

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import random
import warnings
warnings.filterwarnings("ignore")
```

```
from google.colab import files
uploaded=files.upload()
```

Choose Files No file chosen

Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable.

Saving ONLINE EDUCATION SYSTEM REVIEW.csv to ONLINE EDUCATION SYSTEM REVIEW.csv

```
import io
df2=pd.read_csv(io.BytesIO(uploaded['ONLINE EDUCATION SYSTEM REVIEW.csv']))
```

```
df2.head()
```



	Gender	Home Location	Level of Education	Age(Years)	Number of Subjects	Device type used to attend classes	Economic status	Family size	Internet facilities in your locality
0	Male	Urban	Under Graduate	18	11	Laptop	Middle Class	4	Yes
1	Male	Urban	Under Graduate	19	7	Laptop	Middle Class	4	Yes
2	Male	Rural	Under Graduate	18	5	Laptop	Middle Class	5	Yes
3	Male	Urban	Under Graduate	18	5	Laptop	Middle Class	4	Yes
4	Male	Rural	Under Graduate	18	5	Laptop	Middle Class	4	Yes

5 rows x 23 columns

```
#check the shape of dataset
df2.shape
```

(1033, 23)

```
#check the size of dataset
df2.size
```

23759

```
#check the duplicates of dataset
df2.duplicated().sum()
```

0

```
df2.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1033 entries, 0 to 1032
Data columns (total 23 columns):
#   Column                                     Non-Null Count  Dtype
---  -
0   Gender                                     1033 non-null   object
1   Home Location                             1033 non-null   object
2   Level of Education                       1033 non-null   object
3   Age(Years)                               1033 non-null   int64
4   Number of Subjects                       1033 non-null   int64
5   Device type used to attend classes        1033 non-null   object
6   Economic status                           1033 non-null   object
7   Family size                              1033 non-null   int64
```

```

8  Internet facility in your locality          1033 non-null  int64
9  Are you involved in any sports?            1033 non-null  object
10 Do elderly people monitor you?            1033 non-null  object
11 Study time (Hours)                        1033 non-null  int64
12 Sleep time (Hours)                       1033 non-null  int64
13 Time spent on social media (Hours)        1033 non-null  int64
14 Interested in Gaming?                    1033 non-null  object
15 Have separate room for studying?         1033 non-null  object
16 Engaged in group studies?                1033 non-null  object
17 Average marks scored before pandemic in traditional classroom 1033 non-null  object
18 Your interaction in online mode           1033 non-null  int64
19 Clearing doubts with faculties in online mode 1033 non-null  int64
20 Interested in?                           1033 non-null  object
21 Performance in online                   1033 non-null  int64
22 Your level of satisfaction in Online Education 1033 non-null  object
dtypes: int64(10), object(13)
memory usage: 185.7+ KB

```

```
df2.describe()
```

	Age(Years)	Number of Subjects	Family size	Internet facility in your locality	Study time (Hours)	Sleep time (Hours)	Time spent on social media (Hours)	Your interaction in online mode	Clearin with facu onl
count	1033.000000	1033.00000	1033.000000	1033.000000	1033.000000	1033.000000	1033.000000	1033.000000	1033.000000
mean	19.798645	7.03485	4.413359	3.586641	4.325266	6.947725	2.636980	2.930300	2.930300
std	3.199158	2.81034	1.236750	1.026063	2.134233	1.324039	1.859625	1.105387	1.105387
min	9.000000	1.00000	2.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000
25%	18.000000	6.00000	4.000000	3.000000	3.000000	6.000000	1.000000	2.000000	2.000000
50%	19.000000	7.00000	4.000000	4.000000	4.000000	7.000000	2.000000	3.000000	3.000000
75%	20.000000	8.00000	5.000000	4.000000	6.000000	8.000000	3.000000	4.000000	4.000000
max	40.000000	20.00000	10.000000	5.000000	10.000000	10.000000	10.000000	5.000000	5.000000

```
df2.nunique()
```

```

Gender                2
Home Location         2
Level of Education    3
Age(Years)           24
Number of Subjects   20
Device type used to attend classes 3
Economic status       3
Family size           9
Internet facility in your locality 5
Are you involved in any sports?    2
Do elderly people monitor you?    2
Study time (Hours)    10
Sleep time (Hours)    10
Time spent on social media (Hours) 10
Interested in Gaming?  2
Have separate room for studying?  2
Engaged in group studies?  2
Average marks scored before pandemic in traditional classroom 10
Your interaction in online mode    5
Clearing doubts with faculties in online mode 5
Interested in?              3
Performance in online        9
Your level of satisfaction in Online Education 3
dtype: int64

```

```
df2.isnull().sum()
```

```

Gender                0
Home Location         0
Level of Education    0
Age(Years)           0
Number of Subjects    0
Device type used to attend classes 0
Economic status       0
Family size           0
Internet facility in your locality 0
Are you involved in any sports?    0
Do elderly people monitor you?    0
Study time (Hours)    0
Sleep time (Hours)    0

```

```

Time spent on social media (Hours)      0
Interested in Gaming?                   0
Have separate room for studying?        0
Engaged in group studies?                0
Average marks scored before pandemic in traditional classroom  0
Your interaction in online mode          0
Clearing doubts with faculties in online mode  0
Interested in?                           0
Performance in online                    0
Your level of satisfaction in Online Education  0
dtype: int64

```

```
df2.corr()
```

	Age(Years)	Number of Subjects	Family size	Internet facility in your locality	Study time (Hours)	Sleep time (Hours)	Time spent on social media (Hours)	Your interaction in online mode	Clearin with f in onl
Age(Years)	1.000000	-0.011505	0.016893	0.072625	0.005344	-0.072031	0.005455	0.171943	
Number of Subjects	-0.011505	1.000000	0.035997	-0.025915	0.024280	0.051531	0.004833	-0.005144	.
Family size	0.016893	0.035997	1.000000	-0.020996	0.060614	-0.015787	0.053512	-0.024268	.
Internet facility in your locality	0.072625	-0.025915	-0.020996	1.000000	-0.011554	0.051838	0.082772	0.204391	
Study time (Hours)	0.005344	0.024280	0.060614	-0.011554	1.000000	-0.044385	-0.125986	0.148859	
Sleep time (Hours)	-0.072031	0.051531	-0.015787	0.051838	-0.044385	1.000000	0.144587	0.032598	
Time spent on social media (Hours)	0.005455	0.004833	0.053512	0.082772	-0.125986	0.144587	1.000000	-0.083972	.
Your interaction in online mode	0.171943	-0.005144	-0.024268	0.204391	0.148859	0.032598	-0.083972	1.000000	
Clearing doubts with faculties in online mode	0.125559	-0.008595	-0.019461	0.158990	0.175169	0.043402	-0.101847	0.723216	
Performance in online	0.114965	0.012740	0.001549	0.257339	0.134344	0.065020	-0.086568	0.557507	

```
#removing unimportant coloumns
```

```
df2.drop(["Have separate room for studying?"],axis=1,inplace=True)
df2.columns
```

```

Index(['Gender', 'Home Location', 'Level of Education', 'Age(Years)',
      'Number of Subjects', 'Device type used to attend classes',
      'Economic status', 'Family size', 'Internet facility in your locality',
      'Are you involved in any sports?', 'Do elderly people monitor you?',
      'Study time (Hours)', 'Sleep time (Hours)',
      'Time spent on social media (Hours)', 'Interested in Gaming?',
      'Engaged in group studies?',
      'Average marks scored before pandemic in traditional classroom',
      'Your interaction in online mode',
      'Clearing doubts with faculties in online mode', 'Interested in?',
      'Performance in online',
      'Your level of satisfaction in Online Education'],
      dtype='object')

```

```
df2['Gender'].value_counts()
```

```

Male      614
Female    419
Name: Gender, dtype: int64

```

```
#plotting pie chart
```

```
#Gender impacts
```

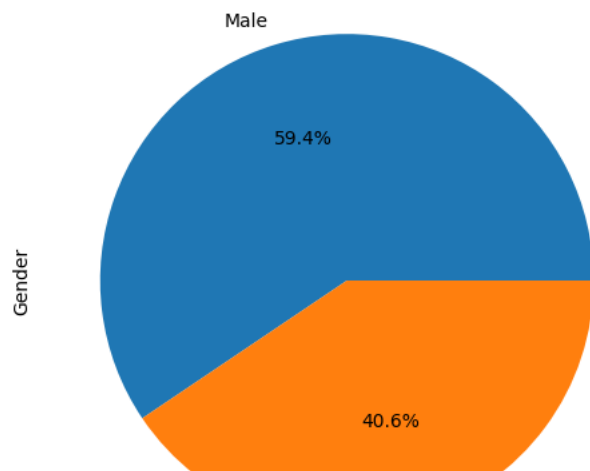
```
plt.figure(figsize=(6,6))
```

```
df2['Gender'].value_counts().plot(kind="pie",autopct='%1.1f%%')
```

```
plt.title('Gender',size=15)
```

```
Text(0.5, 1.0, 'Gender')
```

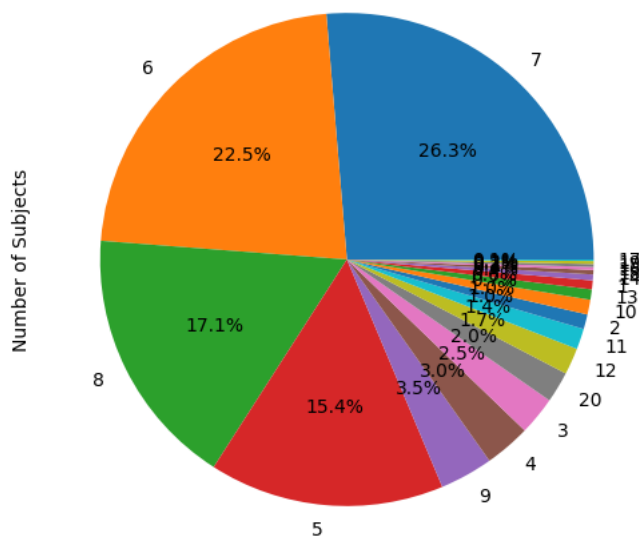
Gender



```
plt.figure(figsize=(6,6))
df2['Number of Subjects'].value_counts().plot(kind="pie", autopct='%1.1f%%')
plt.title('Number of Subjects', size=15)
```

```
Text(0.5, 1.0, 'Number of Subjects')
```

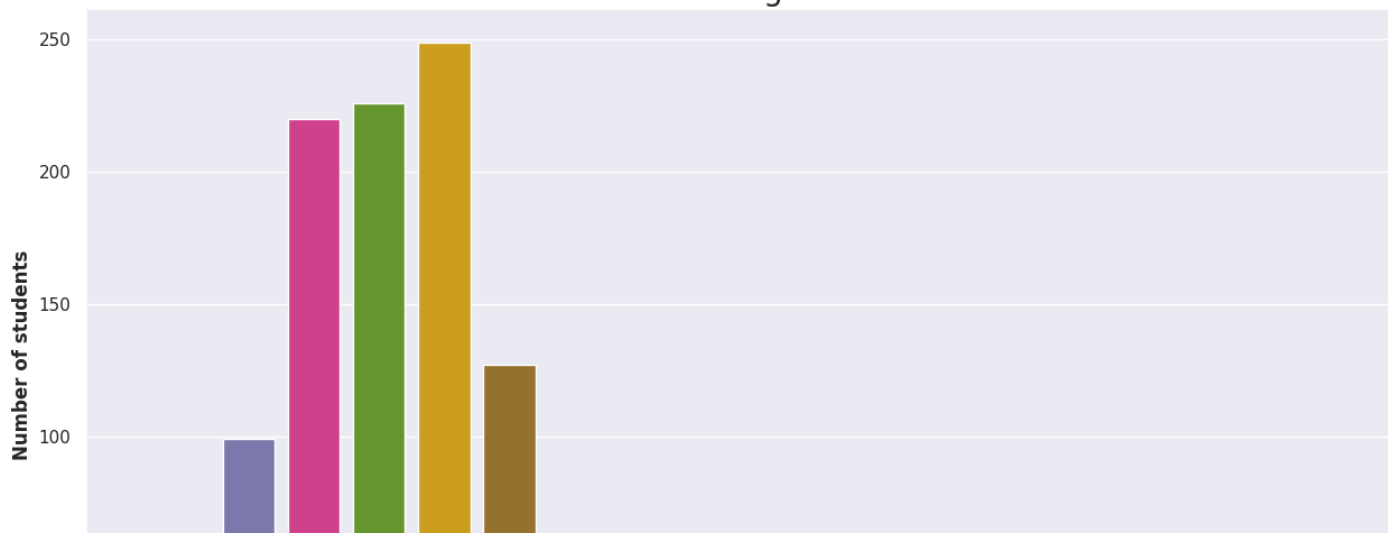
Number of Subjects



```
#age impacts in online
# plotting the count plot
```

```
plt.figure(figsize=(18, 8))
sns.set(style='darkgrid')
plot = sns.countplot(x='Age(Years)', data=df2, palette= 'Dark2') # palette are to control colors or patterns of a columne
plt.title('Age Distribution', size= 20)
plt.xlabel('Age groups', weight='bold')
plt.ylabel('Number of students', weight='bold')
plt.show()
```

Age Distribution



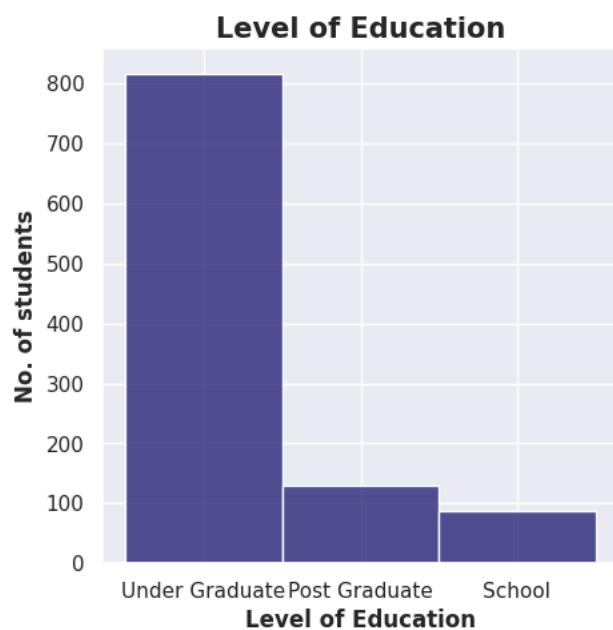
Checking value counts for 'Level of Education' column.

```
df2['Level of Education'].value_counts()
```

```
Under Graduate    817
Post Graduate     129
School            87
Name: Level of Education, dtype: int64
```

plotting the Histplot

```
plt.figure(figsize=(5,5))
sns.histplot(df2, x = df2['Level of Education'], color = 'midnightblue')
plt.title('Level of Education',weight='bold', size = 15)
plt.xlabel('Level of Education',weight='bold')
plt.ylabel('No. of students',weight='bold')
plt.show()
```



#Checking value counts for 'Home Location' column.

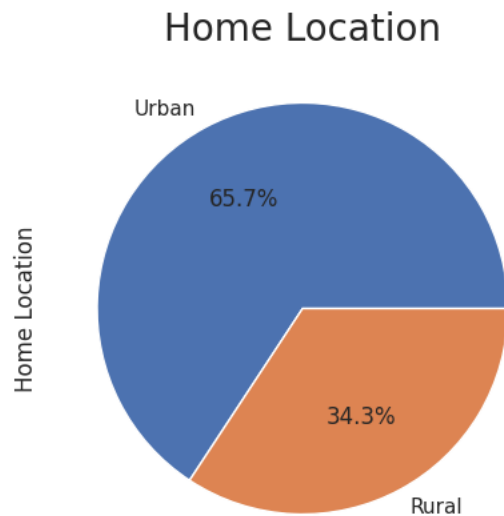
```
df2['Home Location'].value_counts()
```

```
Urban    679
Rural    354
Name: Home Location, dtype: int64
```

```
# plotting the pie chart.
```

```
plt.figure(figsize=(5,5))
df2['Home Location'].value_counts().plot(kind = "pie" ,autopct='%1.1f%%') # autopct display the percent value using string formatting
plt.title('Home Location',size=20)
```

```
Text(0.5, 1.0, 'Home Location')
```



```
# Checking value counts for 'Device type used to attend classes' column.
```

```
df2['Device type used to attend classes'].value_counts()
```

```
Laptop      672
Mobile      334
Desktop       27
Name: Device type used to attend classes, dtype: int64
```

```
# plotting the count plot
```

```
plt.figure(figsize=(10, 6))
sns.set(style='darkgrid')
plot = sns.countplot(x='Device type used to attend classes', data=df2, palette= 'Dark2') # palette are to control colors or patterns of a color
plt.title('Device type', size= 20)
plt.xlabel('Device type used to attend classes', weight='bold')
plt.ylabel('Number of students', weight='bold')
plt.show()
```

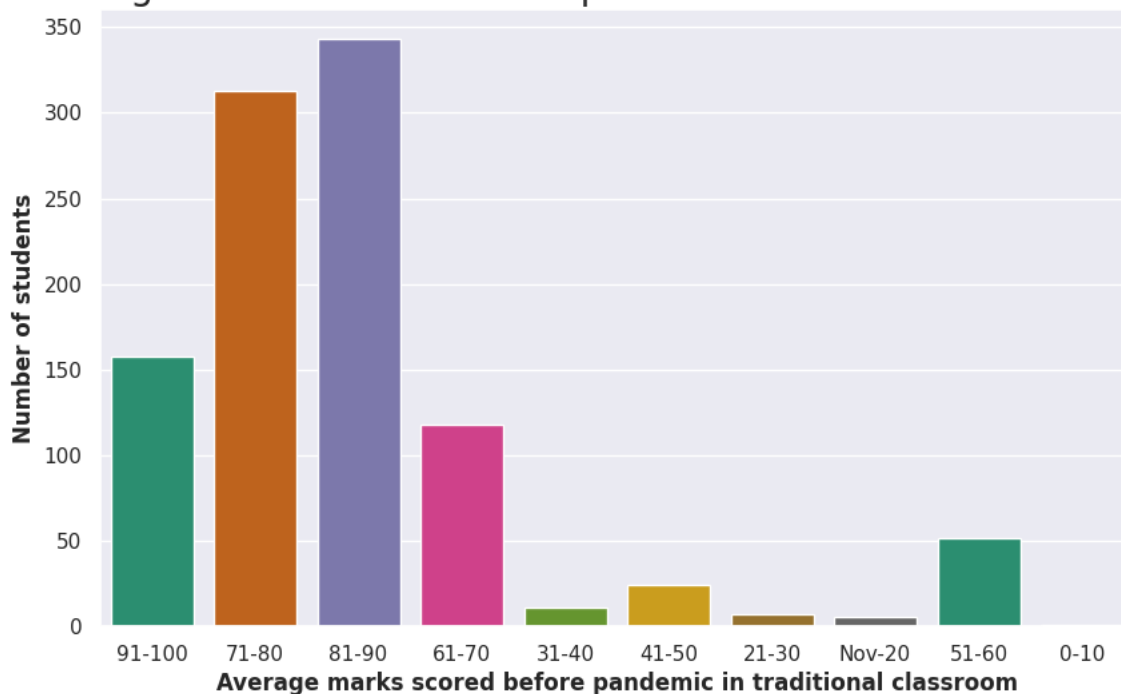
Device type

700

```
# plotting the count plot
```

```
plt.figure(figsize=(10, 6))
sns.set(style='darkgrid')
plot = sns.countplot(x='Average marks scored before pandemic in traditional classroom', data=df2, palette= 'Dark2') # palette are to control
plt.title('Average marks scored before pandemic in traditional classroom', size= 20)
plt.xlabel('Average marks scored before pandemic in traditional classroom', weight='bold')
plt.ylabel('Number of students', weight='bold')
plt.show()
```

Average marks scored before pandemic in traditional classroom



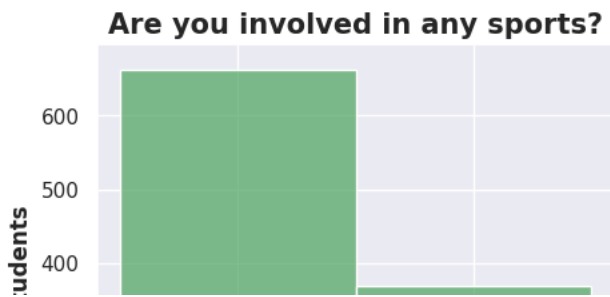
```
# Checking value counts for 'Are you involved in any sports?' column.
```

```
df2['Are you involved in any sports?'].value_counts()
```

```
No      663
Yes     370
Name: Are you involved in any sports?, dtype: int64
```

```
# plotting the histplot
```

```
plt.figure(figsize=(5,5))
sns.histplot(df2, x = df2['Are you involved in any sports?'],color='g')
plt.title('Are you involved in any sports?',weight='bold', size = 15)
plt.xlabel('Are you involved in any sports?',weight='bold')
plt.ylabel('No. of students',weight='bold')
plt.show()
```



```
# Checking value counts for 'Are you involved in any sports?' column.
```

```
df2['Your level of satisfaction in Online Education'].value_counts()
```

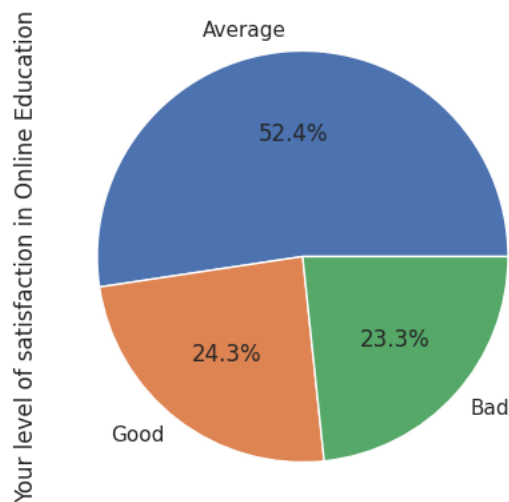
```
Average    541
Good        251
Bad         241
Name: Your level of satisfaction in Online Education, dtype: int64
```

```
# plotting the pie chart.
```

```
plt.figure(figsize=(5,5))
df2['Your level of satisfaction in Online Education'].value_counts().plot(kind = "pie" ,autopct='%1.1f%%') # autopct display the percent value
plt.title('Your level of satisfaction in Online Education',size=20)
```

```
Text(0.5, 1.0, 'Your level of satisfaction in Online Education')
```

Your level of satisfaction in Online Education

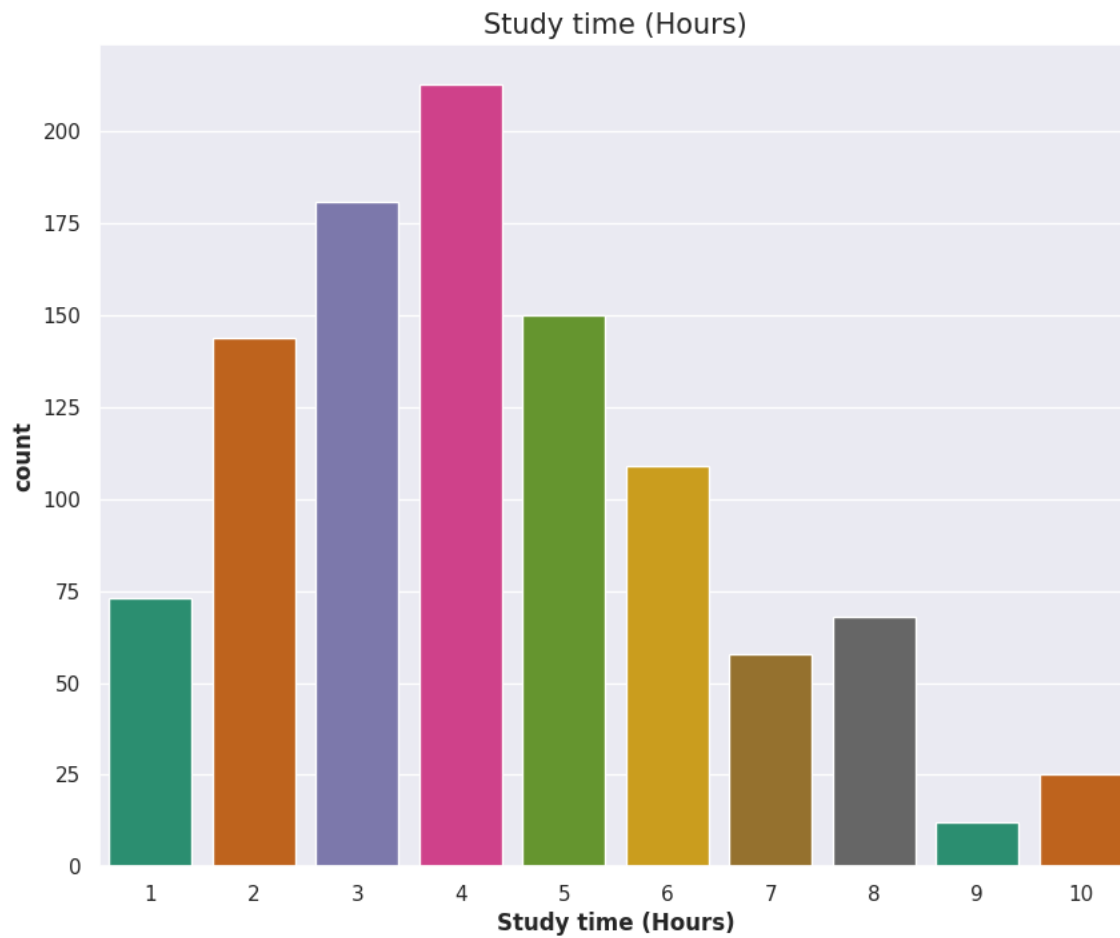


```
df2['Study time (Hours)'].value_counts()
```

```
4    213
3    181
5    150
2    144
6    109
1     73
8     68
7     58
10    25
9     12
Name: Study time (Hours), dtype: int64
```

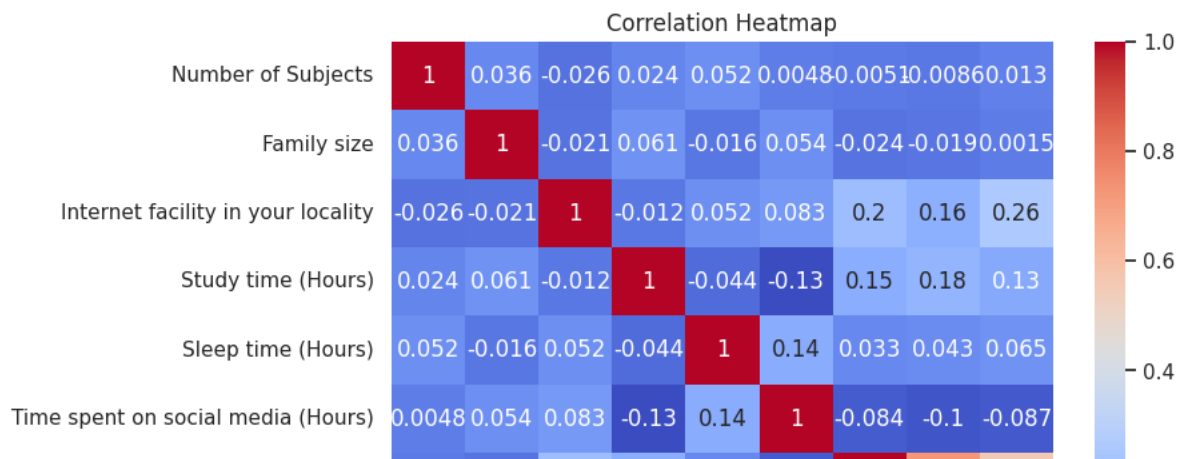
```
# plotting the count plot
```

```
plt.figure(figsize=(10, 8))
sns.set(style='darkgrid')
plot = sns.countplot(x='Study time (Hours)', data=df2, palette= 'Dark2') # palette are to control colors or patterns of a column
plt.title('Study time (Hours)', size= 15)
plt.xlabel('Study time (Hours)', weight='bold')
plt.ylabel('count', weight='bold')
plt.show()
```

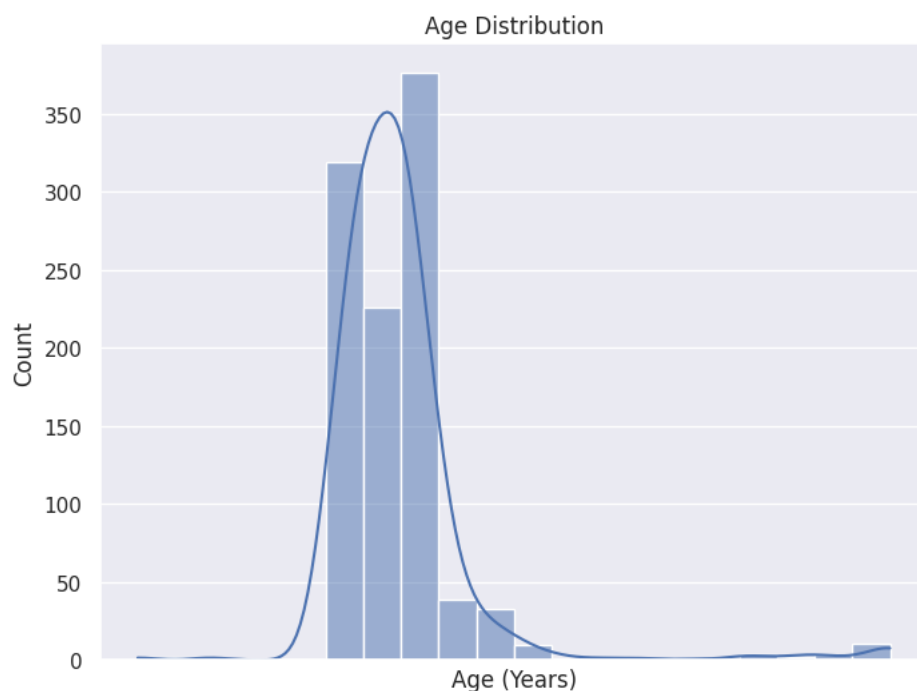



