DATA SCIENCE

WEEK 3 PROJECT

CODE:

```
import matplotlib.pyplot as plt

from sklearn.preprocessing import MinMaxScaler

from sklearn.preprocessing import StandardScaler

import pandas as pd

shopping=pd.read_csv("E:\DATA SCIENCE\week3\shopping.csv")

print(shopping)

print("Min Max Scaler")

numeric_col=shopping.select_dtypes(include='number').columns

scaler=MinMaxScaler()

shopping_normalized=pd.DataFrame(scaler.fit_transform(shopping[numeric_col]),columns=
numeric_col)

print(shopping_normalized.head())

print("Standard Scaler")

numeric_col1=shopping.select_dtypes(include="number").columns

scaler=StandardScaler()

shopping_standardized=pd.DataFrame(scaler.fit_transform(shopping[numeric_col1]),column
s=numeric_col1)

print(shopping_standardized.head())

plt.figure(figsize=(8,6))

plt.hist(shopping['Avg_Price'],bins=10)

plt.title("Distribution of Sales", )

plt.xlabel("Avg_Price")

plt.ylabel("Frequency")

plt.show()
```

OUTPUT:

```
       Unnamed: 0  CustomerID Gender Location  Tenure_Months  ...  Offline_Spend Online_Spend Month Coupon_Code Discount_pct
0               0     17850.0      M  Chicago           12.0  ...         4500.0       2424.5     1      ELEC10         10.0
1               1     17850.0      M  Chicago           12.0  ...         4500.0       2424.5     1      ELEC10         10.0
2               2     17850.0      M  Chicago           12.0  ...         4500.0       2424.5     1      ELEC10         10.0
3               3     17850.0      M  Chicago           12.0  ...         4500.0       2424.5     1      ELEC10         10.0
4               4     17850.0      M  Chicago           12.0  ...         4500.0       2424.5     1      ELEC10         10.0
...           ...         ...    ...      ...            ...  ...            ...          ...   ...         ...          ...
52950       52950         NaN    NaN      NaN            NaN  ...            NaN          NaN    11        GC20         20.0
52951       52951         NaN    NaN      NaN            NaN  ...            NaN          NaN    11        NJ20         20.0
52952       52952         NaN    NaN      NaN            NaN  ...            NaN          NaN    10       AND10         10.0
52953       52953         NaN    NaN      NaN            NaN  ...            NaN          NaN    11       AND20         20.0
52954       52954         NaN    NaN      NaN            NaN  ...            NaN          NaN    12       AND30         30.0

[52955 rows x 21 columns]
Min Max Scaler
   Unnamed: 0  CustomerID  Tenure_Months  Transaction_ID  Quantity  ...       GST  Offline_Spend  Online_Spend  Month  Discount_pct
0    0.000000    0.927068       0.208333        0.000000  0.000000  ...  0.384615       0.888889      0.496674    0.0           0.0
1    0.000019    0.927068       0.208333        0.000031  0.000000  ...  0.384615       0.888889      0.496674    0.0           0.0
2    0.000038    0.927068       0.208333        0.000534  0.001112  ...  0.384615       0.888889      0.496674    0.0           0.0
3    0.000057    0.927068       0.208333        0.000629  0.000000  ...  0.384615       0.888889      0.496674    0.0           0.0
4    0.000076    0.927068       0.208333        0.000660  0.000000  ...  0.384615       0.888889      0.496674    0.0           0.0

[5 rows x 12 columns]
```

```
Standard Scaler
   Unnamed: 0  CustomerID  Tenure_Months  Transaction_ID  Quantity  ...       GST  Offline_Spend  Online_Spend     Month  Discount_pc
t
0   -1.732018    1.417059      -1.048214       -1.818890 -0.173973  ... -0.817509       1.782934      0.658472 -1.695688     -1.22472
6
1   -1.731953    1.417059      -1.048214       -1.818774 -0.173973  ... -0.817509       1.782934      0.658472 -1.695688     -1.22472
6
2   -1.731887    1.417059      -1.048214       -1.816924 -0.124233  ... -0.817509       1.782934      0.658472 -1.695688     -1.22472
6
3   -1.731822    1.417059      -1.048214       -1.816577 -0.173973  ... -0.817509       1.782934      0.658472 -1.695688     -1.22472
6
4   -1.731756    1.417059      -1.048214       -1.816461 -0.173973  ... -0.817509       1.782934      0.658472 -1.695688     -1.22472
6

[5 rows x 12 columns]
```