

```

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

df=pd.read_csv('Bwali Sales Data.csv',encoding='unicode_escape')

In [3]: df.shape
Out[3]: (11251, 15)

In [4]: df.head()
Out[4]:
   User_ID  Cust_name  Product_ID  Gender  Age Group  Age  Marital_Status  State  Zone  Occupation  Product_Category  Orders  Amount  Status  unnamed1
0  1002903  Sarani  P0012942  F  26-35  28  0  Maharashtra  Western  Healthcare  Auto  1  23952.0  NaN  NaN
1  100732  Karik  P0011942  F  26-35  35  1  Andhra Pradesh  Southern  Govt  Auto  3  23934.0  NaN  NaN
2  1001990  Bindu  P0018542  F  26-35  35  1  Uttar Pradesh  Central  Automobile  Auto  3  23924.0  NaN  NaN
3  1001425  Sudevi  P0023742  M  0-17  16  0  Karnataka  Southern  Construction  Auto  2  23912.0  NaN  NaN
4  1002689  Joni  P0025742  M  26-35  28  1  Gujarat  Western  Food Processing  Auto  2  23877.0  NaN  NaN

In [5]: df.head(50)
Out[5]:
   User_ID  Cust_name  Product_ID  Gender  Age Group  Age  Marital_Status  State  Zone  Occupation  Product_Category  Orders  Amount  Status  unnamed1
0  1002903  Sarani  P0012942  F  26-35  28  0  Maharashtra  Western  Healthcare  Auto  1  23952.0  NaN  NaN
1  100732  Karik  P0011942  F  26-35  35  1  Andhra Pradesh  Southern  Govt  Auto  3  23934.0  NaN  NaN
2  1001990  Bindu  P0018542  F  26-35  35  1  Uttar Pradesh  Central  Automobile  Auto  3  23924.0  NaN  NaN
3  1001425  Sudevi  P0023742  M  0-17  16  0  Karnataka  Southern  Construction  Auto  2  23912.0  NaN  NaN
4  1002689  Joni  P0025742  M  26-35  28  1  Gujarat  Western  Food Processing  Auto  2  23877.0  NaN  NaN
5  100588  Jovi  P0025742  M  26-35  28  1  Gujarat  Western  Food Processing  Auto  2  23817.0  NaN  NaN
6  100588  Jovi  P0025742  M  26-35  28  1  Himachal Pradesh  Northern  Food Processing  Auto  1  23877.0  NaN  NaN
7  1001132  Bak  P0010342  F  18-25  25  1  Uttar Pradesh  Central  Lawyer  Auto  4  23841.0  NaN  NaN
8  1002924  Shivang  P0023442  F  55+  61  0  Maharashtra  Western  IT Sector  Auto  1  NaN  NaN  NaN
9  1002224  Kushal  P0020542  M  26-35  35  0  Uttar Pradesh  Central  Govt  Auto  2  23809.0  NaN  NaN
10  100850  Giniy  P0003142  F  26-35  26  1  Andhra Pradesh  Southern  Media  Auto  4  23799.99  NaN  NaN
11  100829  Harsha  P0020042  M  26-35  34  0  Delhi  Central  Banking  Auto  1  23770.0  NaN  NaN
12  1002114  Kargela  P0011942  F  18-25  20  0  Andhra Pradesh  Southern  Retail  Auto  2  23750.0  NaN  NaN
13  1004035  Rishi  P0000042  F  18-25  20  1  Andhra Pradesh  Southern  IT Sector  Auto  2  23730.0  NaN  NaN
14  1001690  Vasudh  P0020442  M  26-35  26  1  Andhra Pradesh  Southern  Automobile  Auto  4  23718.0  NaN  NaN
15  1003858  Cano  P0029742  M  46-50  46  1  Madhya Pradesh  Central  Hospitality  Auto  3  NaN  NaN  NaN
16  1008113  Lauren  P0039942  F  18-25  24  0  Andhra Pradesh  Southern  Govt  Auto  2  23664.0  NaN  NaN
17  1005447  Amy  P0025642  F  46-50  48  1  Andhra Pradesh  Southern  IT Sector  Auto  3  NaN  NaN  NaN
18  1001193  Mick  P0005442  F  26-35  29  0  Andhra Pradesh  Southern  Aviation  Auto  1  23619.00  NaN  NaN
19  1001883  Pranant  P0002942  M  51-55  54  1  Uttar Pradesh  Central  Hospitality  Auto  1  23568.0  NaN  NaN
20  1001883  Pranant  P0002942  M  51-55  54  1  Uttar Pradesh  Central  Hospitality  Auto  1  23568.0  NaN  NaN
21  1001113  Elie  P0018542  F  18-25  19  1  Andhra Pradesh  Southern  Govt  Auto  4  23546.0  NaN  NaN
22  1000416  Hishesh  P0018342  F  46-50  46  1  Uttar Pradesh  Central  Banking  Auto  2  23525.0  NaN  NaN
23  1002556  Grant  P001142  F  26-35  30  0  Andhra Pradesh  Southern  IT Sector  Auto  1  23518.0  NaN  NaN
24  1001505  Gireesh  P0017342  F  51-55  53  0  Uttar Pradesh  Central  Automobile  Auto  2  23510.0  NaN  NaN
25  1000900  Skara  P0017342  M  55+  83  0  Karnataka  Southern  Automobile  Auto  3  23513.0  NaN  NaN
26  1006989  Eric  P0028242  F  26-35  33  0  Andhra Pradesh  Southern  IT Sector  Auto  3  23462.0  NaN  NaN
27  1001101  Gibson  P0023742  F  36-45  40  0  Uttar Pradesh  Central  Banking  Auto  3  23450.0  NaN  NaN
28  1004995  Madhu  P0005942  F  18-25  25  1  Andhra Pradesh  Southern  Banking  Auto  2  23451.00  NaN  NaN
29  1004037  Ehsal  P0019042  M  51-55  54  1  Andhra Pradesh  Southern  Govt  Hard & Power Tools  Auto  2  23434.00  NaN  NaN
30  1002340  James  P0011942  F  36-45  39  1  Andhra Pradesh  Southern  Aviation  Auto  3  23389.00  NaN  NaN

In [6]: df.info()
Out[6]:
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 11251 entries, 0 to 11250
Data columns (total 15 columns):
#   Column                Non-Null Count  Dtype
---  --
0   User_ID               11251 non-null  int64
1   Cust_name            11251 non-null  object
2   Product_ID           11251 non-null  object
3   Gender               11251 non-null  object
4   Age Group            11251 non-null  object
5   Age                 11251 non-null  int64
6   Marital_Status       11251 non-null  int64
7   State               11251 non-null  object
8   Zone               11251 non-null  object
9   Occupation           11251 non-null  object
10  Product_Category     11251 non-null  object
11  Orders              11251 non-null  int64
12  Amount              11239 non-null  float64
13  Status              6 non-null      float64
14  unnamed1            6 non-null      float64
dtypes: float64(3), int64(4), object(8)
memory usage: 1.1+ MB

In [7]: df.dropna(inplace=True)
df.drop('Status', inplace=True)

In [8]: df.head()
Out[8]:
   User_ID  Cust_name  Product_ID  Gender  Age Group  Age  Marital_Status  State  Zone  Occupation  Product_Category  Orders  Amount  Status  unnamed1
0  1002903  Sarani  P0012942  F  26-35  28  0  Maharashtra  Western  Healthcare  Auto  1  23952.0  NaN  NaN
1  100732  Karik  P0011942  F  26-35  35  1  Andhra Pradesh  Southern  Govt  Auto  3  23934.0  NaN  NaN
2  1001990  Bindu  P0018542  F  26-35  35  1  Uttar Pradesh  Central  Automobile  Auto  3  23924.0  NaN  NaN
3  1001425  Sudevi  P0023742  M  0-17  16  0  Karnataka  Southern  Construction  Auto  2  23912.0  NaN  NaN
4  1002689  Joni  P0025742  M  26-35  28  1  Gujarat  Western  Food Processing  Auto  2  23877.0  NaN  NaN
...
11246  1008085  Manning  P0029642  M  18-25  19  0  Andhra Pradesh  Southern  Banking  Auto  1  23450.0  NaN  NaN
11247  1004889  Reshenbach  P0017342  F  26-35  33  0  Andhra Pradesh  Southern  Banking  Auto  2  23451.00  NaN  NaN
11248  1001299  Oshin  P0028342  F  36-45  40  1  Uttar Pradesh  Central  Banking  Auto  2  23525.0  NaN  NaN
11249  1004823  Norman  P0025642  M  36-45  37  0  Andhra Pradesh  Southern  Govt  Hard & Power Tools  Auto  2  23434.00  NaN  NaN
11250  1002744  Brumley  P00281742  F  18-25  19  0  Andhra Pradesh  Southern  Aviation  Auto  3  23389.00  NaN  NaN

In [9]: df.info()
Out[9]:
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 11251 entries, 0 to 11250
Data columns (total 13 columns):
#   Column                Non-Null Count  Dtype
---  --
0   User_ID               11251 non-null  int64
1   Cust_name            11251 non-null  object
2   Product_ID           11251 non-null  object
3   Gender               11251 non-null  object
4   Age Group            11251 non-null  object
5   Age                 11251 non-null  int64
6   Marital_Status       11251 non-null  int64
7   State               11251 non-null  object
8   Zone               11251 non-null  object
9   Occupation           11251 non-null  object
10  Product_Category     11251 non-null  object
11  Orders              11251 non-null  int64
12  Amount              11239 non-null  float64
dtypes: float64(1), int64(4), object(8)
memory usage: 1.1+ MB

In [10]: #to check null values
pd.isnull(df).sum()
Out[10]:
User_ID      0
Cust_name    0
Product_ID   0
Gender       0
Age Group    0
Age          0
Marital_Status  0
State        0
Zone         0
Occupation   0
Product_Category  0
Orders       0
Amount       0
dtype: int64

In [11]: df.shape
Out[11]: (11251, 13)

In [12]: #drop null values
df.dropna(inplace=True)

In [13]: df.shape
Out[13
```