```
import seaborn as sns

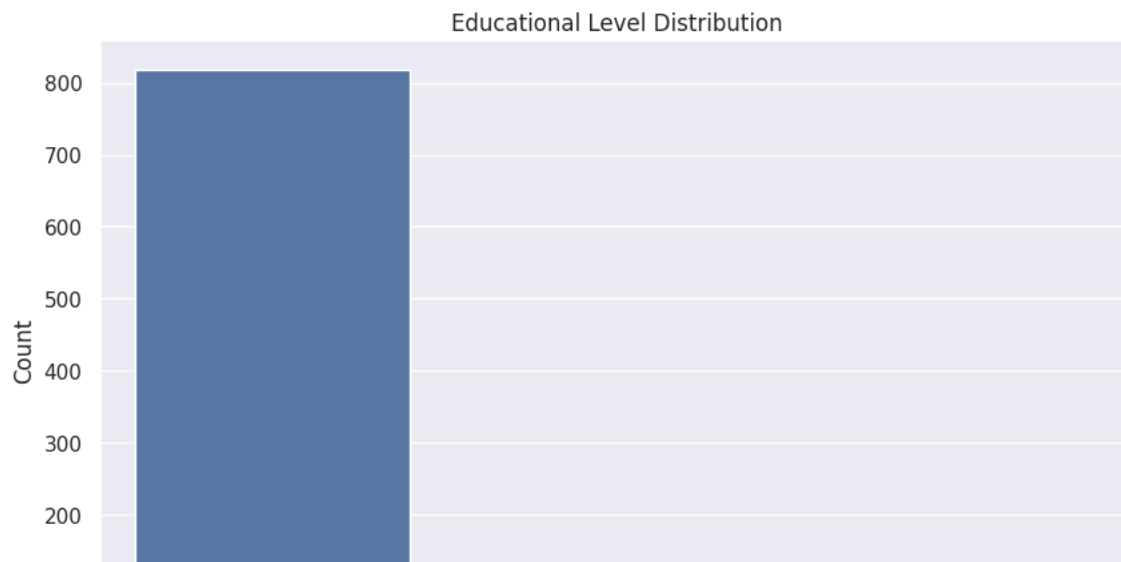
# Assuming 'df2' is your DataFrame
correlation_matrix = df2.corr()
plt.figure(figsize=(8, 6))
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm')
plt.title('Correlation Heatmap')
plt.show()
```



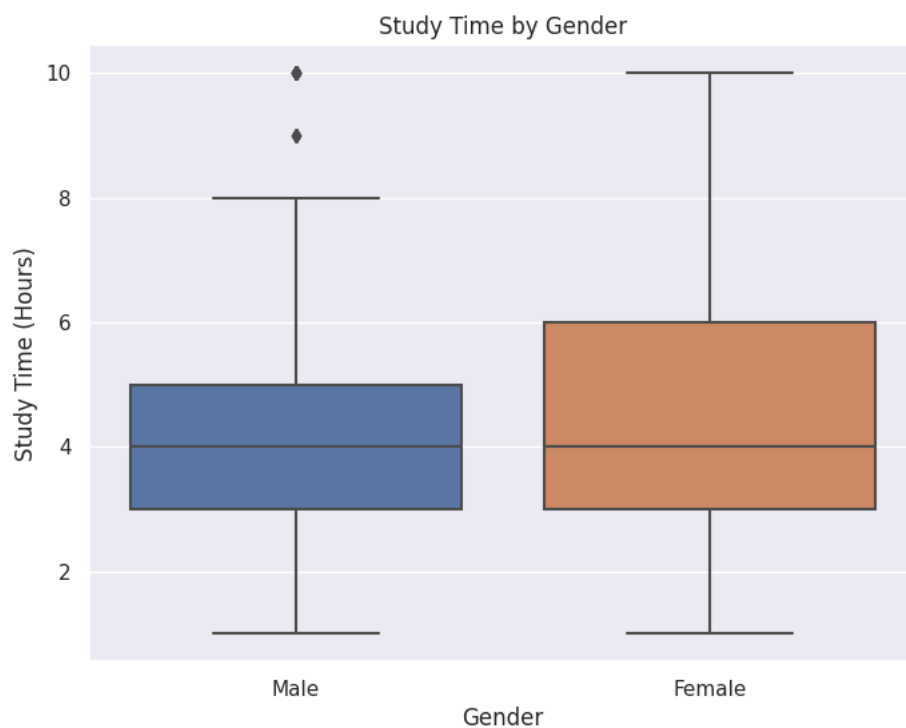
```
plt.figure(figsize=(8, 6))
sns.histplot(data=df2, x="Age(Years)", bins=20, kde=True)
plt.title("Age Distribution")
plt.xlabel("Age (Years)")
plt.ylabel("Count")
plt.show()
```



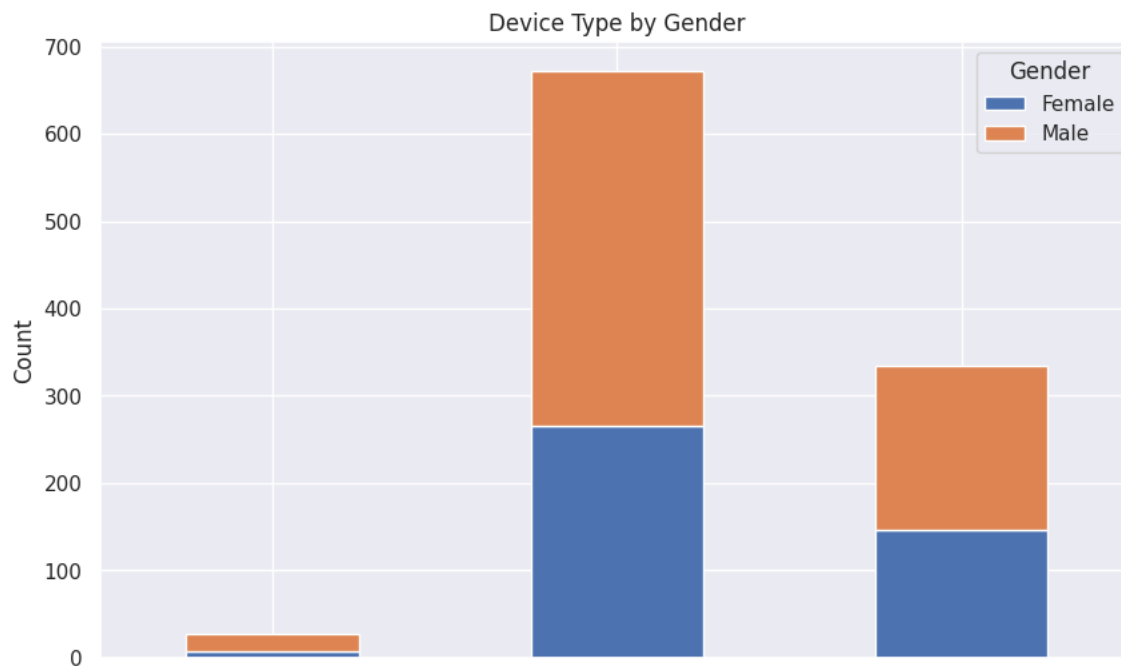
```
plt.figure(figsize=(10, 6))
sns.countplot(data=df2, x="Level of Education", order=df2["Level of Education"].value_counts().index)
plt.title("Educational Level Distribution")
plt.xlabel("Level of Education")
plt.ylabel("Count")
plt.xticks(rotation=45)
plt.show()
```



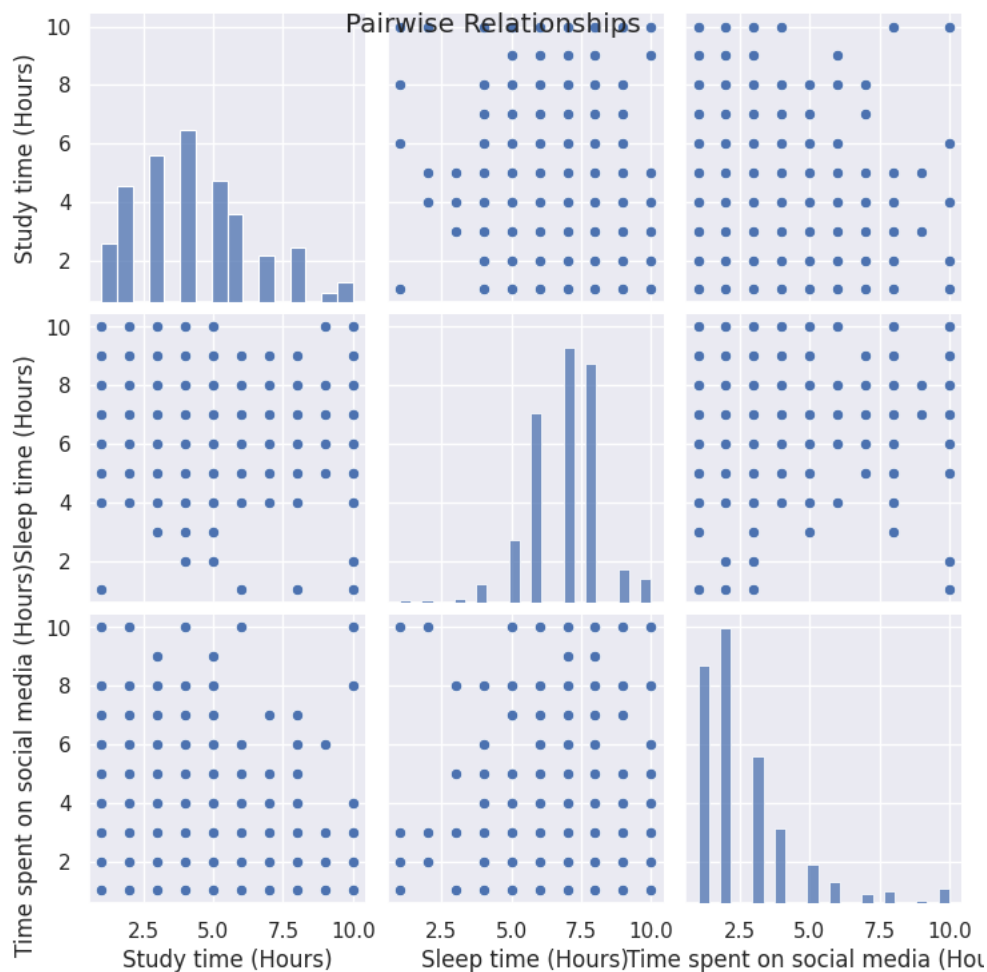
```
plt.figure(figsize=(8, 6))
sns.boxplot(data=df2, x="Gender", y="Study time (Hours)")
plt.title("Study Time by Gender")
plt.xlabel("Gender")
plt.ylabel("Study Time (Hours)")
plt.show()
```



```
device_gender_counts = df2.groupby(['Device type used to attend classes', 'Gender']).size().unstack()
device_gender_counts.plot(kind='bar', stacked=True, figsize=(10, 6))
plt.title("Device Type by Gender")
plt.xlabel("Device Type")
plt.ylabel("Count")
plt.xticks(rotation=45)
plt.legend(title='Gender')
plt.show()
```

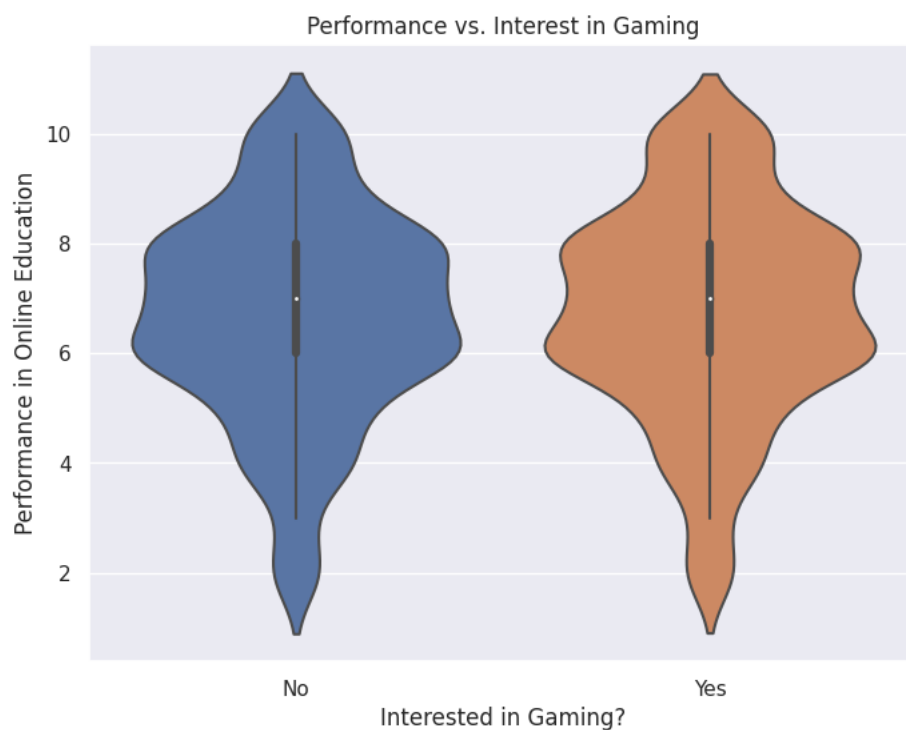


```
sns.pairplot(df2[['Study time (Hours)', 'Sleep time (Hours)', 'Time spent on social media (Hours)']])
plt.suptitle("Pairwise Relationships")
plt.show()
```

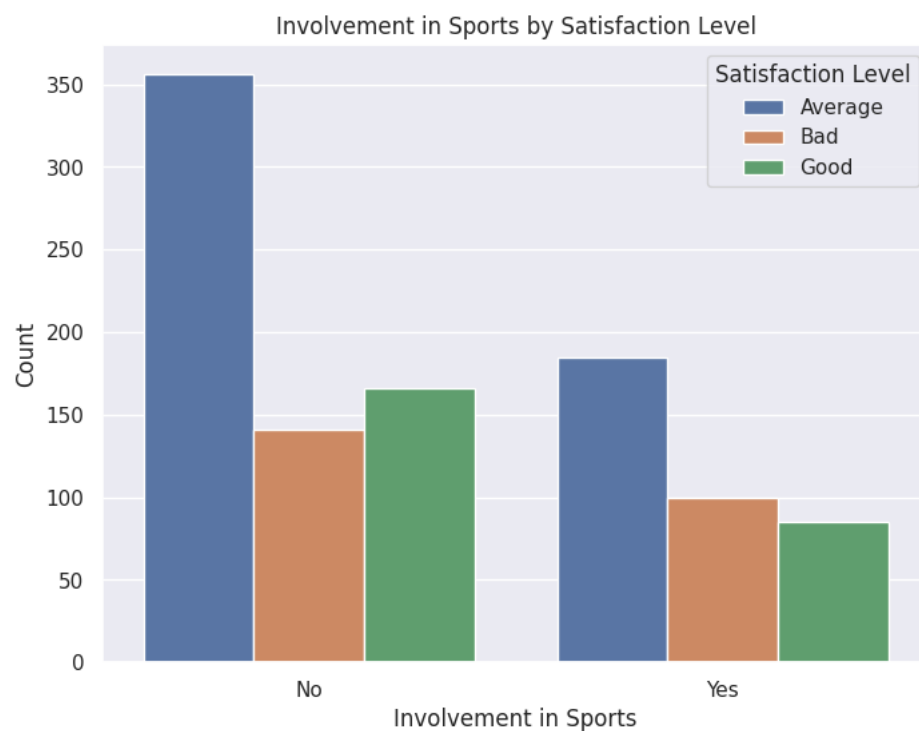


```
plt.figure(figsize=(8, 6))
sns.violinplot(data=df2, x="Interested in Gaming?", y="Performance in online")
plt.title("Performance vs. Interest in Gaming")
plt.xlabel("Interested in Gaming?")
```

```
plt.ylabel("Performance in Online Education")
plt.show()
```

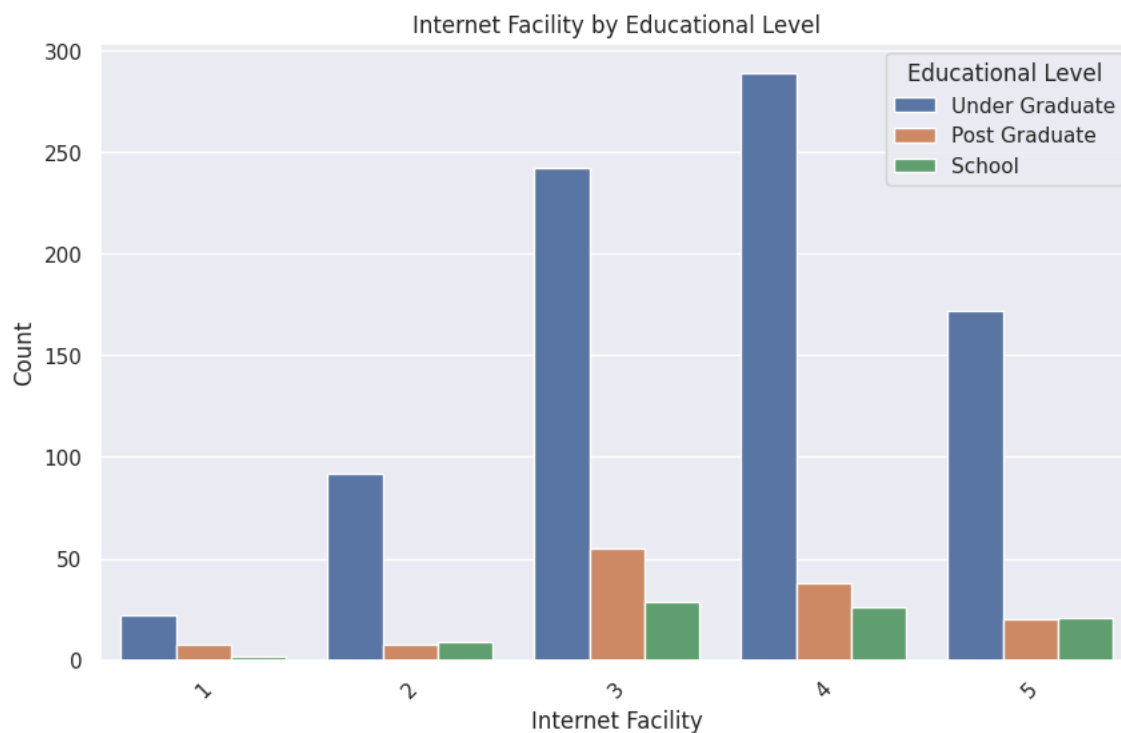


```
plt.figure(figsize=(8, 6))
sns.countplot(data=df2, x="Are you involved in any sports?", hue="Your level of satisfaction in Online Education")
plt.title("Involvement in Sports by Satisfaction Level")
plt.xlabel("Involvement in Sports")
plt.ylabel("Count")
plt.legend(title="Satisfaction Level")
plt.show()
```



