

```
1 #importing Lib
2
3 import pandas as pd
4 import seaborn as sns
5 import matplotlib.pyplot as plt
6
7 #Import DataSet
8
9 ds = pd.read_excel("C:/Users/VANSH SHARMA/OneDrive/
    Desktop/DataAnalyst/Intenship/Quantium/
    QVI_transaction_data.xlsx")
10 print(ds.head(10))
11
12 #Summarize Data
13
14 descr = ds.describe()
15 print(descr)
16
17 #Finding for Null Values
18
19 print(ds.isnull().sum())
20
21 #Checking Outliers
22
23 sns.displot(ds.TOT_SALES, kde = True)
24 plt.show()
25
26 #Removing Outliers
27
28 nudata = ds.select_dtypes(['float','int'])
29 print(nudata.head(10))
30
31 fl = nudata[nudata['TOT_SALES'] < 8.000]
32 sns.displot(fl.TOT_SALES , kde = True)
33 plt.show()
34
35 #Checking Data Formats
36
37 print(ds.dtypes)
38
```