CHAPTER -I
INTRODUCTION TO DATA ANALYTICS

Introduction

The "Operating System and Smartphone Sales Analysis" project aims to address this challenge by conducting a comprehensive and in-depth analysis of a dataset containing detailed information about smartphone sales. This project seeks to uncover key insights into consumer preferences, market trends, and the overall landscape of the smartphone industry. By meticulously examining various aspects of smartphone sales data, including operating systems, brand popularity, feature evaluations, and pricing, the project aims to provide a holistic view of the market dynamics. Through this analysis, the project aims to guide consumers in making informed purchasing decisions. By understanding which smartphone brands and features are most preferred, consumers can better navigate the multitude of options available to them and choose devices that align with their needs and preferences. The "Smartphone Sales Analysis" project delves into several key areas to provide a comprehensive understanding of the smartphone market. One primary focus is Operating System Preferences, aiming to identify which operating systems, such as Android and iOS, are most popular among consumers. This helps in recognizing market segmentation and potential growth areas for each platform. Another critical area is Brand Analysis, determining which smartphone brands, like Apple, Samsung, and Huawei, are preferred by customers based on sales data. This includes evaluating brand loyalty, market share, and regional performance, helping manufacturers strategize their market approaches. The project also emphasizes Feature Evaluation, assessing which features—such as camera quality, battery life, and processor speed—are most valued by consumers. By examining customer ratings and reviews, the project identifies the features that drive purchasing decisions and contribute to customer satisfaction. Lastly, the focus on Affordability and Specifications aims to find the balance between price and specifications, identifying affordable smartphones that still provide high customer satisfaction and good performance. This helps consumers make cost-effective decisions without compromising on quality.

1.2 Data Analytics Approaches

a) **Descriptive Analytics:** This approach focuses on summarizing and describing historical data to gain insights into past events and trends.

Descriptive analytics techniques include basic statistical measures, data visualization, and summary reports. It helps to understand what has happened and provides a foundation for further analysis.

b) **Diagnostic Analytics:** Diagnostic analytics aims to identify the reasons behind past events or trends. It involves analysing data to uncover patterns, correlations, or anomalies that can explain specific outcomes. Diagnostic analytics often utilizes statistical analysis, root cause analysis, and exploratory data analysis techniques.

c) **Predictive Analytics:** Predictive analytics involves using historical data to make predictions or forecasts about future events or outcomes. It leverages statistical modelling, machine learning algorithms, and data mining techniques to identify patterns and build predictive models. Predictive analytics helps organizations anticipate future trends, make informed decisions, and take proactive actions.

d) **Prescriptive Analytics:** Prescriptive analytics goes beyond prediction and provides recommendations on the best course of action to achieve desired outcomes. It uses optimization algorithms, simulation models, and decision analysis to generate actionable insights. Prescriptive analytics helps organizations optimize resources, streamline processes, and make data-driven decisions.

e) **Exploratory Analytics**: Exploratory analytics is an open-ended approach used to discover hidden patterns, relationships, or insights in data. Exploratory analytics is often used in research or when dealing with unstructured or large datasets

f) **Cognitive analytics:** It goes beyond traditional data analytics approaches by incorporating natural language processing, machine learning, pattern recognition, and contextual understanding to mimic human thought processes

1.3 Steps of Data Analytics

The steps of data analytics typically involve a systematic process that encompasses various stages. Here are the key steps involved in data analytics:

a) **Define the Problem:** Clearly articulate the problem or question you want to address through data analytics. Understand the objectives, scope, and desired

outcomes of the analysis. Formulate specific and measurable goals to guide your analysis.

Example: Analyze customer churn in a telecommunications company and identify factors influencing churn rates.

b) **Data Collection:** Identify and gather relevant data from various sources. This may include internal databases, external datasets, surveys, APIs, or other data collection methods. Ensure that the data collected aligns with the problem statement and objectives.

Example: Gather customer data, including demographics, usage patterns, service history, and churn status.

c) **Data Cleaning and Preprocessing:** Clean the collected data to ensure its quality and usability. This involves tasks such as handling missing values, removing duplicates, addressing outliers, standardizing formats, and transforming variables as needed

Example: Handle missing values, remove duplicates, and transform variables if necessary.

d) **Exploratory Data Analysis (EDA):** Perform exploratory data analysis to understand the characteristics of the dataset. Explore the data through statistical summaries, visualization techniques, and basic data exploration methods. Identify patterns, trends, relationships, and potential outliers in the data. EDA helps in gaining insights and formulating hypotheses for further analysis.