```
plt.figure(figsize=(10, 6))
sns.countplot(data=df2, x="Internet facility in your locality", hue="Level of Education")
```

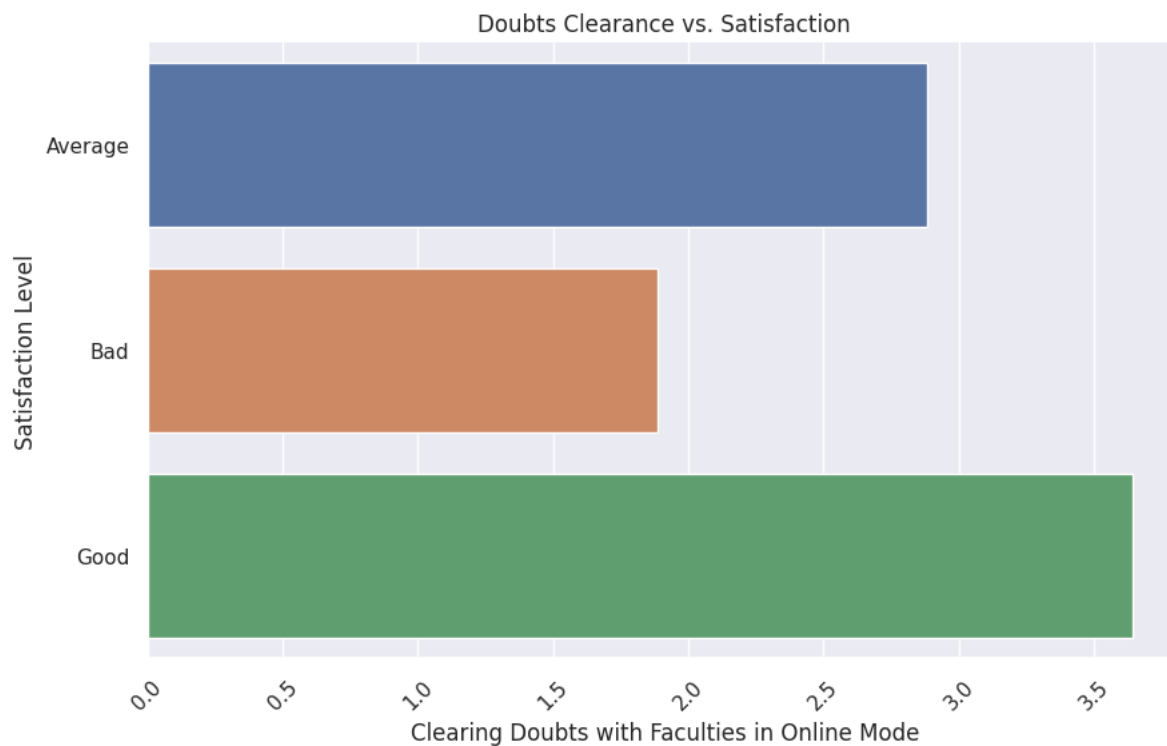
```
plt.title("Internet Facility by Educational Level")
plt.xlabel("Internet Facility")
plt.ylabel("Count")
plt.legend(title="Educational Level")
plt.xticks(rotation=45)
plt.show()
```



```
plt.figure(figsize=(8, 6))
sns.countplot(data=df2, x="Engaged in group studies?", hue="Performance in online")
plt.title("Group Studies vs. Performance")
plt.xlabel("Engaged in Group Studies")
plt.ylabel("Count")
plt.legend(title="Performance")
plt.xticks(rotation=45)
plt.show()
```

Group Studies vs. Performance

```
plt.figure(figsize=(10, 6))
sns.barplot(data=df2, x="Clearing doubts with faculties in online mode", y="Your level of satisfaction in Online Education", ci=None)
plt.title("Doubts Clearance vs. Satisfaction")
plt.xlabel("Clearing Doubts with Faculties in Online Mode")
plt.ylabel("Satisfaction Level")
plt.xticks(rotation=45)
plt.show()
```



```
plt.figure(figsize=(10, 6))
sns.barplot(data=df2, x="Clearing doubts with faculties in online mode", y="Your level of satisfaction in Online Education", ci=None)
plt.title("Doubts Clearance vs. Satisfaction")
plt.xlabel("Clearing Doubts with Faculties in Online Mode")
plt.ylabel("Satisfaction Level")
plt.xticks(rotation=45)
plt.show()
```

Doubts Clearance vs. Satisfaction

Average

```

import pandas as pd
from sklearn.tree import DecisionTreeClassifier
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score
from sklearn.tree import export_text

# Load your dataset
# Assuming you have a DataFrame named 'df2'

# Select the features and target variable
features = df2.drop("Your level of satisfaction in Online Education", axis=1)
target = df2["Your level of satisfaction in Online Education"]

# Convert categorical variables to numerical using one-hot encoding
features_encoded = pd.get_dummies(features)

# Split the dataset into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(features_encoded, target, test_size=0.2, random_state=42)

# Create a decision tree classifier
clf = DecisionTreeClassifier()

# Fit the classifier to the training data
clf.fit(X_train, y_train)

# Make predictions on the test data
y_pred = clf.predict(X_test)

# Evaluate the model
accuracy = accuracy_score(y_test, y_pred)
print(f"Accuracy: {accuracy}")

# Print the decision tree rules
tree_rules = export_text(clf, feature_names=list(features_encoded.columns))
print("Decision Tree Rules:")
print(tree_rules)

```

```

Accuracy: 0.5700483091787439
Decision Tree Rules:
|--- Clearing doubts with faculties in online mode <= 2.50
|   |--- Performance in online <= 2.50
|   |   |--- Sleep time (Hours) <= 3.00
|   |   |   |--- class: Good
|   |   |--- Sleep time (Hours) > 3.00
|   |   |   |--- Level of Education_Post Graduate <= 0.50
|   |   |   |   |--- class: Bad
|   |   |   |--- Level of Education_Post Graduate > 0.50
|   |   |   |   |--- class: Good
|   |--- Performance in online > 2.50
|   |   |--- Device type used to attend classes_Laptop <= 0.50
|   |   |   |--- Clearing doubts with faculties in online mode <= 1.50
|   |   |   |   |--- Home Location_Urban <= 0.50
|   |   |   |   |   |--- Level of Education_Post Graduate <= 0.50
|   |   |   |   |   |   |--- class: Average
|   |   |   |   |   |--- Level of Education_Post Graduate > 0.50
|   |   |   |   |   |   |--- class: Bad
|   |   |   |   |--- Home Location_Urban > 0.50
|   |   |   |   |   |--- Family size <= 6.00
|   |   |   |   |   |   |--- Average marks scored before pandemic in traditional classroom_71-80 <= 0.50
|   |   |   |   |   |   |   |--- Average marks scored before pandemic in traditional classroom_61-70 <= 0.50
|   |   |   |   |   |   |   |   |--- Device type used to attend classes_Mobile <= 0.50
|   |   |   |   |   |   |   |   |   |--- Gender_Male <= 0.50
|   |   |   |   |   |   |   |   |   |   |--- class: Average
|   |   |   |   |   |   |   |   |   |--- Gender_Male > 0.50
|   |   |   |   |   |   |   |   |   |   |--- class: Bad
|   |   |   |   |   |   |   |--- Device type used to attend classes_Mobile > 0.50
|   |   |   |   |   |   |   |   |--- class: Bad
|   |   |   |   |   |--- Average marks scored before pandemic in traditional classroom_61-70 > 0.50
|   |   |   |   |   |   |--- class: Average
|   |   |   |   |--- Average marks scored before pandemic in traditional classroom_71-80 > 0.50
|   |   |   |   |   |--- Your interaction in online mode <= 1.50
|   |   |   |   |   |   |--- class: Bad

```

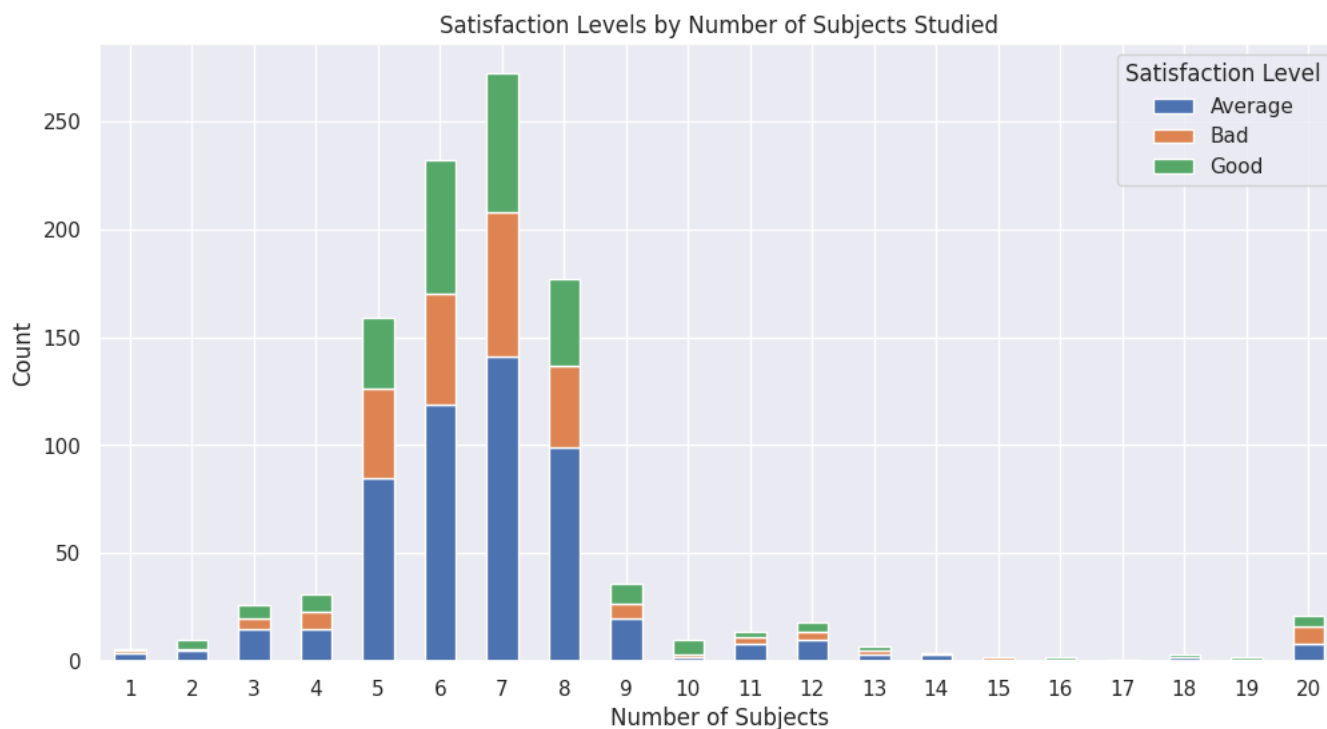

[illegible]

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

# Load your dataset
# Assuming you have a DataFrame named 'df2'

# Create a crosstab to count satisfaction levels by the number of subjects
cross_tab = pd.crosstab(df2['Number of Subjects'], df2['Your level of satisfaction in Online Education'])

# Plot a stacked bar chart
cross_tab.plot(kind='bar', stacked=True, figsize=(12, 6))
plt.title("Satisfaction Levels by Number of Subjects Studied")
plt.xlabel("Number of Subjects")
plt.ylabel("Count")
plt.legend(title="Satisfaction Level")
plt.xticks(rotation=0)
plt.show()
```



```
import pandas as pd
import matplotlib.pyplot as plt

# Load your dataset
```

```
# Assuming you have a DataFrame named 'df2'

# Select relevant attributes for the radar chart
attributes = ["Study time (Hours)", "Performance in online", "Your interaction in online mode", "Sleep time (Hours)"]
data = df2[attributes]

# Normalize the data (optional but recommended)
min_val = data.min()
max_val = data.max()
data = (data - min_val) / (max_val - min_val)

# Create a radar chart
fig, ax = plt.subplots(subplot_kw={'polar': True}, figsize=(8, 8))

# Number of attributes
categories = list(data.columns)
N = len(categories)

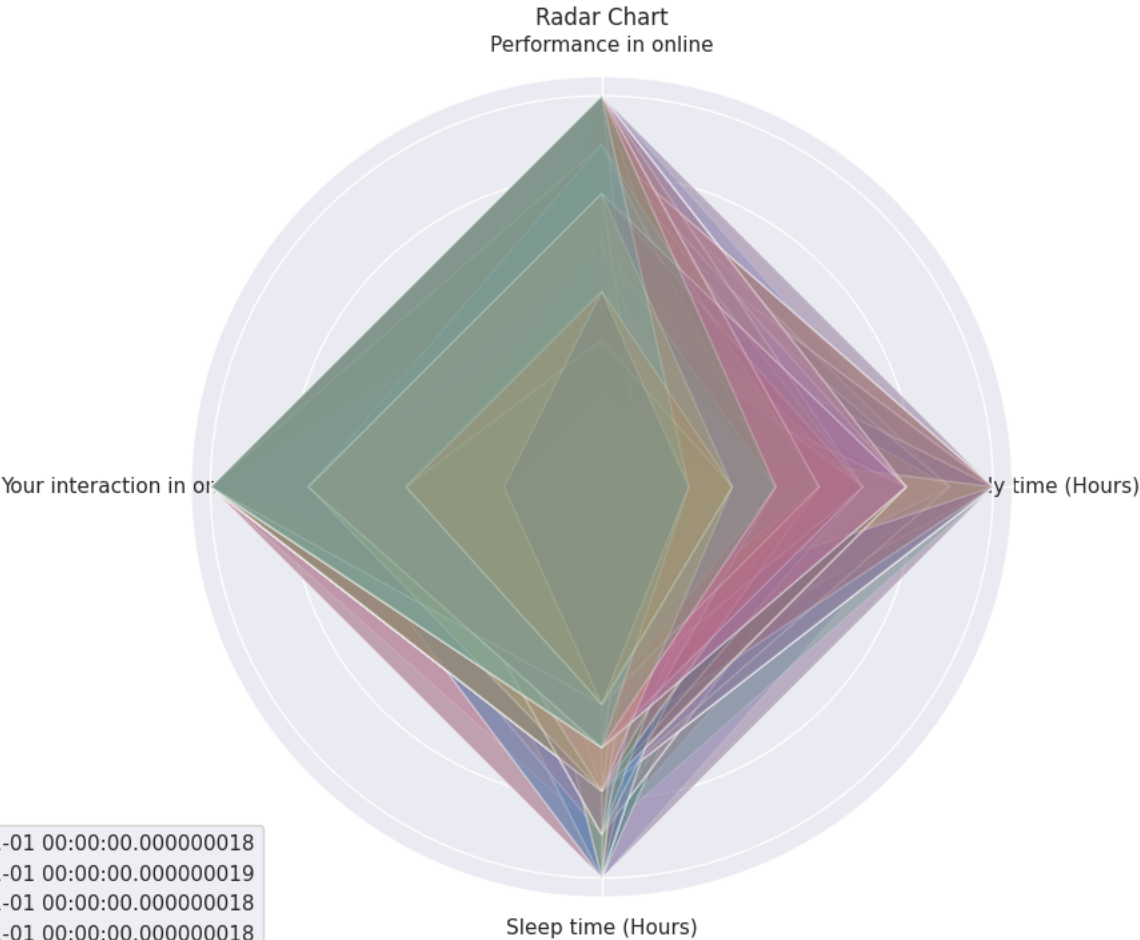
# Calculate angle values
angles = [n / float(N) * 2 * 3.14159265359 for n in range(N)]
angles += angles[:1]

# Plot the data
for i, row in data.iterrows():
    values = row.values.flatten().tolist()
    values += values[:1]
    ax.fill(angles, values, alpha=0.25, label=i)

# Set labels
ax.set_xticks(angles[:-1])
ax.set_xticklabels(categories)
ax.set_yticklabels([])

# Add a legend
ax.legend(loc='upper right', bbox_to_anchor=(0.1, 0.1))

# Show the radar chart
plt.title("Radar Chart")
plt.show()
```



1970-01-01 00:00:00.000000018
1970-01-01 00:00:00.000000019
1970-01-01 00:00:00.000000018
1970-01-01 00:00:00.000000018
1970-01-01 00:00:00.000000018
1970-01-01 00:00:00.000000018
1970-01-01 00:00:00.000000019
1970-01-01 00:00:00.000000017
1970-01-01 00:00:00.000000019
1970-01-01 00:00:00.000000020
1970-01-01 00:00:00.000000018
1970-01-01 00:00:00.000000025
1970-01-01 00:00:00.000000020
1970-01-01 00:00:00.000000017
1970-01-01 00:00:00.000000021
1970-01-01 00:00:00.000000020
1970-01-01 00:00:00.000000020
1970-01-01 00:00:00.000000023
1970-01-01 00:00:00.000000018
1970-01-01 00:00:00.000000019
1970-01-01 00:00:00.000000024
1970-01-01 00:00:00.000000022
1970-01-01 00:00:00.000000019
1970-01-01 00:00:00.000000024
1970-01-01 00:00:00.000000018
1970-01-01 00:00:00.000000021
1970-01-01 00:00:00.000000019
1970-01-01 00:00:00.000000021
1970-01-01 00:00:00.000000020
1970-01-01 00:00:00.000000017
1970-01-01 00:00:00.000000022
1970-01-01 00:00:00.000000018
1970-01-01 00:00:00.000000021
1970-01-01 00:00:00.000000019
1970-01-01 00:00:00.000000018
1970-01-01 00:00:00.000000020
1970-01-01 00:00:00.000000018