iPhone Sales Analysis

Name of Students:

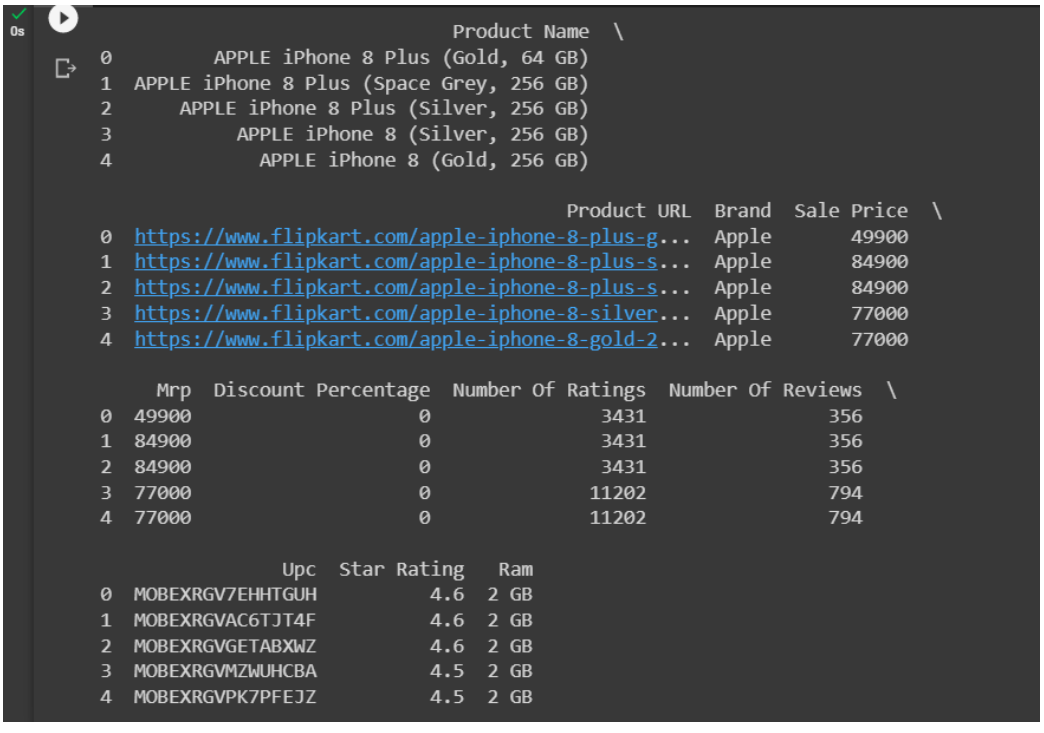
Sr No.	Name	Roll No.
1	Chordiya Yashkumar Umesh	2
2	Jain Shruti Anup	4
3	Desarda Khushi Sandeep	7
4	Surywanshi Vedant Sunil	61

Code of iphone sales Analysis:

1. import pandas as pd
   import numpy as np
   import plotly.express as px
   import plotly.graph_objects as go
   from sklearn.cluster import KMeans

   data = pd.read_csv("apple_products.csv")
   print(data.head())
   //This code reads in a csv filenamed "apple_products.csv".

Output:



```
0s
                                Product Name  \
0        APPLE iPhone 8 Plus (Gold, 64 GB)
1  APPLE iPhone 8 Plus (Space Grey, 256 GB)
2     APPLE iPhone 8 Plus (Silver, 256 GB)
3           APPLE iPhone 8 (Silver, 256 GB)
4             APPLE iPhone 8 (Gold, 256 GB)

                                     Product URL  Brand  Sale Price  \
0  https://www.flipkart.com/apple-iphone-8-plus-g...  Apple       49900
1  https://www.flipkart.com/apple-iphone-8-plus-s...  Apple       84900
2  https://www.flipkart.com/apple-iphone-8-plus-s...  Apple       84900
3  https://www.flipkart.com/apple-iphone-8-silver...  Apple       77000
4  https://www.flipkart.com/apple-iphone-8-gold-2...  Apple       77000

     Mrp  Discount Percentage  Number Of Ratings  Number Of Reviews  \
0  49900                    0               3431                356
1  84900                    0               3431                356
2  84900                    0               3431                356
3  77000                    0              11202                794
4  77000                    0              11202                794

                Upc  Star Rating   Ram
0  MOBEXRGV7EHHTGUH          4.6  2 GB
1  MOBEXRGVAC6TJT4F          4.6  2 GB
2  MOBEXRGVGETABXWZ          4.6  2 GB
3  MOBEXRGVMZWUHCBA          4.5  2 GB
4  MOBEXRGVPK7PFEJZ          4.5  2 GB
```

2. print(data.isnull().sum())

   //It Checks whether this dataset contains any null values or not.

Output:

```
Product Name            0
Product URL             0
Brand                   0
Sale Price              0
Mrp                     0
Discount Percentage     0
Number Of Ratings       0
Number Of Reviews       0
Upc                     0
Star Rating             0
Ram                     0
dtype: int64
```

3. print(data.describe())

   //The descriptive statistics of the data.

Output:

```
          Sale Price            Mrp  Discount Percentage  Number Of Ratings  \
count      62.000000      62.000000            62.000000          62.000000
mean    80073.887097   88058.064516             9.951613       22420.403226
std     34310.446132   34728.825597             7.608079       33768.589550
min     29999.000000   39900.000000             0.000000         542.000000
25%     49900.000000   54900.000000             6.000000         740.000000
50%     75900.000000   79900.000000            10.000000        2101.000000
75%    117100.000000  120950.000000            14.000000       43470.000000
max    140900.000000  149900.000000            29.000000       95909.000000

       Number Of Reviews  Star Rating
count          62.000000    62.000000
mean         1861.677419     4.575806
std          2855.883830     0.059190
min            42.000000     4.500000
25%            64.000000     4.500000
50%           180.000000     4.600000
75%          3331.000000     4.600000
max          8161.000000     4.700000
```

iPhone Sales Analysis in India :

4. highest_rated = data.sort_values(by=["Star Rating"], ascending=False)
   highest_rated = highest_rated.head(10)
   print(highest_rated['Product Name'])

//Now I will create a new dataframe by storing all the data about the top 10 highest-rated iPhones in India on Flipkart. It will help in understanding what kind of iPhones are liked the most in India.

Output:

```
20     APPLE iPhone 11 Pro Max (Midnight Green, 64 GB)
17         APPLE iPhone 11 Pro Max (Space Grey, 64 GB)
16    APPLE iPhone 11 Pro Max (Midnight Green, 256 GB)
15                APPLE iPhone 11 Pro Max (Gold, 64 GB)
14               APPLE iPhone 11 Pro Max (Gold, 256 GB)
0                     APPLE iPhone 8 Plus (Gold, 64 GB)
29                      APPLE iPhone 12 (White, 128 GB)
32           APPLE iPhone 12 Pro Max (Graphite, 128 GB)
35                      APPLE iPhone 12 (Black, 128 GB)
36                       APPLE iPhone 12 (Blue, 128 GB)
Name: Product Name, dtype: object
```

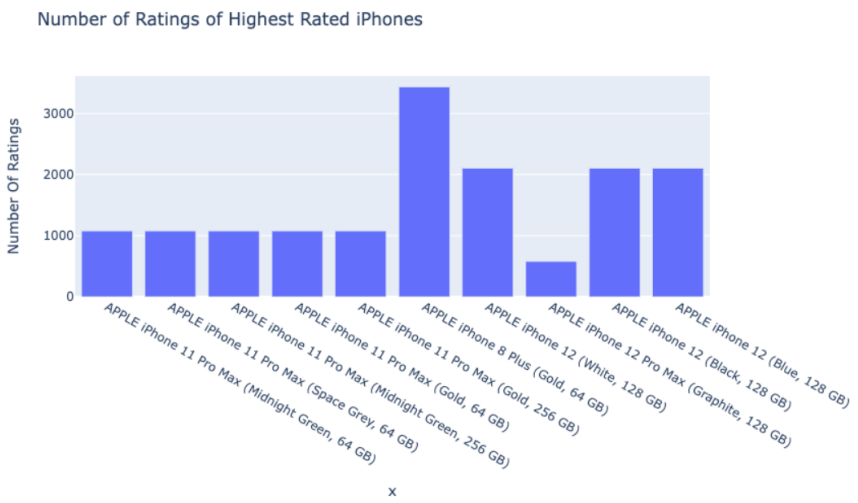
According to the above data, below are the top 5 most liked iPhones in India:

- **APPLE iPhone 11 Pro Max (Midnight Green, 64 GB)**
- **APPLE iPhone 11 Pro Max (Space Grey, 64 GB)**
- **APPLE iPhone 11 Pro Max (Midnight Green, 256 GB)**
- **APPLE iPhone 11 Pro Max (Gold, 64 GB)**
- **APPLE iPhone 11 Pro Max (Gold, 256 GB)**

5. iphones = highest_rated["Product Name"].value_counts()
   label = iphones.index
   counts = highest_rated["Number Of Ratings"]
   figure = px.bar(highest_rated, x=label,
   y = counts,
   title="Number of Ratings of Highest Rated iPhones")
   figure.show()

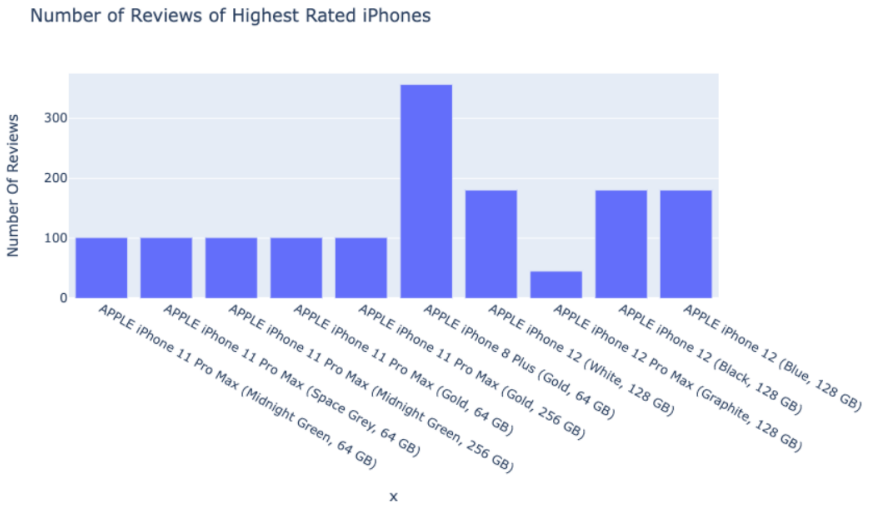
//This code generates a bar plot to display the number of ratings of the highest-rated iPhones.

Output:



6. iphones = highest_rated["Product Name"].value_counts()
   label = iphones.index
   counts = highest_rated["Number Of Reviews"]
   figure = px.bar(highest_rated, x=label, y = counts,
   title="Number of Reviews of Highest Rated iPhones")
   figure.show()
   //This code creates a bar chart using Plotly Express that displays the number of ratings for each iPhone model in the highest_rated dataset.
   The value_counts() method is used to count the number of occurrences of each unique value in the "Product Name" column, and the resulting counts are used as the y-axis values. The x-axis values are the unique product names.

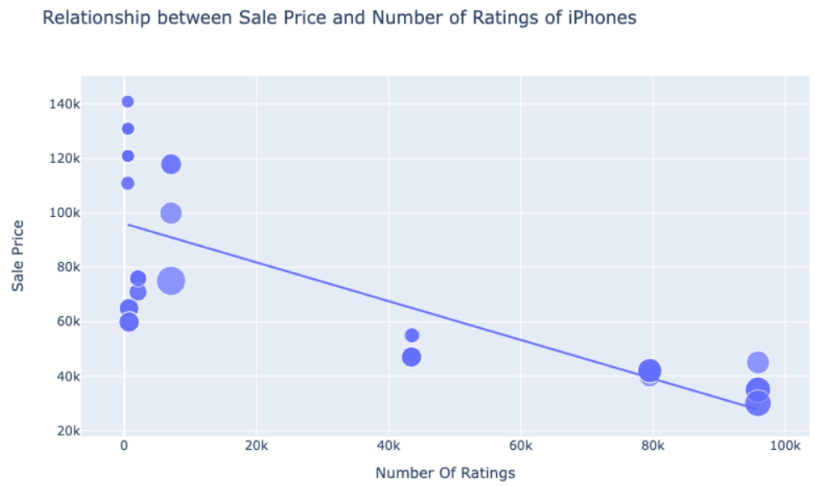
Output:



7. figure = px.scatter(data_frame = data, x="Number Of Ratings",
   y="Sale Price", size="Discount Percentage",
   trendline="ols",
   title="Relationship between Sale Price and Number of Ratings of iPhones")
   figure.show()

   //This code is used to create a scatter plot of the relationship between the sale price and number of ratings of iPhones. The x-axis represents the number of ratings, the y-axis represents the sale price, and the size of each data point represents the discount percentage.

Output:




```
# extract "Product Name" and "Number Of Ratings" columns
product_names = highest_rated["Product Name"]
num_ratings = highest_rated["Number Of Ratings"]

# add clustering for "Discount Percentage" column
X = highest_rated[["Discount Percentage"]]
kmeans = KMeans(n_clusters=3, random_state=0).fit(X)
labels = kmeans.labels_
highest_rated["Cluster Label"] = labels

# create bar plot with color-coded clusters
```

```
fig = px.bar(highest_rated, x=product_names, y=num_ratings, color="Cluster Label",
title="Number of Ratings of Highest Rated iPhones")
fig.show()
```

Output: