```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np
```

```
data=pd.read_csv('user_behavior_dataset.csv')
data.head(10)
```

	User ID	Device Model	Operating System	App Usage Time (min/day)	Screen On Time (hours/day)	Battery Drain (mAh/day)	Number of Apps Installed	(M
0	1	Google Pixel 5	Android	393	6.4	1872	67	
1	2	OnePlus 9	Android	268	4.7	1331	42	
2	3	Xiaomi Mi 11	Android	154	4.0	761	32	
3	4	Google Pixel 5	Android	239	4.8	1676	56	
4	5	iPhone 12	iOS	187	4.3	1367	58	
5	6	Google Pixel 5	Android	99	2.0	940	35	
6	7	Samsung Galaxy S21	Android	350	7.3	1802	66	
7	8	OnePlus 9	Android	543	11.4	2956	82	

Next steps: Generate code with data | View recommended plots | New interactive sheet

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 700 entries, 0 to 699
Data columns (total 11 columns):
 #   Column                      Non-Null Count  Dtype
---  ------                      --------------  -----
 0   User ID                     700 non-null    int64
 1   Device Model                700 non-null    object
 2   Operating System            700 non-null    object
 3   App Usage Time (min/day)    700 non-null    int64
 4   Screen On Time (hours/day)  700 non-null    float64
 5   Battery Drain (mAh/day)     700 non-null    int64
 6   Number of Apps Installed    700 non-null    int64
 7   Data Usage (MB/day)         700 non-null    int64
 8   Age                         700 non-null    int64
 9   Gender                      700 non-null    object
 10  User Behavior Class         700 non-null    int64
dtypes: float64(1), int64(7), object(3)
memory usage: 60.3+ KB
```

```
data.describe()
```

	User ID	App Usage Time (min/day)	Screen On Time (hours/day)	Battery Drain (mAh/day)	Number of Apps Installed	Data Usage (MB/day)
count	700.00000	700.000000	700.000000	700.000000	700.000000	700.000000
mean	350.50000	271.128571	5.272714	1525.158571	50.681429	929.742857
std	202.21688	177.199484	3.068584	819.136414	26.943324	640.451729
min	1.00000	30.000000	1.000000	302.000000	10.000000	102.000000
25%	175.75000	113.250000	2.500000	722.250000	26.000000	373.000000
50%	350.50000	227.500000	4.900000	1502.500000	49.000000	823.500000
75%	525.25000	434.250000	7.400000	2229.500000	74.000000	1341.000000
max	700.00000	598.000000	12.000000	2993.000000	99.000000	2497.000000

```
data.isnull().sum()
```

	0
User ID	0
Device Model	0
Operating System	0
App Usage Time (min/day)	0
Screen On Time (hours/day)	0
Battery Drain (mAh/day)	0
Number of Apps Installed	0
Data Usage (MB/day)	0
Age	0
Gender	0
User Behavior Class	0

```
data.shape
```

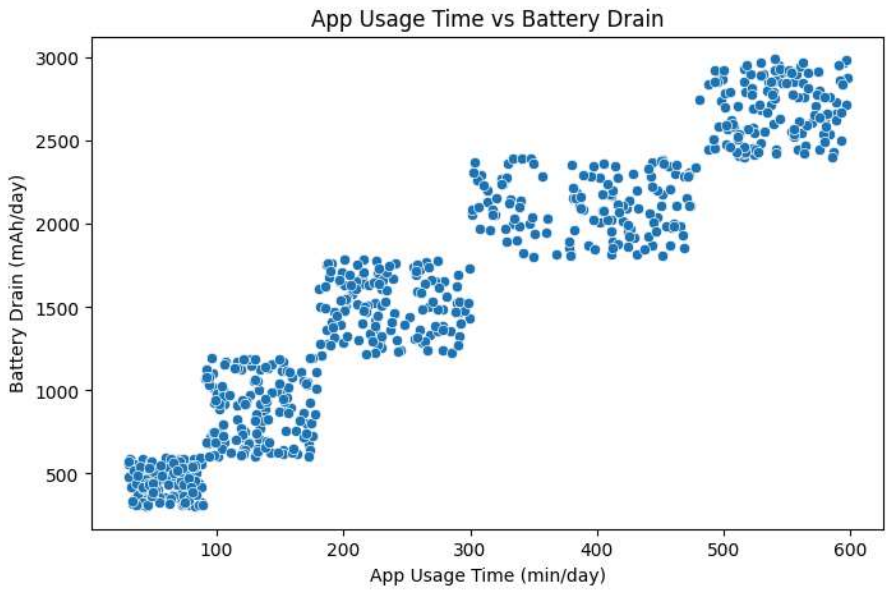
```
(700, 11)
```

Exploratory Data Analysis (EDA) & Visualizations

1 App Usage Time vs. Battery Drain

Insight Goal: Are users with high app usage consuming more battery power?

```
plt.figure(figsize=(8, 5))
sns.scatterplot(x='App Usage Time (min/day)', y='Battery Drain (mAh/day)', data=da
plt.title('App Usage Time vs Battery Drain')
plt.xlabel('App Usage Time (min/day)')
plt.ylabel('Battery Drain (mAh/day)')
plt.show()
```



Hassan Ayaz
12:53 Today

What Does a 0.96 Correlation Mean Between App Usage and Battery Drain?

The more time people spend using apps, the faster their phone's battery drains. This means heavy phone use leads to quicker battery loss.

```
correlation = data['App Usage Time (min/day)'].corr(data['Battery Drain (mAh/day)
print(f'Correlation: {correlation:.2f}')
```

```
Correlation: 0.96
```

2 Average Screen-On Time by Gender

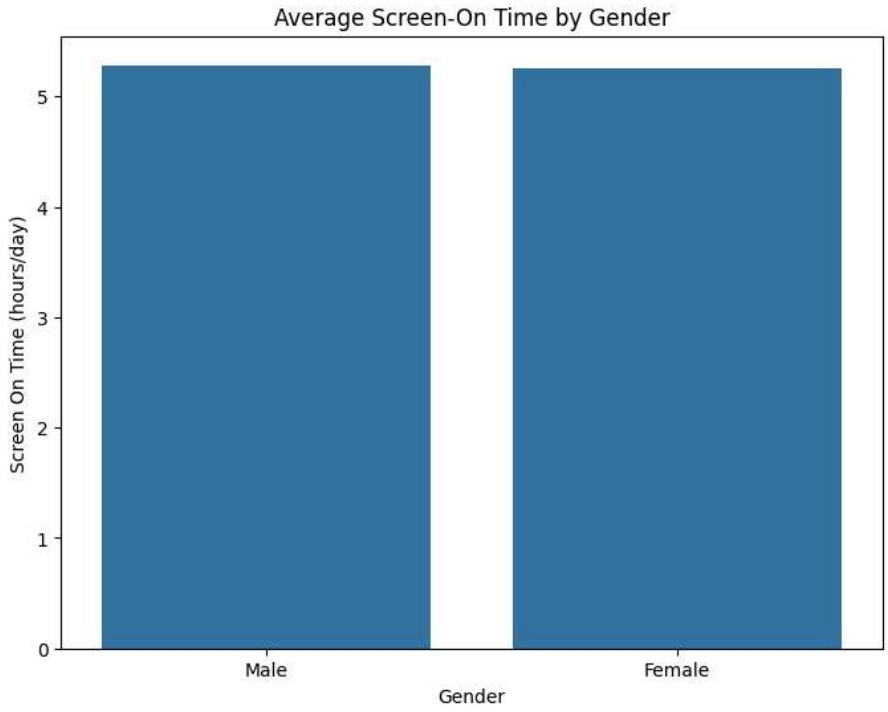
Do male and female users differ in screen-on time?

```
plt.figure(figsize=(8, 6))
sns.barplot(x='Gender', y='Screen On Time (hours/day)', data=data, ci=None)
plt.title('Average Screen-On Time by Gender')
plt.show()
```

```
Correlation: 0.96
<ipython-input-8-faa795c32a9e>:2: FutureWarning:

The `ci` parameter is deprecated. Use `errorbar=None` for the same effect.

  sns.barplot(x='Gender', y='Screen On Time (hours/day)', data=data, ci=None)
```



3. What is the Average Screen-On Time by Age Group?

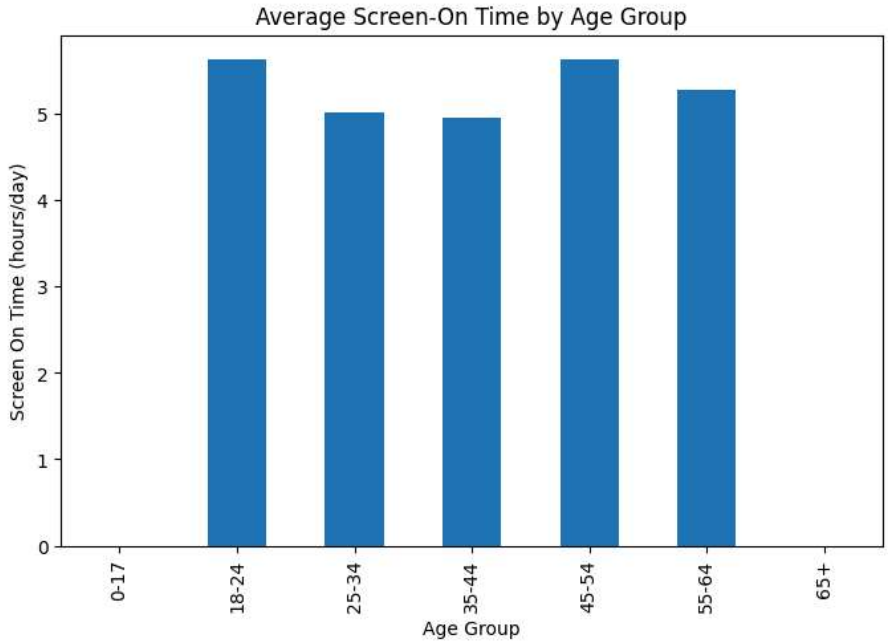
Insight Goal: Are younger users spending more time on their screens?

```
bins = [0, 18, 25, 35, 45, 55, 65, 100]
labels = ['0-17', '18-24', '25-34', '35-44', '45-54', '55-64', '65+']
data['Age Group'] = pd.cut(data['Age'], bins=bins, labels=labels, right=False)


avg_screen_time = data.groupby('Age Group')['Screen On Time (hours/day)'].mean()
print(avg_screen_time)

avg_screen_time.plot(kind='bar',  figsize=(8, 5))
plt.title('Average Screen-On Time by Age Group')
plt.xlabel('Age Group')
plt.ylabel('Screen On Time (hours/day)')
plt.show()
```

```
<ipython-input-12-d938cc4e8e04>:6: FutureWarning: The default of observed=Fals
  avg_screen_time = data.groupby('Age Group')['Screen On Time (hours/day)'].me
Age Group
0-17          NaN
18-24    5.630909
25-34    5.021978
35-44    4.948148
45-54    5.628049
55-64    5.279268
65+           NaN
Name: Screen On Time (hours/day), dtype: float64
```



4. How Does the Number of Installed Apps Relate to Data Usage?

Insight Goal: Does installing more apps lead to higher data consumption?

```
plt.figure(figsize=(8, 5))
sns.scatterplot(x='Number of Apps Installed', y='Data Usage (MB/day)', data=data)
plt.title('Number of Apps Installed vs Data Usage')
plt.xlabel('Number of Apps Installed')
plt.ylabel('Data Usage (MB/day)')
plt.show()
```

H **Hassan Ayaz**
12:52 Today

What Does a 0.93 Correlation Mean Between Number of Apps and Data Usage? The more apps people install, the more data they use. This means having many apps often leads to higher internet consumption, likely because more apps run in the background or need updates.

H **Hassan Ayaz**
13:07 Today

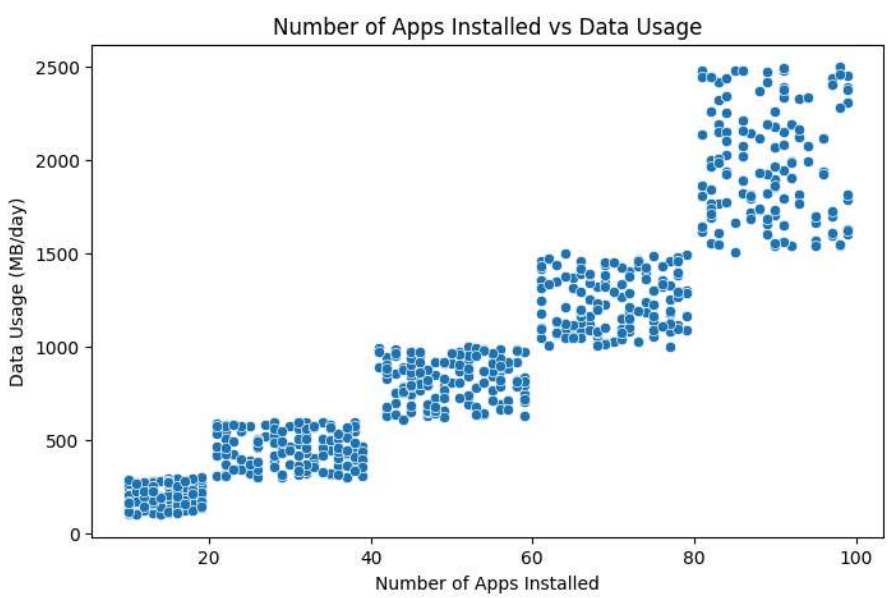
This section analyzes the average battery drain per day for different operating systems.
The average battery drain is calculated in milliampere-hours (mAh/day).

Results:
- Android has an average battery drain of 1508.20 mAh/day.
- iOS has a higher average battery drain of 1589.51 mAh/day.

Interpretation:

The data suggests that iOS devices drain more battery on average than Android devices.

This implies that iOS may struggle more with battery life compared to Android,



requiring more frequent charging under similar usage conditions.

```
correlation_apps_data = data['Number of Apps Installed'].corr(data['Data Usage (MI
print(f'Correlation: {correlation_apps_data:.2f}')
```

```
Correlation: 0.93
```

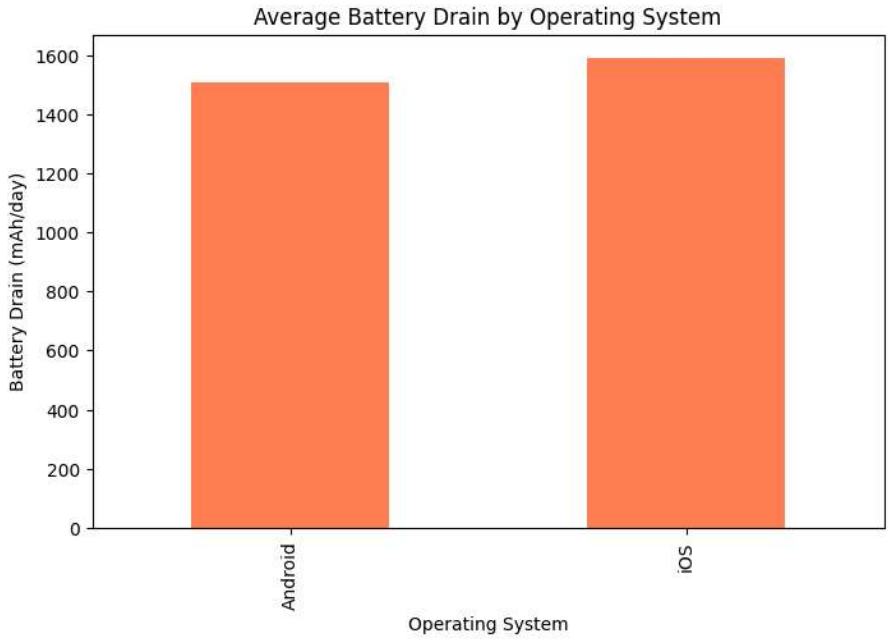
5. Average Battery Drain by Operating System

Insight Goal: Identify which OS struggles more with battery life.

```
avg_battery_drain_os = data.groupby('Operating System')['Battery Drain (mAh/day)']
print(avg_battery_drain_os)
```

```
avg_battery_drain_os.plot(kind='bar', color='coral', figsize=(8, 5))
plt.title('Average Battery Drain by Operating System')
plt.xlabel('Operating System')
plt.ylabel('Battery Drain (mAh/day)')
plt.show()
```

```
Operating System
Android    1508.198556
iOS        1589.513699
Name: Battery Drain (mAh/day), dtype: float64
```



6. How Does Gender Affect App Usage Time?

Insight Goal: Compare app usage time between males and females.

```
avg_app_usage_gender = data.groupby('Gender')['App Usage Time (min/day)'].mean()
print(avg_app_usage_gender)
```

```
avg_app_usage_gender.plot(kind='bar', color=['blue', 'orange'], figsize=(8, 5))
plt.title('Average App Usage Time by Gender')
plt.xlabel('Gender')
plt.ylabel('App Usage Time (min/day)')
plt.xticks(rotation=0)
plt.show()
```