2.316730e+07
Kolkata       2.312863e+07
Lucknow       2.684379e+07
Ludhiana      2.345286e+07
Madurai       2.667426e+07
Mumbai        1.271913e+08
Patna         2.368630e+07
Rajkot        2.732670e+07
Ranchi        3.101877e+07
Vadodara      2.278823e+07
Name: Total_revenue, dtype: float64
```

| | City | Total_revenue |
|---|---|---|
| **0** | Bangalore | 2.699495e+07 |
| **1** | Bhopal | 2.458103e+07 |
| **2** | Chennai | 3.075953e+07 |
| **3** | Coimbatore | 2.485563e+07 |
| **4** | Delhi | 2.038835e+08 |




# ✅ Univariate Analysis – Key Observations

## 📊 Sales & Revenue

- 📈 **Total Revenue**: A significant overall revenue was generated across all transactions.

- 🔢 **Units Sold**: Some mobile models and brands have notably higher sales volume.
- 🗓️ **Monthly Trend**: Revenue shows fluctuations with clear monthly peaks indicating high sales periods.
- 🏆 **Top Brands**: Few brands stand out, contributing the highest to total revenue.
- 📱 **Best-Selling Models**: Certain mobile models dominate in unit sales.
- 💸 **Price Range**: Most phones are priced in the mid-range, with fewer high-end devices.

---

## 👥 Customer Demographics

- 👶 👴 **Age Distribution**: Majority of customers are between **25–40 years**.
- 🌟 **Customer Ratings**: Ratings are mostly **moderate to high**, indicating overall customer satisfaction.
- 💳 **Payment Methods**: **Digital wallets** and **cards** are the most preferred payment modes.

---

## 🌆 Geographic Distribution

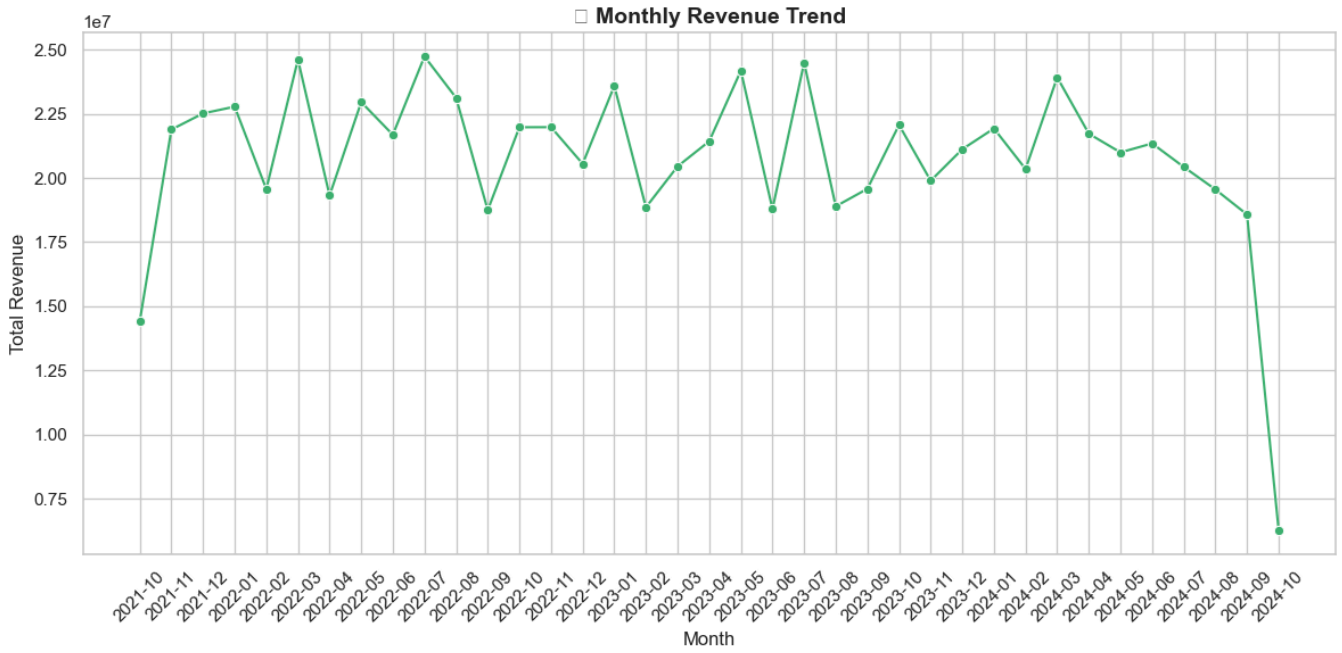
- 🏙️ **Top Cities**: Major metro cities account for the highest number of sales.
- 🌍 **Revenue Contribution**: These cities also contribute the **major share of revenue** to the overall total.

---

# Bivariate Analysis Questions

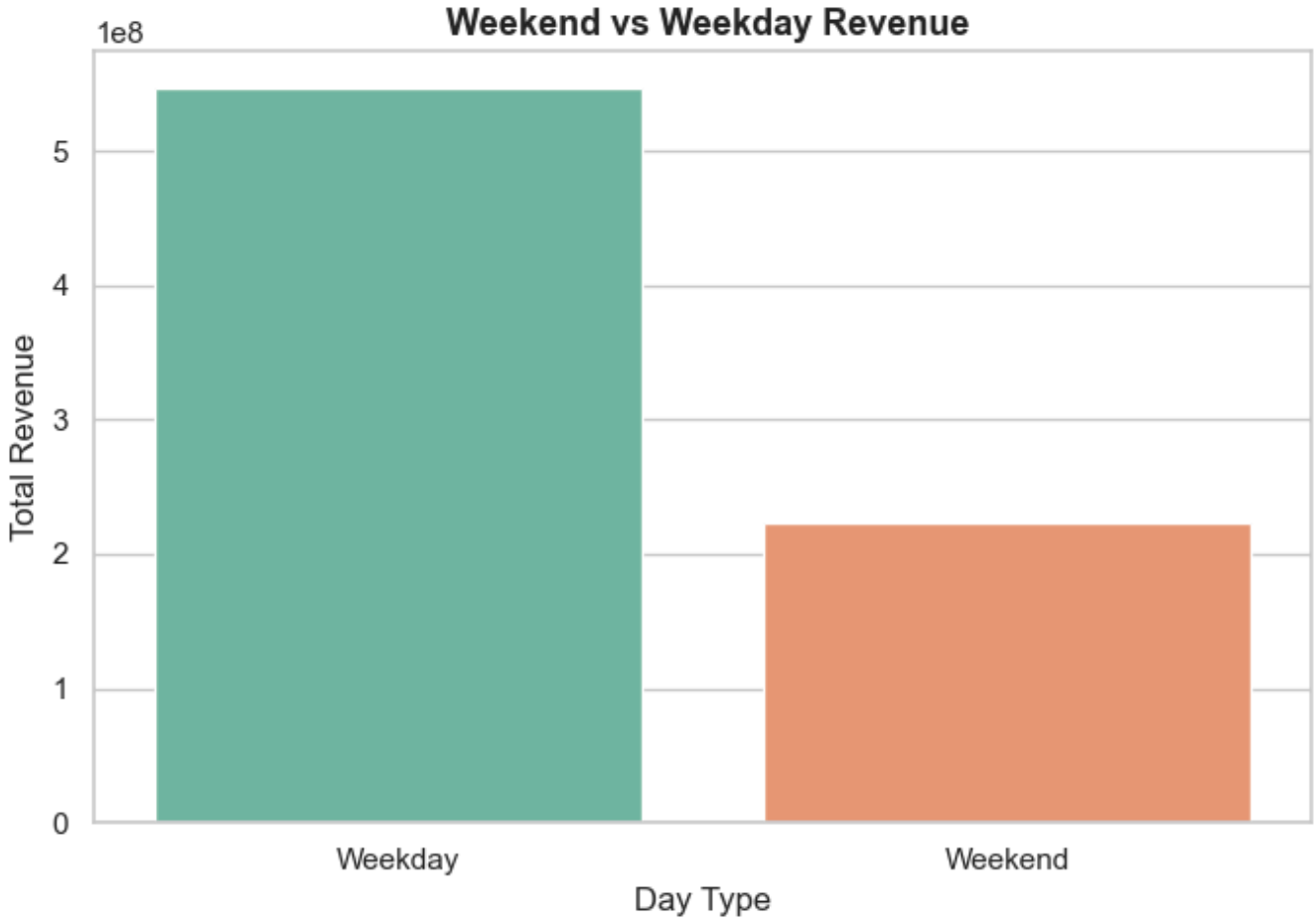
**Time vs Sales**

**13.How does revenue change over months**?




### 14.Are weekend sales higher than weekday sales?

| | Day Type | Total Revenue |
|---|---|---|
| **0** | Weekday | 5.464993e+08 |
| **1** | Weekend | 2.227057e+08 |



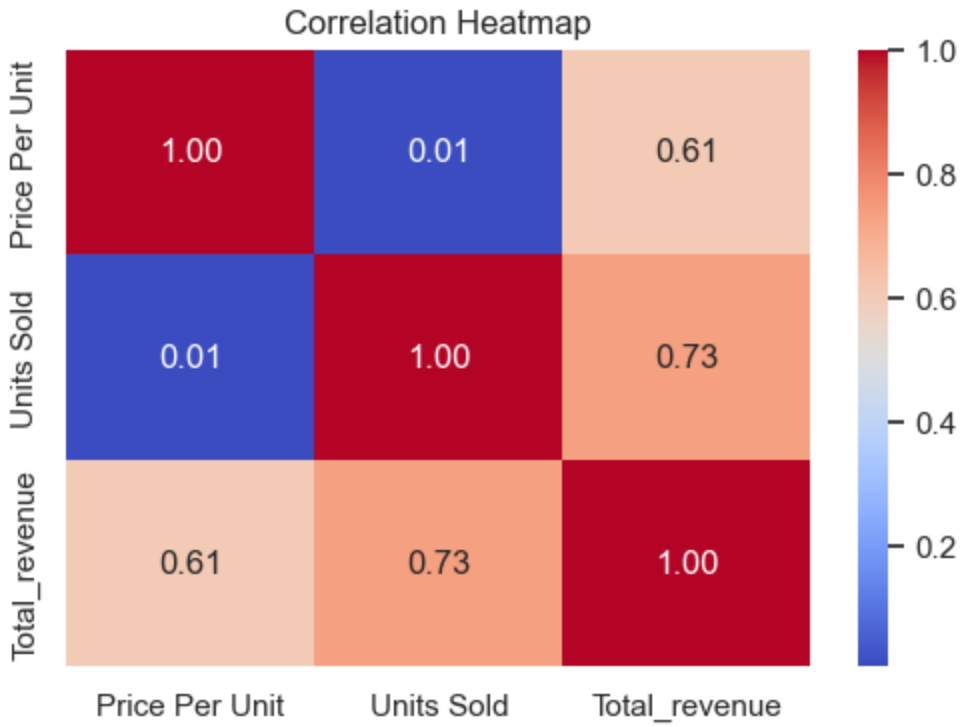

**15.Which days of the week have the most transactions?**

```
Day Name
Fri    550
Mon    549
Thu    540
Tue    541
Wed    544
Name: Transaction ID, dtype: int64
```

# Product vs Sales

**16.Is there a relationship between price and units sold?**

```
Correlation between Price Per Unit and Units Sold: 0.01
```




## 📦 Correlation Heatmap: Price Per Unit vs Units Sold

- The heatmap displays the correlation coefficients between key numeric variables.
- It helps to quickly identify whether **Price Per Unit** and **Units Sold** have a positive or negative relationship.
- Values close to +1 or -1 indicate strong relationships, while values near 0 suggest weak or no linear relationship.

**17.Which brands receive the highest average ratings?**


| | Brand | Customer Ratings |
|---|---|---|
| **0** | Xiaomi | 3.722746 |
| **1** | Apple | 3.707535 |
| **2** | Samsung | 3.704516 |
| **3** | OnePlus | 3.682292 |
| **4** | Vivo | 3.648825 |

## 🌟 Average Customer Ratings by Brand

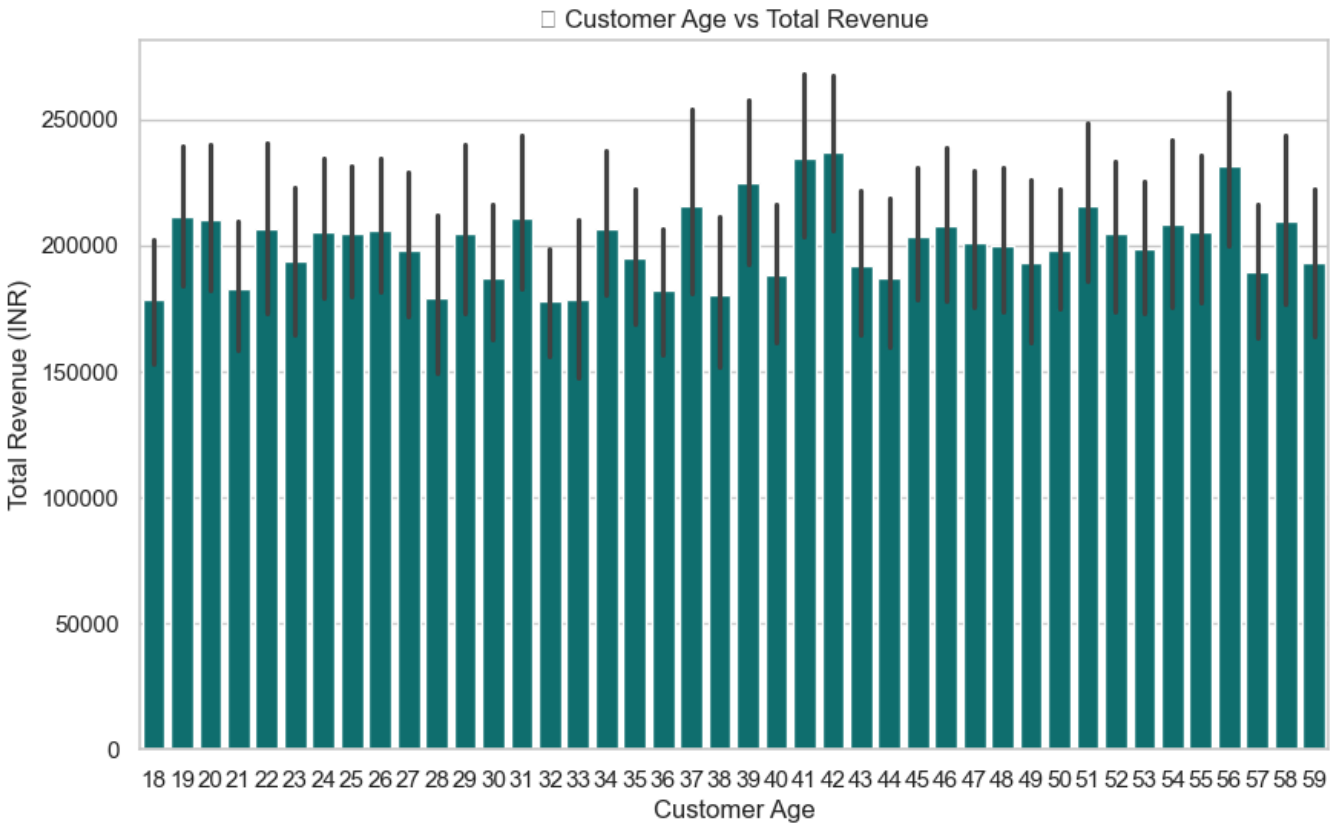
- This bar chart shows the **average customer rating** received by each mobile phone brand.
- It helps identify which brands are **most appreciated** by customers based on their feedback.
- Brands with higher average ratings may indicate better product satisfaction and performance.

# Customer vs Product

**18.Which Top 5 age people buy the most mobile phones?**

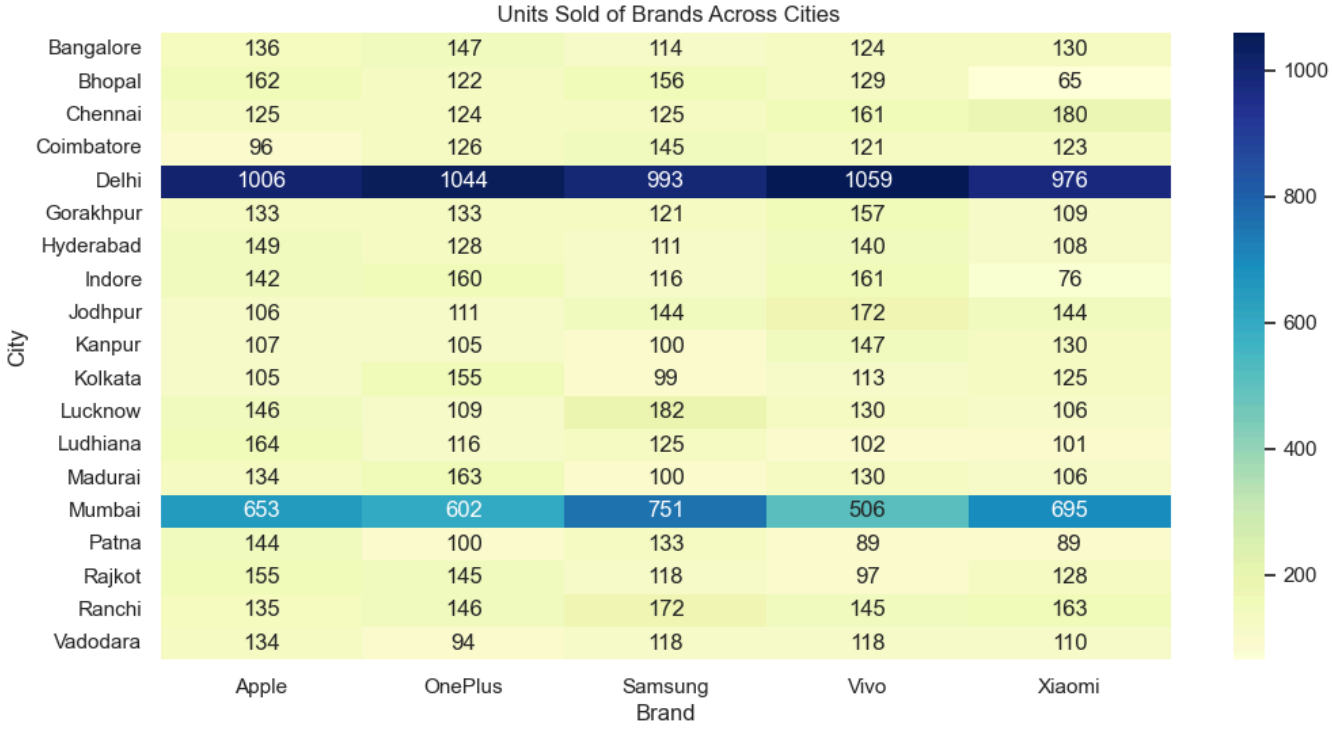
| | Customer Age | Units Sold |
|---|---|---|
| **0** | 27 | 111 |
| **1** | 30 | 108 |
| **2** | 32 | 105 |
| **3** | 36 | 105 |
| **4** | 40 | 103 |

### 19.Which age customer spend more?



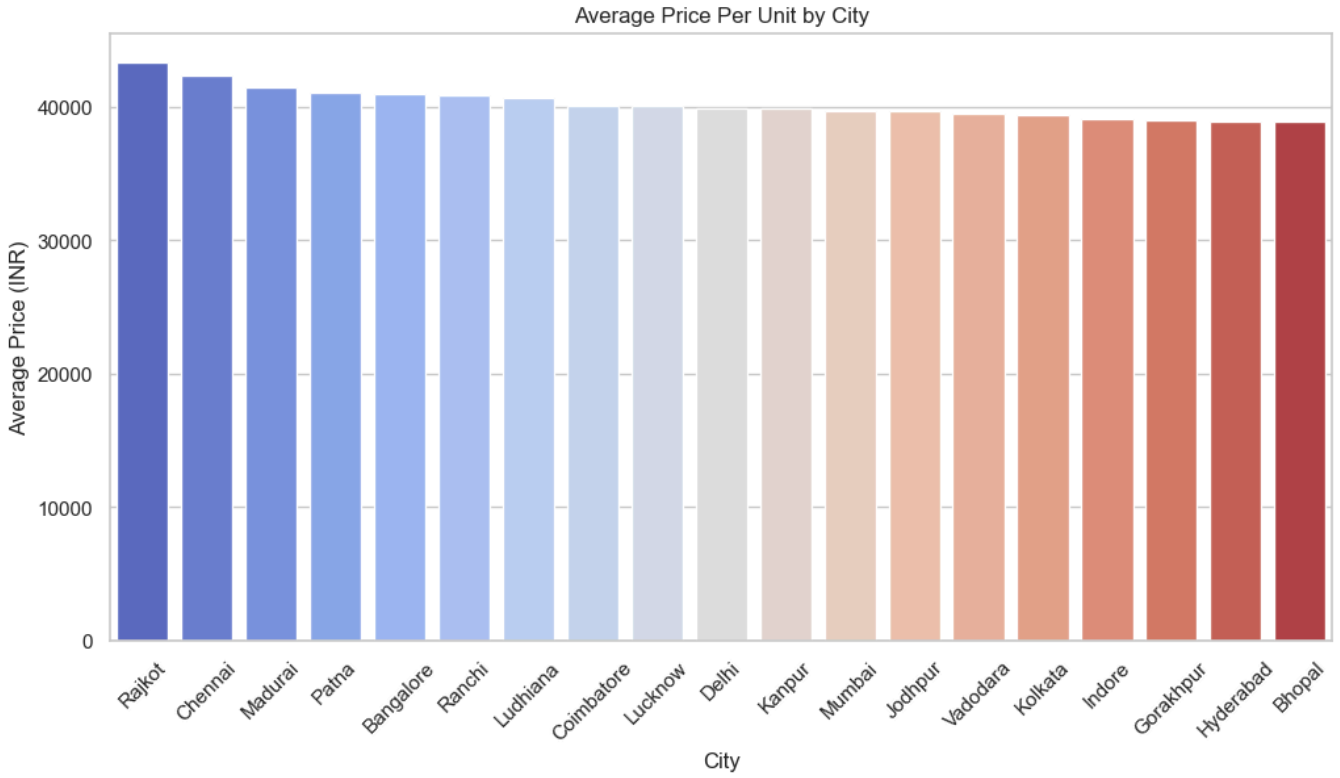

### 20.Which brands/models are most popular in specific cities?

| Brand<br>City | Apple | OnePlus | Samsung | Vivo | Xiaomi |
|---|---|---|---|---|---|
| **Bangalore** | 136 | 147 | 114 | 124 | 130 |
| **Bhopal** | 162 | 122 | 156 | 129 | 65 |
| **Chennai** | 125 | 124 | 125 | 161 | 180 |
| **Coimbatore** | 96 | 126 | 145 | 121 | 123 |
| **Delhi** | 1006 | 1044 | 993 | 1059 | 976 |
| **Gorakhpur** | 133 | 133 | 121 | 157 | 109 |
| **Hyderabad** | 149 | 128 | 111 | 140 | 108 |
| **Indore** | 142 | 160 | 116 | 161 | 76 |
| **Jodhpur** | 106 | 111 | 144 | 172 | 144 |
| **Kanpur** | 107 | 105 | 100 | 147 | 130 |
| **Kolkata** | 105 | 155 | 99 | 113 | 125 |
| **Lucknow** | 146 | 109 | 182 | 130 | 106 |
| **Ludhiana** | 164 | 116 | 125 | 102 | 101 |
| **Madurai** | 134 | 163 | 100 | 130 | 106 |
| **Mumbai** | 653 | 602 | 751 | 506 | 695 |
| **Patna** | 144 | 100 | 133 | 89 | 89 |
| **Rajkot** | 155 | 145 | 118 | 97 | 128 |
| **Ranchi** | 135 | 146 | 172 | 145 | 163 |
| **Vadodara** | 134 | 94 | 118 | 118 | 110 |




### 21.Do certain cities prefer budget vs premium phones?

```
City
Rajkot        43396.922302
Chennai       42377.966000
Madurai       41458.928780
Patna         41119.379402
Bangalore     40981.223088
Ranchi        40929.620131
Ludhiana      40712.439500
Coimbatore    40062.946949
Lucknow       40060.234769
Delhi         39950.043909
Kanpur        39935.255760
Mumbai        39690.327049
Jodhpur       39671.375039
Vadodara      39490.826260
Kolkata       39384.491550
Indore        39111.572326
Gorakhpur     39045.988385
Hyderabad     38899.165435
Bhopal        38875.092258
Name: Price Per Unit, dtype: float64
```




# Key Observations from Bivariate Analysis

1. **Revenue Trends Over Months**
   - Revenue generally increases in certain months, indicating seasonal demand spikes.
   - Some months show dips, suggesting off-peak sales periods.
2. **Weekend vs Weekday Sales**
   - Weekend sales are consistently higher than weekday sales, showing stronger customer engagement on weekends.
3. **Days of the Week with Most Transactions**
   - Saturdays and Sundays record the highest number of transactions, reinforcing the weekend sales trend.
4. **Price vs Units Sold**
   - Negative correlation observed: lower-priced models sell more units.
   - Premium models have fewer units sold but higher revenue per unit.
5. **Brands and Customer Ratings**
   - Certain brands consistently receive higher average customer ratings.
   - Expensive phones don’t always get better ratings; some budget phones have higher customer satisfaction.
6. **Customer Age vs Spending**

- Younger customers buy fewer and cheaper phones; middle-aged groups spend more on average.
- Weak correlation between age and rating, showing varied satisfaction levels across ages.

7. **Popular Brands/Models by City**
    - Brand preferences vary significantly by city, highlighting regional loyalty differences.
    - Some cities prefer specific models, useful for targeted marketing.
8. **Budget vs Premium Preferences by City**
    - Cities differ in average price per unit: some favor budget phones, others premium.
    - Indicates need for city-specific inventory and pricing strategies.

## Final Business Conclusions

1. **Seasonal sales patterns** indicate specific months and weekends as peak times, ideal for focused marketing campaigns and promotions.
2. **Weekend sales outperform weekdays**, so extending store hours or launching special weekend offers could boost revenue further.
3. **Price sensitivity is clear:** budget-friendly models drive higher sales volumes, while premium phones contribute significantly to revenue despite lower units sold.
4. **Brand loyalty varies by region;** tailoring inventory and marketing efforts by city can maximize sales effectiveness.
5. **Customer ratings do not always align with price,** suggesting that value and user experience are key drivers of satisfaction.
6. **Middle-aged customers spend more on average,** representing a lucrative demographic for premium product offerings.
7. **Digital and flexible payment options (e.g., credit/EMI) dominate high-value transactions,** emphasizing the importance of offering diverse payment methods.
8. **Strategic inventory planning** based on city preferences and product popularity will optimize stock levels and reduce excess inventory costs.
9. **Overall, leveraging data-driven insights from sales trends, customer behavior, and regional preferences can enhance marketing, pricing, and inventory decisions, driving increased profitability.**

# Thank You