Example: Explore the data using visualizations and statistical summaries to identify patterns and relationships.

1.4 Applications of Data Analytics

a) **Education:** Educational institutions predict which students might drop out due to academic difficulties or financial issues. Targeted support services, such as tutoring or financial aid, can help at-risk students stay enrolled. Insights from churn analysis inform curriculum development, ensuring programs meet student needs and interests.

b) **Business Intelligence:** Data analytics is extensively used in business intelligence to gain insights into market trends, customer behavior, and

competitive landscapes. It helps businesses make informed decisions, optimize operations, and improve overall performance.

c) **Financial Analysis:** Data analytics plays a crucial role in financial analysis by analyzing financial data, detecting fraud, predicting market trends, and assessing investment opportunities. It helps financial institutions, investment firms, and businesses make data-driven decisions in areas such as risk management, portfolio optimization, and fraud detection.

CHAPTER -II

OVERVIEW OF THE PROBLEM

2.1. Problem Study

In the ever-evolving smartphone market, consumers are faced with a multitude of choices, each featuring different specifications, operating systems, and price points. This abundance of options makes it increasingly challenging for consumers to determine which smartphone best meets their needs and preferences. On the other hand, manufacturers need detailed insights into consumer behaviour and preferences to design products that cater to the market demands and stay competitive. The core problem addressed by the "Smartphone Sales Analysis" project is the need for a comprehensive understanding of these consumer preferences and market trends, to bridge the gap between consumer expectations and the products offered by manufacturers.

2.2. Existing & Proposed System

Existing System: Various analyses and studies have been conducted on smartphone sales and consumer preferences. Existing systems typically rely on sales data from manufacturers or third-party retailers to assess market trends. However, these systems often suffer from several limitations:

- Data Limitation: Limited access to comprehensive and up-to-date datasets.
- Static Analysis: Analyses are not regularly updated, leading to outdated conclusions.
- Lack of Detailed Specifications: Many studies focus solely on sales figures, ignoring important details like specifications and customer reviews.

Proposed System: The "Smartphone Sales Analysis" project proposes a more comprehensive approach by utilizing a detailed dataset obtained through web scraping from reputable online sources. This dataset includes information about various smartphone models, their specifications, features, pricing, and customer ratings.

The proposed system aims to:

- Identify Operating System Preferences: Determine the most popular operating systems (e.g., Android, iOS) among consumers.
- Analyze Brand Preferences: Ascertain the preferred smartphone brands (e.g., Apple, Samsung, Huawei) based on sales data.
- Evaluate Key Features: Assess which features (e.g., camera quality, battery life, processor speed) are most valued by consumers.

- Balance Affordability and Specifications: Identify affordable smartphones that offer good performance and high customer ratings.

2.3. Challenges/Need of the study

The study on smartphone sales presents several challenges and needs that must be addressed for effective analysis. A significant challenge is the availability and quality of data; incomplete or missing data can impede the analytical process, emphasizing the need for a comprehensive dataset. Data collected from various sources often requires extensive cleaning and processing to ensure consistency and accuracy, highlighting the necessity for efficient data cleaning techniques. The rapid pace of technological advancements in the smartphone industry poses another challenge, necessitating continuous updates to the dataset to stay relevant. Understanding consumer preferences, which can vary widely across different demographics and regions, is crucial for tailoring marketing strategies. The competitive nature of the smartphone market also underscores the need to identify key factors influencing consumer purchasing decisions. The analysis revealed that the top five most sold and popular smartphone brands primarily use the Android OS, with Snapdragon being the most used processor, while the Bionic processor, exclusive to Apple iPhones, is noted as the fastest. SQL was utilized to identify the most purchased brands and to introduce features that highlight cheaper phone models with good specifications.

2.4. Hardware / System Requirements

System requirements are the configuration that a system must have in order for a hardware or software application to run smoothly and efficiently. The system used in this project is Windows 11. It has 11th Gen Intel(R) Core (TM) i3-1115G4 @ 3.00GHz processor with 8.00GB RAM and 64-bit operating system, x64-based processor system type. GPU is Optional for large-scale data processing.

2.5. Software, Tools and Libraries Requirements

- Operating System: Windows 11
- Programming Language: Python 3.11.4
- Libraries:

  ✦ pandas: For data manipulation and analysis.

- ✦ numpy: For numerical computations.

- ✦ Seaborn: For data visualization.

- ✦ Matplotlib: For graphical plotting.

- ✦ SQL:For connecting the database

- Development Environment: Jupyter Notebook or any Python IDE.

- Data Source: Excel, CSV files.

PYTHON:

Python is a high-level programming language that is widely used for various purposes, such as web development, data analysis, machine learning, scientific computing, and more.

Jupyter Notebook is an open-source web application that allows you to create and share documents that contain live code, equations, visualizations, and narrative text. It is commonly used for data analysis, scientific computing, machine learning, and more.

CHAPTER - III

DATA PREPARATION

3.1. Data Collection Approaches

The dataset used for this project is sourced from Kaggle. It comprises 981 entries and 22 variables. It provides various features that are significant for predicting smartphone sales.

Dataset link:

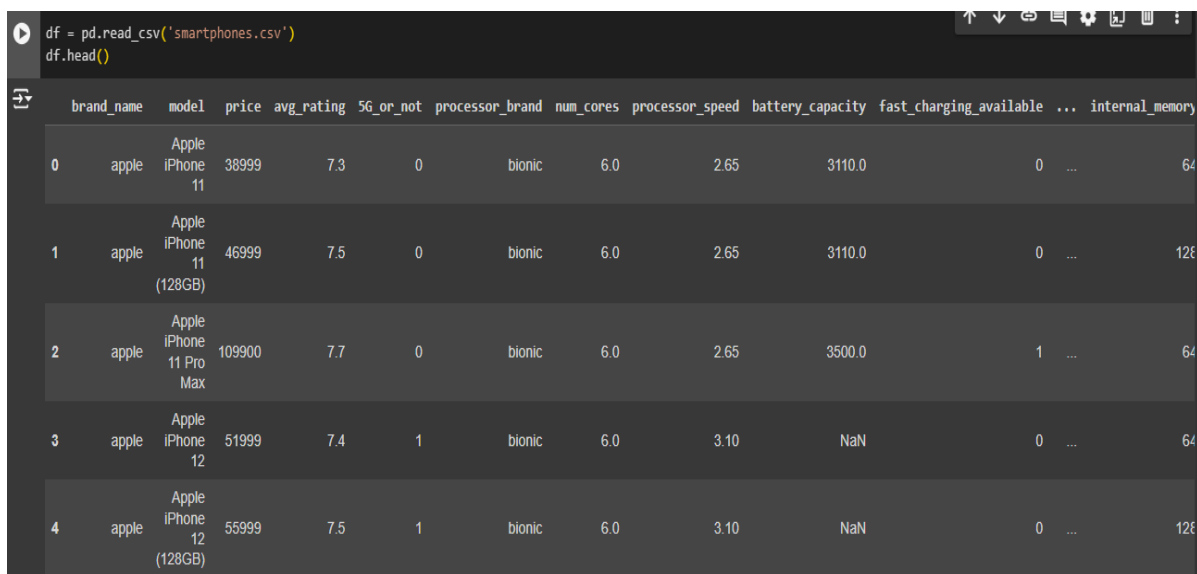https://www.kaggle.com/datasets/real-world-smartphones-dataset

3.2. Data Method

Exploratory Data Analysis: Exploratory Data Analysis is a method of evaluating or comprehending data in order to derive insights or key characteristics. EDA can be divided into two categories: graphical analysis and non-graphical analysis. EDA is a critical component of any data science or machine learning process. You must explore the data, understand the relationships between variables, and the underlying structure of the data in order to build a reliable and valuable output based on it.

Data Loading and Exploration:

The dataset was loaded, and initial exploration was performed to understand the data structure and distributions.

```
df = pd.read_csv('smartphones.csv')
df.head()
```

	brand_name	model	price	avg_rating	5G_or_not	processor_brand	num_cores	processor_speed	battery_capacity	fast_charging_available	...	internal_memory
0	apple	Apple iPhone 11	38999	7.3	0	bionic	6.0	2.65	3110.0	0	...	64
1	apple	Apple iPhone 11 (128GB)	46999	7.5	0	bionic	6.0	2.65	3110.0	0	...	128
2	apple	Apple iPhone 11 Pro Max	109900	7.7	0	bionic	6.0	2.65	3500.0	1	...	64
3	apple	Apple iPhone 12	51999	7.4	1	bionic	6.0	3.10	NaN	0	...	64
4	apple	Apple iPhone 12 (128GB)	55999	7.5	1	bionic	6.0	3.10	NaN	0	...	128

EXPLORATION DATA ANALYSIS

1.Decribe the info of the dataset

```
df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 980 entries, 0 to 979
Data columns (total 22 columns):
 #   Column                    Non-Null Count  Dtype
---  ------                    --------------  -----
 0   brand_name                980 non-null    object
 1   model                     980 non-null    object
 2   price                     980 non-null    int64
 3   avg_rating                879 non-null    float64
 4   5G_or_not                 980 non-null    int64
 5   processor_brand           960 non-null    object
 6   num_cores                 974 non-null    float64
 7   processor_speed           938 non-null    float64
 8   battery_capacity          969 non-null    float64
 9   fast_charging_available   980 non-null    int64
 10  fast_charging             769 non-null    float64
 11  ram_capacity              980 non-null    int64
 12  internal_memory           980 non-null    int64
 13  screen_size               980 non-null    float64
 14  refresh_rate              980 non-null    int64
 15  num_rear_cameras          980 non-null    int64
 16  os                        966 non-null    object
 17  primary_camera_rear       980 non-null    float64
 18  primary_camera_front      975 non-null    float64
 19  extended_memory_available 980 non-null    int64
 20  resolution_height         980 non-null    int64
```

2.Decribe the dataset by showing values

df.describe()

	price	avg_rating	5G_or_not	num_cores	processor_speed	battery_capacity	fast_charging_available	fast_charging	ram_capacity	internal_memory
count	980.000000	879.000000	980.000000	974.000000	938.000000	969.000000	980.000000	769.000000	980.000000	980.000000
mean	32520.504082	7.825825	0.560204	7.772074	2.427217	4817.748194	0.854082	46.126138	6.560204	141.036735
std	39531.812669	0.740285	0.496616	0.836845	0.464090	1009.540054	0.353205	34.277870	2.744378	107.134516
min	3499.000000	6.000000	0.000000	4.000000	1.200000	1821.000000	0.000000	10.000000	1.000000	8.000000
25%	12999.000000	7.400000	0.000000	8.000000	2.050000	4500.000000	1.000000	18.000000	4.000000	64.000000
50%	19994.500000	8.000000	1.000000	8.000000	2.300000	5000.000000	1.000000	33.000000	6.000000	128.000000
75%	35491.500000	8.400000	1.000000	8.000000	2.840000	5000.000000	1.000000	66.000000	8.000000	128.000000
max	650000.000000	8.900000	1.000000	8.000000	3.220000	22000.000000	1.000000	240.000000	18.000000	1024.000000

9

Data cleaning

3.Finding the null values

```
df.isnull().sum()
```

```
brand_name                    0
model                         0
price                         0
avg_rating                  101
5G_or_not                     0
processor_brand              20
num_cores                     6
processor_speed              42
battery_capacity             11
fast_charging_available       0
fast_charging               211
ram_capacity                  0
internal_memory               0
screen_size                   0
refresh_rate                  0
num_rear_cameras              0
os                           14
primary_camera_rear           0
primary_camera_front          5
extended_memory_available     0
resolution_height             0
resolution_width              0
dtype: int64
```

4.Filling the null values

```
df = df.ffill(axis= 0)
df.head()
```

	brand_name	model	price	avg_rating	5G_or_not	processor_brand	num_cores	processor_speed	battery_capacity	fast_charging_available	...	internal_memory
0	apple	Apple iPhone 11	38999	7.3	0	bionic	6.0	2.65	3110.0	0	...	64
1	apple	Apple iPhone 11 (128GB)	46999	7.5	0	bionic	6.0	2.65	3110.0	0	...	128
2	apple	Apple iPhone 11 Pro Max	109900	7.7	0	bionic	6.0	2.65	3500.0	1	...	64
3	apple	Apple iPhone 12	51999	7.4	1	bionic	6.0	3.10	3500.0	0	...	64
4	apple	Apple iPhone 12 (128GB)	55999	7.5	1	bionic	6.0	3.10	3500.0	0	...	128

```
df.isnull().sum()
```

```
brand_name                    0
model                         0
price                         0
avg_rating                    0
5G_or_not                     0
processor_brand               0
num_cores                     0
processor_speed               0
battery_capacity              0
fast_charging_available       0
fast_charging                 2
ram_capacity                  0
internal_memory               0
screen_size                   0
refresh_rate                  0
num_rear_cameras              0
os                            0
primary_camera_rear           0
primary_camera_front          0
extended_memory_available     0
resolution_height             0
resolution_width              0
dtype: int64
```
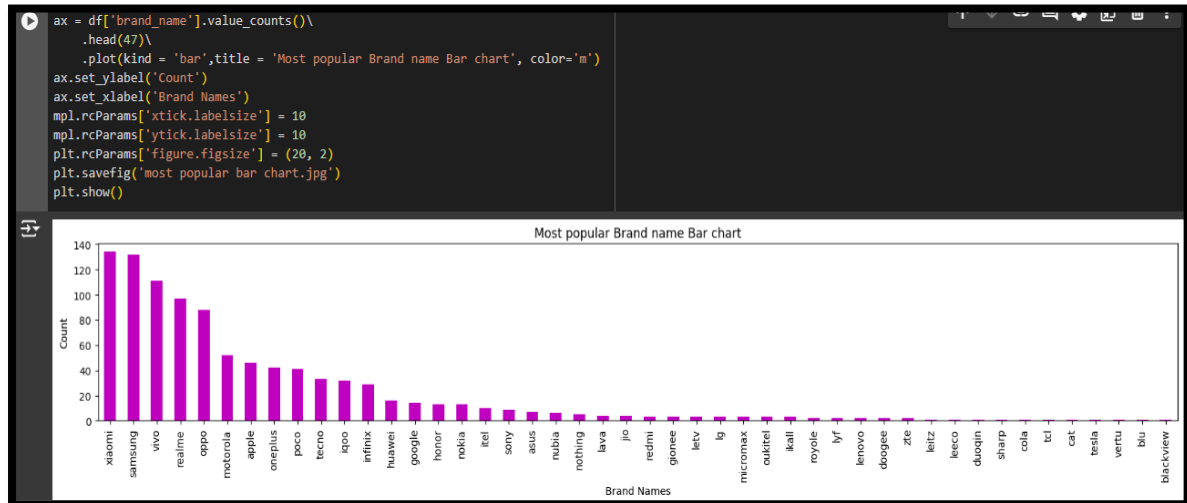
```
df.brand_name.value_counts()
```

```
brand_name
xiaomi      134
samsung     132
vivo        111
realme       97
oppo         88
motorola     52
apple        46
oneplus      42
poco         41
tecno        33
iqoo         32
infinix      29
huawei       16
google       14
honor        13
nokia        13
itel         10
sony          9
asus          7
nubia         6
nothing       5
lava          4
jio           4
redmi         3
gionee        3
```

Data Visualization

Data visualization is the graphical representation of information and data. It involves creating visual elements like charts, graphs, and maps to make complex data more accessible, understandable, and usable.


```
ax = df['brand_name'].value_counts()\
    .head(47)\
    .plot(kind = 'bar',title = 'Most popular Brand name Bar chart', color='m')
ax.set_ylabel('Count')
ax.set_xlabel('Brand Names')
mpl.rcParams['xtick.labelsize'] = 10
mpl.rcParams['ytick.labelsize'] = 10
plt.rcParams['figure.figsize'] = (20, 2)
plt.savefig('most popular bar chart.jpg')
plt.show()
```

3.3. Purpose of Data

The dataset used in the "Smartphone Sales Analysis" project serves to provide a comprehensive understanding of the smartphone market by identifying consumer preferences, evaluating key features, balancing affordability and specifications, and analyzing market trends. It includes detailed information on smartphone models, operating systems, brands, features, and pricing, enabling the identification of popular operating systems and brands, as well as the assessment of valued features like camera quality and battery life. The dataset helps find affordable smartphones offering high performance and customer satisfaction, aids in understanding broader market trends, and supports strategic decision-making for manufacturers. For consumers, it enhances their knowledge, enabling informed purchasing decisions by highlighting top-rated smartphones and preferred features. Overall, the dataset forms the backbone of the analysis, offering valuable insights that benefit both consumers and manufacturers in navigating the complex smartphone market.

CHAPTER - IV

METHODOLOGY

Descriptive Analysis / Model Used

Descriptive analysis in python involves summarizing and examining data to gain insights and understand its characteristics. It helps in understanding the central tendencies, distributions, and variability of the data.

The "Smartphone Sales Analysis" project employs a structured and systematic methodology to ensure comprehensive analysis and reliable insights. The methodology is divided into several key steps, including data collection, data cleaning, data analysis, and result interpretation. Each step is crucial to achieving the project's objectives.

a) **Data Collection**

The dataset for this project was created through web scraping from Kaggle. This involves various smartphone models, their specifications, features, pricing, and customer ratings The collected data includes details such as brand, operating system, camera quality, battery life, processor speed, display size, and customer reviews.

b) **Data Cleaning**

Once the data is collected, it undergoes a thorough cleaning process to ensure accuracy and completeness. This step involves handling missing values, removing duplicates, and correcting inconsistencies. Techniques such as forward filling and interpolation are used to address missing data, ensuring that the dataset is robust and reliable for analysis.

c) **Data Analysis**

The cleaned dataset is then subjected to detailed analysis using various statistical and visualization techniques. Python libraries such as NumPy, pandas, matplotlib, seaborn, and plotly are employed for data manipulation and visualization. The analysis focuses on several key areas:

- Operating System Preferences: Identifying the most popular operating systems (e.g., Android,iOS) among consumers.
- Brand Analysis: Determining which smartphone brands (e.g., Apple, Samsung, Huawei) are preferred by customers based on sales data
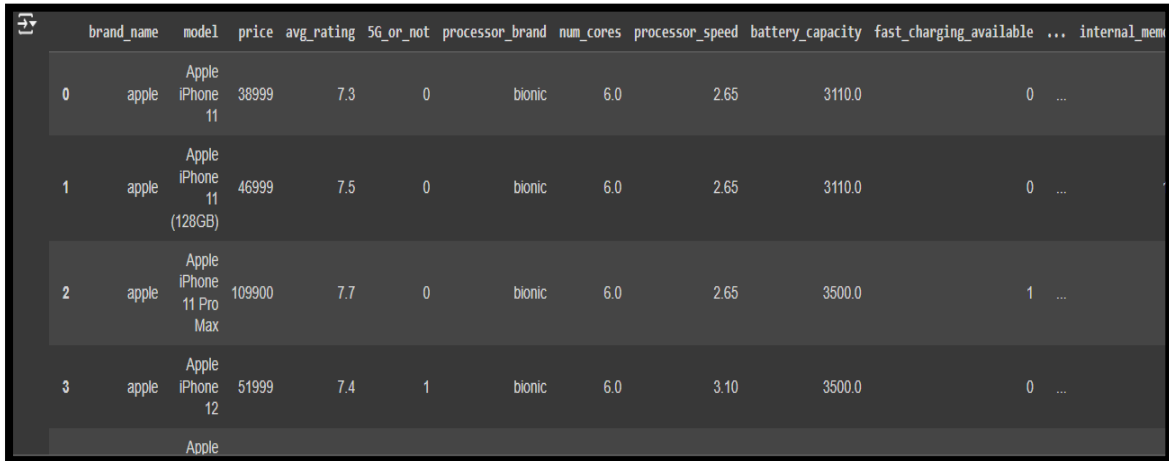
- Feature Evaluation: Assessing which features (e.g., camera quality, battery life, processor speed) are most valued by consumers by examining customer ratings and reviews.

d) Affordability and Specifications: Finding the balance between price and specifications to identify affordable smartphones that offer good performance and high customer satisfaction.

e) Result Interpretation: The results of the data analysis are interpreted to draw meaningful conclusions about consumer preferences and market trends. This involves synthesizing the findings to identify patterns, correlations, and actionable insights. The interpretation helps in understanding the factors that influence purchasing decisions and market dynamics.

f) Reporting: Finally, the findings are compiled into a comprehensive report that details the methodology, analysis, results, and conclusions. The report provides valuable insights for both consumers and manufacturers, guiding informed purchasing decisions and planning.

RESULTS & FINDING INSIGHTS

Result-Structure query language And Python

1.Read the csv dataset by using structure query language

	brand_name	model	price	avg_rating	5G_or_not	processor_brand	num_cores	processor_speed	battery_capacity	fast_charging_available	...	internal_memo
0	apple	Apple iPhone 11	38999	7.3	0	bionic	6.0	2.65	3110.0	0	...	
1	apple	Apple iPhone 11 (128GB)	46999	7.5	0	bionic	6.0	2.65	3110.0	0	...	
2	apple	Apple iPhone 11 Pro Max	109900	7.7	0	bionic	6.0	2.65	3500.0	1	...	
3	apple	Apple iPhone 12	51999	7.4	1	bionic	6.0	3.10	3500.0	0	...	
		Apple										

The above output read the csv dataset by us structural query language and display the table

2.Brand names that use snapdragon processor

| index	brand_name
0	asus
1	cat
2	google
3	honor
4	huawei
5	ikall
6	iqoo
7	jio
8	leitz
9	lenovo
10	lg
11	lyf
12	motorola
13	nokia

1 to 25 of 29 entries   Filter

The above output represents that which smartphone brand that use snapdragon processor

3.Cheap phones with good specification

	brand_name	model	price	avg_rating	5G_or_not	processor_brand	num_cores	processor_speed	battery_capacity	fast_charging_available	...	internal_me
0	blackview	Blackview BV5200 Pro	8990	6.7	1	helio	8.0	2.3	5180.0	1	...	
1	honor	Honor Play 30	9999	6.8	1	snapdragon	8.0	2.2	5000.0	0	...	
2	infinix	Infinix Hot 12	9999	7.2	0	helio	8.0	2.3	6000.0	1	...	
3	infinix	Infinix Hot 20 Play	8999	6.7	0	helio	8.0	2.3	6000.0	1	...	
4	lava	Lava Blaze NXT	8999	6.6	0	helio	8.0	2.3	5000.0	0	...	
5	motorola	Motorola Moto E22s	9022	7.0	0	helio	8.0	2.3	5000.0	1	...	

The above output represents that blackview pro smartphone has low price with good specification

4.Ploting affordable phones with good specification
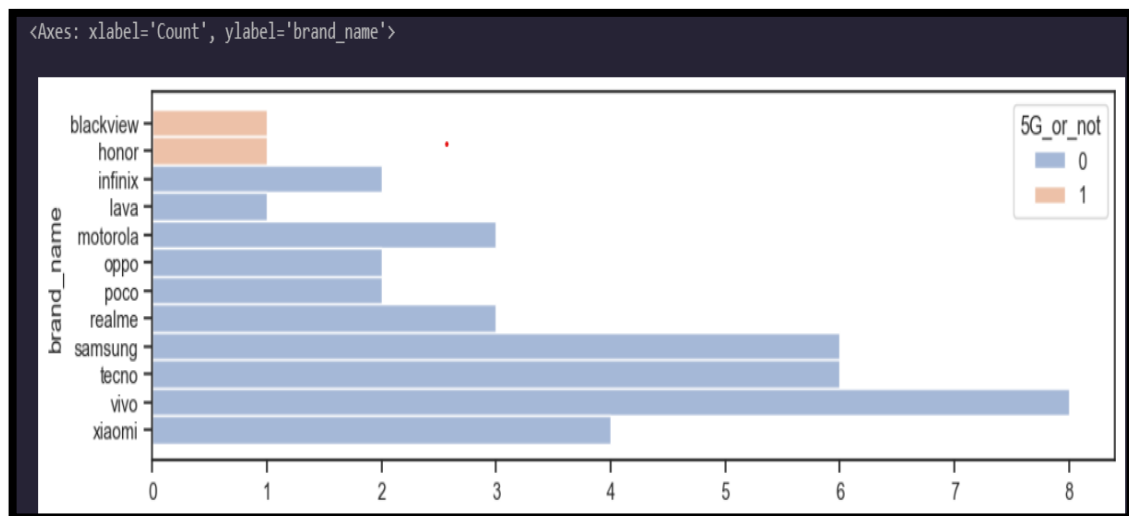
The above output represents that xiaomi has the most affordable phones with good specification

CHAPTER - VI

DICUSSION & FUTURE WORK

The analysis of smartphone sales data has provided several valuable insights into market trends and consumer preferences. Key findings include the dominance of Android OS among the top-selling smartphone brands and the prevalence of Snapdragon processors. However, certain limitations in the analysis, such as potential data quality issues, sample size constraints, and lack of contextual information, must be acknowledged.

The analysis revealed that while Snapdragon processors are widely used, Apple's Bionic processor stands out for its performance, suggesting a competitive edge for iPhones in terms of speed and efficiency. Additionally, the SQL-based analysis highlighted the most popular phone brands and identified cost-effective models with good specifications, which can guide consumers in making informed purchasing decisions.

Future Work

To build upon the current analysis and address its limitations, the following future work is recommended:

1. Enhancing Data Quality and Sources

Objective: Integrate additional data sources and ensure the dataset is comprehensive, accurate, and up-to-date.
Implementation: Collaborate with market research firms, utilize APIs from major retailers, and include data from more diverse regions and demographics.

2. Expanding Sample Size

Objective: Increase the dataset size to improve the robustness and generalizability of the findings.
Implementation: Collect data over a longer period and from multiple sources to ensure a more representative sample of the market.

3. Incorporating Contextual Factors

Objective: Consider external factors such as marketing campaigns, seasonal trends, economic conditions, and new product releases.

Implementation: Develop a framework to integrate and analyze these factors alongside sales data, potentially using time series analysis or econometric modelling.

4.Utilizing Advanced Analytical Method

Objective: Employ more sophisticated analytical techniques to uncover deeper insights and more complex relationships within the data.
Implementation: Use machine learning models for predictive analytics, cluster analysis to identify consumer segments, and network analysis to explore brand interconnections.

5.Dynamic Trend Analysis

Objective: Continuously update the analysis to reflect real-time changes and trends in the smartphone market.
Implementation: Set up automated data pipelines and dashboards that provide live updates and visualizations of key metrics and trends.

CHAPTER - VII

SUMMARY & CONCLUSION

Summary

The project involves analyzing a smartphone sales dataset to determine which operating system and phone brand customers prefer and rate the most, which brands offer the best features, and which phones are affordable with good specifications. The dataset was created by web scraping reputable online sources to gather accurate and up-to-date information about various smartphone models, their specifications, features, and pricing. Missing values in the dataset were handled by interpolating empty cells with average values from surrounding data points

Key analysis points included:

- Total sales
- Most popular brand name (Xiaomi)
- Most popular operating system
- Brands with high customer ratings
- Processor brands with the highest speeds
- Phone brands with expensive prices
- Analyzing data using SQL
- Identifying cheap phones with good specifications

Conclusion

The analysis of the smartphone sales data revealed that the top five most sold and popular phone brands predominantly use the Android OS. The Snapdragon processor is the most commonly used, while the Bionic processor, exclusive to Apple iPhones, is the fastest. Using SQL for data analysis, the project identified the most purchased phone brands and highlighted cheaper phone models with good specifications. This comprehensive analysis provides insights into customer preferences and market trends in the smartphone industry.

CHAPTER - VIII
REFERENCES

Web links:

1.Introduction to data analytics

https://careerfoundry.com/en/blog/data-analytics/what-is-data-analytics/

2. Approach of analytics

https://www.kdnuggets.com/2023/04/data-analytics-four-approaches-analyzing-data-effectively.html

3.Step of data analytics

https://www.geeksforgeeks.org/six-steps-of-data-analysis-process/

4.Application of data analytics

https://www.spiceworks.com/tech/data-management/articles/what-is-data-analytics/

5.Dataset

https://www.kaggle.com/datasets/real-world-smartphones-dataset

Research paper 1

https://www.academia.edu/3983376/Smartphone_Market_Research_Report

Books:

"Python for Data Analysis" by Wes McKinney,

"Python Data Visualization Cookbook" by Mario (Author),

"SQL AND Python Programming" by brain Johnson (Author)

1.Read the csv dataset by using structure query language

```
import sqlite3
conn = sqlite3.connect(':memory:')
c = conn.cursor()
df.to_sql('df', conn, if_exists='replace', index=False)
pd.read_sql('SELECT * FROM df', conn)
```

2.Brand names that use snapdragon processor

```
brand = pd.read_sql("""
    SELECT brand_name FROM df WHERE processor_brand = 'snapdragon'
    GROUP BY brand_name """, conn)
brand
```

3.Cheap phones with good specification

```
brand_price = pd.read_sql("""SELECT * FROM df WHERE price <= 10000 AND
processor_speed > 2.0""", conn)

brand_price
```

4.Ploting affordable phones with good specification

```
sns.set_theme(style='ticks')

sns.histplot(

    brand_price,

    y= 'brand_name', hue= '5G_or_not')
```