assignment1

February 18, 2026

[18]:
```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
```

[35]:
```
df = pd.read_csv("Students Social Media Addiction1.csv")
```

[36]:
```
df.head()
```

[36]:

	Student_ID	Age	Gender	Academic_Level	Country	Avg_Daily_Usage_Hours	Most_Used_Platform	Affects_Academic_Performance	Sleep_Hours_Per_Night	Mental_Health_Score	Relationship_Status	Conflicts_Over_Social_Media	Addicted_Score
0	1	19	Female	Undergraduate	Bangladesh	5.2	Instagram	Yes	6.5	6	In Relationship	3	100
1	2	22	Male	Graduate	India	2.1	Twitter	No	7.5	8	Single	0	3
2	3	20	Female	Undergraduate	USA	6.0	TikTok	Yes	5.0	5	Complicated	4	9
3	4	18	Male	High School	UK	3.0	YouTube	No	7.0	7	Single	1	4
4	5	21	Male	Graduate	Canada	4.5	Facebook	Yes	6.0	6	In Relationship	2	7

[37]:
```
df.shape
```

[37]:
```
(705, 13)
```

[38]:
```
df.columns
```

[38]:
```
Index(['Student_ID', 'Age', 'Gender', 'Academic_Level', 'Country',
       'Avg_Daily_Usage_Hours', 'Most_Used_Platform',
       'Affects_Academic_Performance', 'Sleep_Hours_Per_Night',
       'Mental_Health_Score', 'Relationship_Status',
       'Conflicts_Over_Social_Media', 'Addicted_Score'],
      dtype='object')
```

[39]:
```
df.isnull().sum()
```

[39]:
```
Student_ID                      0
Age                             0
Gender                          0
Academic_Level                  0
Country                         0
Avg_Daily_Usage_Hours           0
Most_Used_Platform              0
Affects_Academic_Performance    0
Sleep_Hours_Per_Night           0
Mental_Health_Score             0
Relationship_Status             0
Conflicts_Over_Social_Media     0
Addicted_Score                  0
dtype: int64
```

[40]:
```
df.describe()
```

[40]:

	Student_ID	Age	Avg_Daily_Usage_Hours	Sleep_Hours_Per_Night	Mental_Health_Score	Conflicts_Over_Social_Media	Addicted_Score
count	705.000000	705.000000	705.000000	705.000000	705.000000	705.000000	705.000000
mean	353.000000	20.659574	4.918723	6.868936	6.226950	2.849645	7.260993
std	203.660256	1.399217	1.257395	1.126848	1.105055	0.957968	13.780234
min	1.000000	18.000000	1.500000	3.800000	4.000000	0.000000	2.000000
25%	177.000000	19.000000	4.100000	6.000000			
50%	353.000000	21.000000	4.800000	6.900000			
75%	529.000000	22.000000	5.800000	7.700000			
max	705.000000	24.000000	8.500000	9.600000			

```
25%              5.000000                    2.000000        5.000000
50%              6.000000                    3.000000        7.000000
75%              7.000000                    4.000000        8.000000
max              9.000000                    5.000000      300.000000
```

[41]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 705 entries, 0 to 704
Data columns (total 13 columns):
 #   Column                          Non-Null Count  Dtype
---  ------                          --------------  -----
 0   Student_ID                      705 non-null    int64
 1   Age                             705 non-null    int64
 2   Gender                          705 non-null    object
 3   Academic_Level                  705 non-null    object
 4   Country                         705 non-null    object
 5   Avg_Daily_Usage_Hours           705 non-null    float64
 6   Most_Used_Platform              705 non-null    object
 7   Affects_Academic_Performance    705 non-null    object
 8   Sleep_Hours_Per_Night           705 non-null    float64
 9   Mental_Health_Score             705 non-null    int64
 10  Relationship_Status             705 non-null    object
 11  Conflicts_Over_Social_Media     705 non-null    int64
 12  Addicted_Score                  705 non-null    int64
dtypes: float64(2), int64(5), object(6)
memory usage: 71.7+ KB
```

[42]:
```
df['Age'] = df['Age'].astype(int)
df['Gender'] = df['Gender'].astype('category')
```

[43]:
```
mean = df.mean
```

[44]:
```
mean
```

[44]:
```
<bound method DataFrame.mean of      Student_ID  Age  Gender Academic_Level
Country  \
0             1   19  Female  Undergraduate  Bangladesh
1             2   22    Male       Graduate       India
2             3   20  Female  Undergraduate         USA
3             4   18    Male     High School          UK
4             5   21    Male       Graduate      Canada
..          ...  ...     ...             ...         ...
700         701   20  Female  Undergraduate       Italy
701         702   23    Male       Graduate      Russia
702         703   21  Female  Undergraduate       China
703         704   24    Male       Graduate       Japan
```

```
704         705   19  Female  Undergraduate      Poland

     Avg_Daily_Usage_Hours Most_Used_Platform Affects_Academic_Performance  \
0                      5.2          Instagram                          Yes
1                      2.1            Twitter                           No
2                      6.0             TikTok                          Yes
3                      3.0            YouTube                           No
4                      4.5           Facebook                          Yes
..                     …                 …                          …
700                    4.7             TikTok                           No
701                    6.8          Instagram                          Yes
702                    5.6             WeChat                          Yes
703                    4.3            Twitter                           No
704                    6.2           Facebook                          Yes

     Sleep_Hours_Per_Night  Mental_Health_Score Relationship_Status  \
0                      6.5                    6     In Relationship
1                      7.5                    8              Single
2                      5.0                    5         Complicated
3                      7.0                    7              Single
4                      6.0                    6     In Relationship
..                     …                    …                 …
700                    7.2                    7     In Relationship
701                    5.9                    4              Single
702                    6.7                    6     In Relationship
703                    7.5                    8              Single
704                    6.3                    5              Single

     Conflicts_Over_Social_Media  Addicted_Score
0                              3             100
1                              0               3
2                              4               9
3                              1               4
4                              2               7
..                           …               …
700                            2               5
701                            5               9
702                            3               7
703                            2               4
704                            4               8

[705 rows x 13 columns]>
```

[45]:
```
data = df['Addicted_Score']
```

[46]:
```
print("Mean:", np.mean(data))
print("Std:", np.std(data))
```

```
print("Min:", np.min(data))
print("Max:", np.max(data))
```

```
Mean: 7.260992907801419
Std: 13.770457498694586
Min: 2
Max: 300
```

[47]:
```
mean = data.mean()
std = data.std()
```

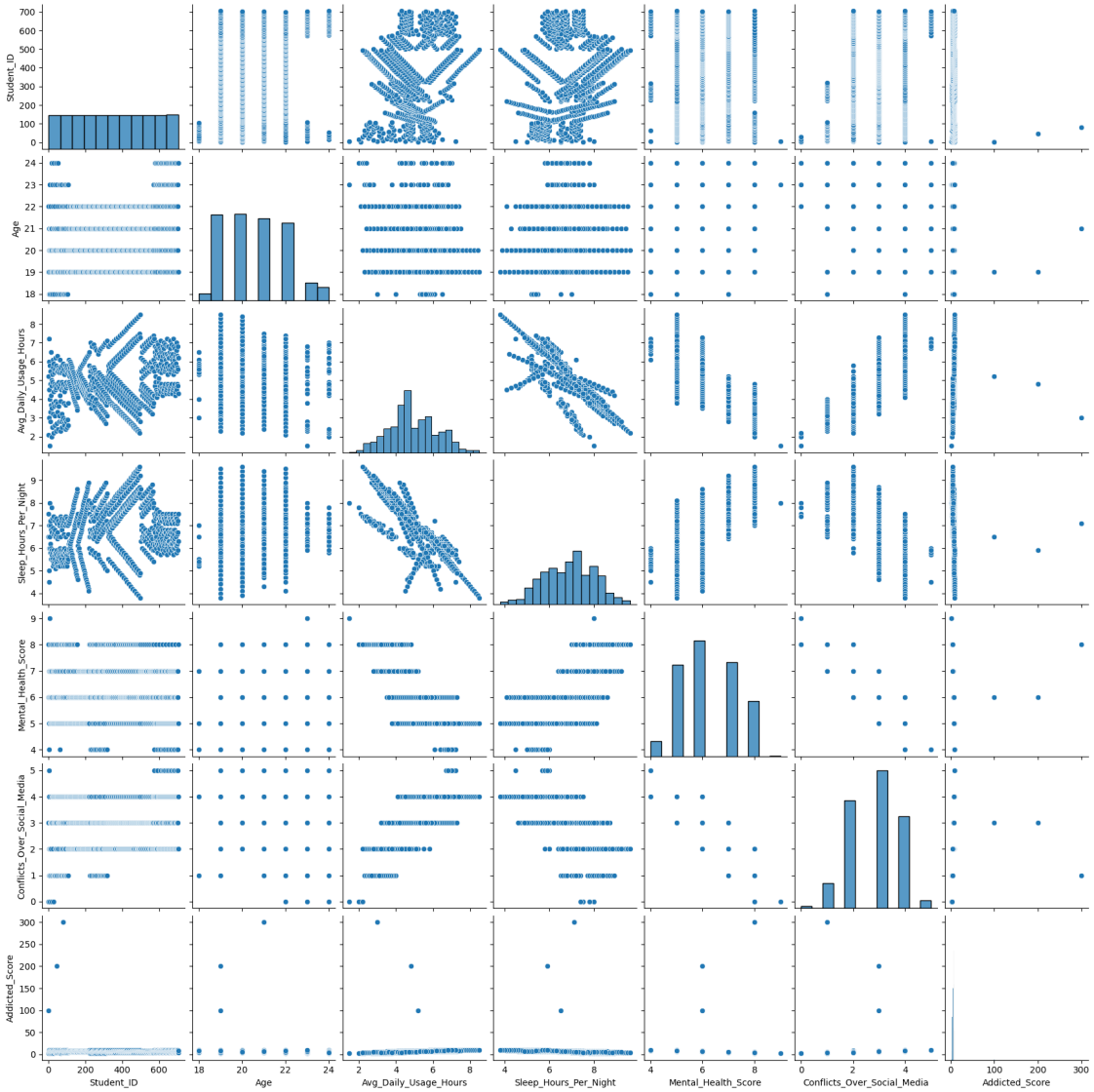
[48]:
```
threshold = 3
outliers = []

for i in data:
    z_score = (i - mean) / std
    if np.abs(z_score) > threshold:
        outliers.append(i)

outliers
```

[48]:
```
[100, 200, 300]
```

[50]:
```
sns.pairplot(df)
plt.show()
```


Student_ID
Age
Avg_Daily_Usage_Hours
Sleep_Hours_Per_Night
Mental_Health_Score
Conflicts_Over_Social_Media
Addicted_Score

[ ]: