

EXPERIMENT N0-12

WRITE PYTHON PROGRAM TO READ CSV FILE AND PERFORM FOLLOWING OPERATION

1:READ CSV FILE:

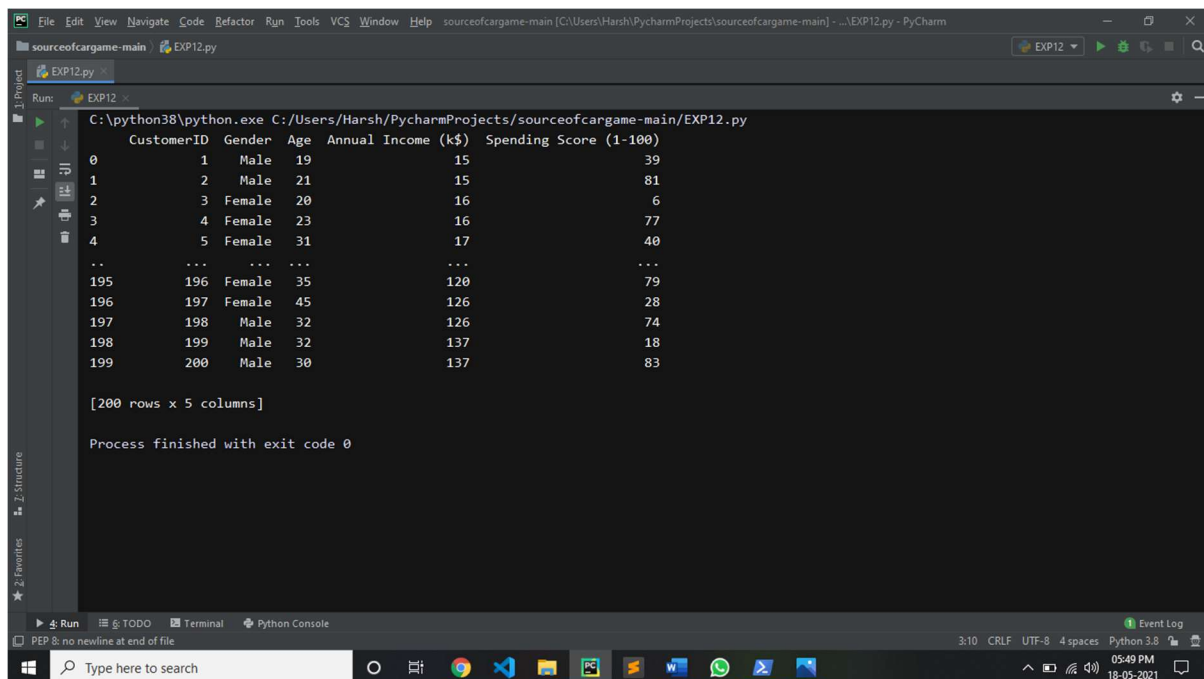
CODE:

```
import pandas as pd

df = pd.read_csv('Mall_Customers.csv')

print(df)
```

OUTPUT:



```
C:\python38\python.exe C:/Users/Harsh/PycharmProjects/sourceofcargame-main/EXP12.py
CustomerID Gender Age Annual Income (k$) Spending Score (1-100)
0 1 Male 19 15 39
1 2 Male 21 15 81
2 3 Female 20 16 6
3 4 Female 23 16 77
4 5 Female 31 17 40
.. ..
195 196 Female 35 120 79
196 197 Female 45 126 28
197 198 Male 32 126 74
198 199 Male 32 137 18
199 200 Male 30 137 83

[200 rows x 5 columns]

Process finished with exit code 0
```

2:READ FIRST 10,20,50 RECORDS:

CODE:

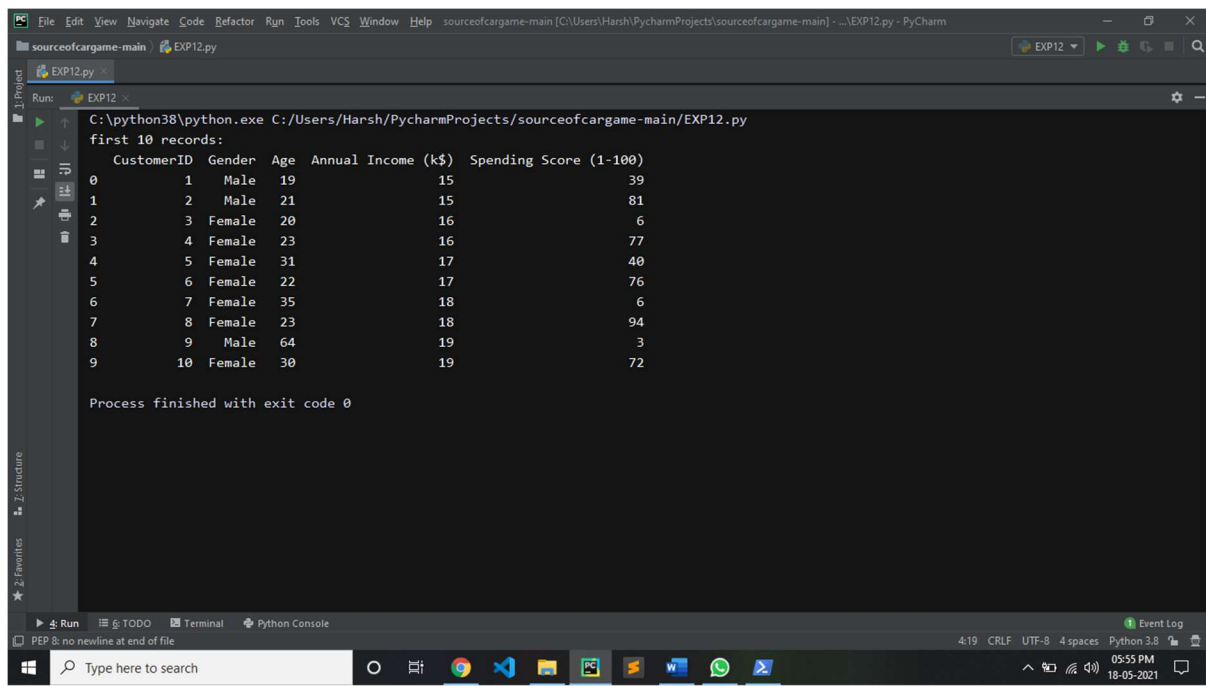
```
import pandas as pd

df = pd.read_csv('Mall_Customers.csv')

print("first 10 records:")

print(df.head(10))
```

OUTPUT:



```
Run: EXP12
C:\python38\python.exe C:/Users/Harsh/PycharmProjects/sourceofcargame-main/EXP12.py
first 10 records:
  CustomerID  Gender  Age  Annual Income (k$)  Spending Score (1-100)
0           1   Male   19                15                39
1           2   Male   21                15                81
2           3  Female  20                16                 6
3           4  Female  23                16                77
4           5  Female  31                17                40
5           6  Female  22                17                76
6           7  Female  35                18                 6
7           8  Female  23                18                94
8           9   Male  64                19                 3
9          10  Female  30                19                72

Process finished with exit code 0
```

CODE:

```
import pandas as pd

df = pd.read_csv('Mall_Customers.csv')

print("first 20 records:")

print(df.head(20))
```

OUTPUT:

```
Run: EXP12.py
C:\python38\python.exe C:/Users/Harsh/PycharmProjects/sourceofcargame-main/EXP12.py
first 20 records:
  CustomerID  Gender  Age  Annual Income (k$)  Spending Score (1-100)
0           1    Male   19                15                39
1           2    Male   21                15                81
2           3  Female   20                16                 6
3           4  Female   23                16                77
4           5  Female   31                17                40
5           6  Female   22                17                76
6           7  Female   35                18                 6
7           8  Female   23                18                94
8           9    Male   64                19                 3
9          10  Female   30                19                72
10          11    Male   67                19                14
11          12  Female   35                19                99
12          13  Female   58                20                15
13          14  Female   24                20                77
14          15    Male   37                20                13
15          16    Male   22                20                79
16          17  Female   35                21                35
17          18    Male   20                21                66
18          19    Male   52                23                29
19          20  Female   35                23                98

Process finished with exit code 0
```

CODE:

```
import pandas as pd

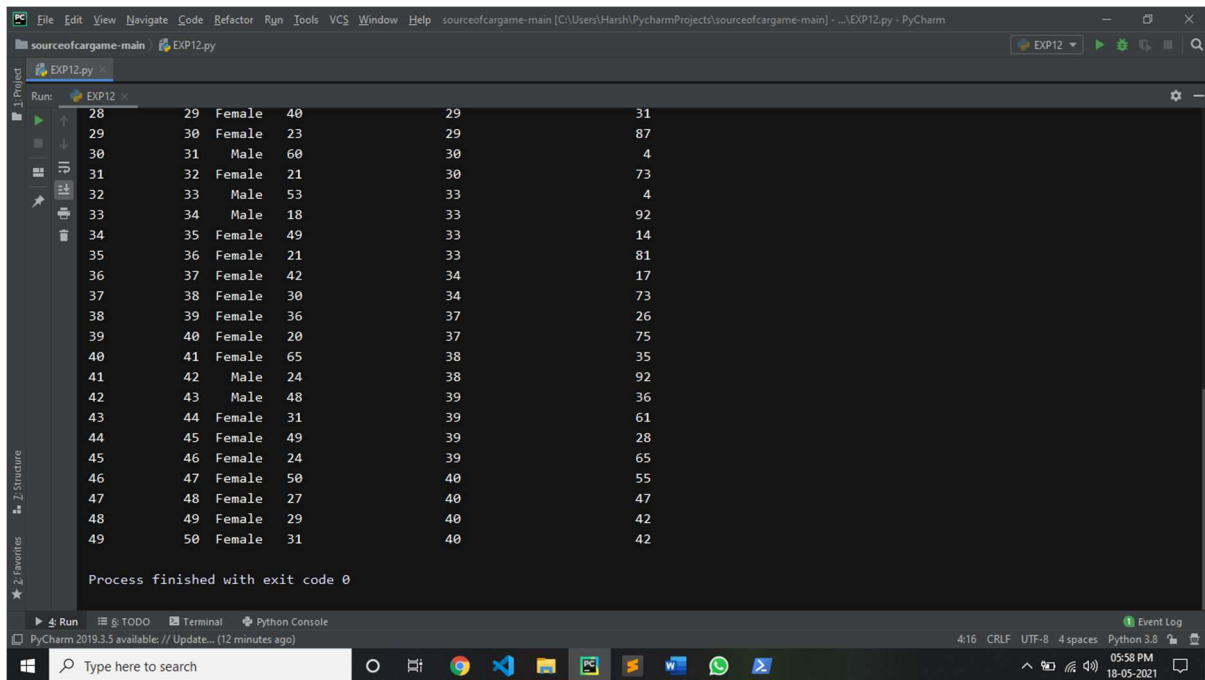
df = pd.read_csv('Mall_Customers.csv')

print("first 50 records:")

print(df.head(50))
```

OUTPUT:

```
Run: EXP12.py
C:\python38\python.exe C:/Users/Harsh/PycharmProjects/sourceofcargame-main/EXP12.py
first 50 records:
  CustomerID  Gender  Age  Annual Income (k$)  Spending Score (1-100)
0           1    Male   19                15                39
1           2    Male   21                15                81
2           3  Female   20                16                 6
3           4  Female   23                16                77
4           5  Female   31                17                40
5           6  Female   22                17                76
6           7  Female   35                18                 6
7           8  Female   23                18                94
8           9    Male   64                19                 3
9          10  Female   30                19                72
10          11    Male   67                19                14
11          12  Female   35                19                99
12          13  Female   58                20                15
13          14  Female   24                20                77
14          15    Male   37                20                13
15          16    Male   22                20                79
16          17  Female   35                21                35
17          18    Male   20                21                66
18          19    Male   52                23                29
19          20  Female   35                23                98
20          21    Male   35                24                35
21          22    Male   25                24                73
```



3: CHECK DATATYPE OF ALL THE COLUMNS:

CODE:

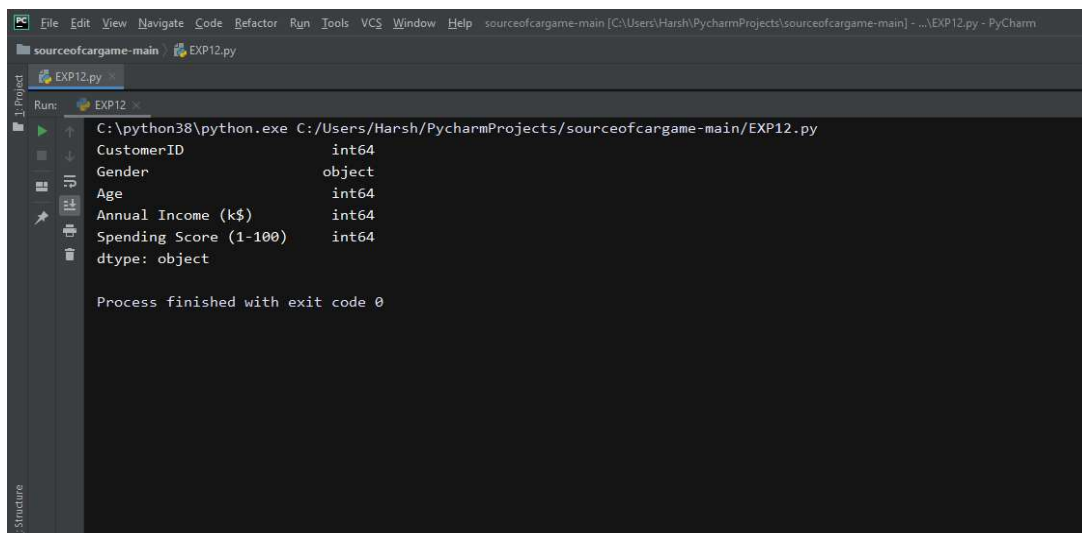
```
import pandas as pd

df = pd.read_csv('Mall_Customers.csv')

res = df.dtypes

print(res)
```

OUTPUT:



4:

**DISPLAY MIN,MAX,SUM,COUNT,PROD,MEAN,MEDIAN,MODE,STD
OF GIVEN DATASET:**

CODE:

```
import pandas as pd
df = pd.read_csv('Mall_Customers.csv')
print("mean is:")
print(df.mean())
print("-----")
print("min is:")
print(df.min())
print("-----")
print("max is:")
print(df.max())
print("-----")
print("count is:")
print(df.count())
print("-----")
print("sum is:")
print(df.sum())
print("-----")
print("prod is:")
print(df.prod())
print("-----")
print("median is:")
print(df.median())
print("-----")
print("mode is:")
```

```

print(df.mode())

print("-----")

print("standard deviation is:")

print(df.std())

```

OUTPUT:

The first screenshot shows the output of the following code:

```

mean is:
CustomerID      100.50
Age             38.85
Annual Income (k$)  60.56
Spending Score (1-100)  50.20
dtype: float64

min is:
CustomerID      1
Gender          Female
Age            18
Annual Income (k$)  15
Spending Score (1-100)  1
dtype: object

max is:
CustomerID      200
Gender          Male
Age            70
Annual Income (k$)  137
Spending Score (1-100)  99
dtype: object

count is:

```

The second screenshot shows the output of the following code:

```

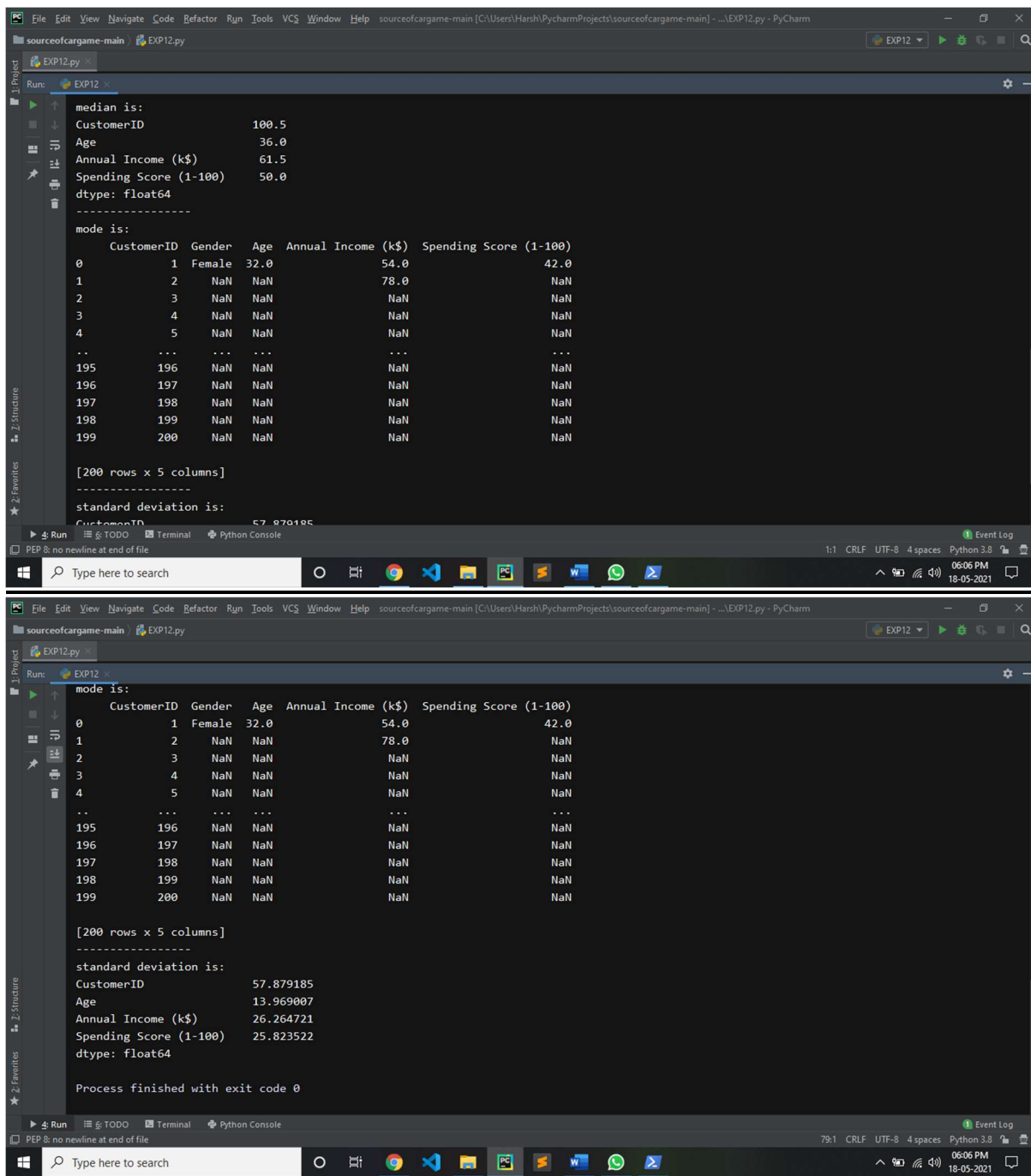
count is:
CustomerID      200
Gender          200
Age            200
Annual Income (k$)  200
Spending Score (1-100)  200
dtype: int64

sum is:
CustomerID      20100
Gender          MaleMaleFemaleFemaleFemaleFemaleFemaleFemale...
Age            7770
Annual Income (k$)  12112
Spending Score (1-100)  10040
dtype: object

prod is:
CustomerID      0
Age            0
Annual Income (k$)  0
Spending Score (1-100)  0
dtype: int64

median is:
CustomerID      100.5

```



5: FIND HOW MANY VALUES IN ANNUAL INCOME COLUMN:

CODE:

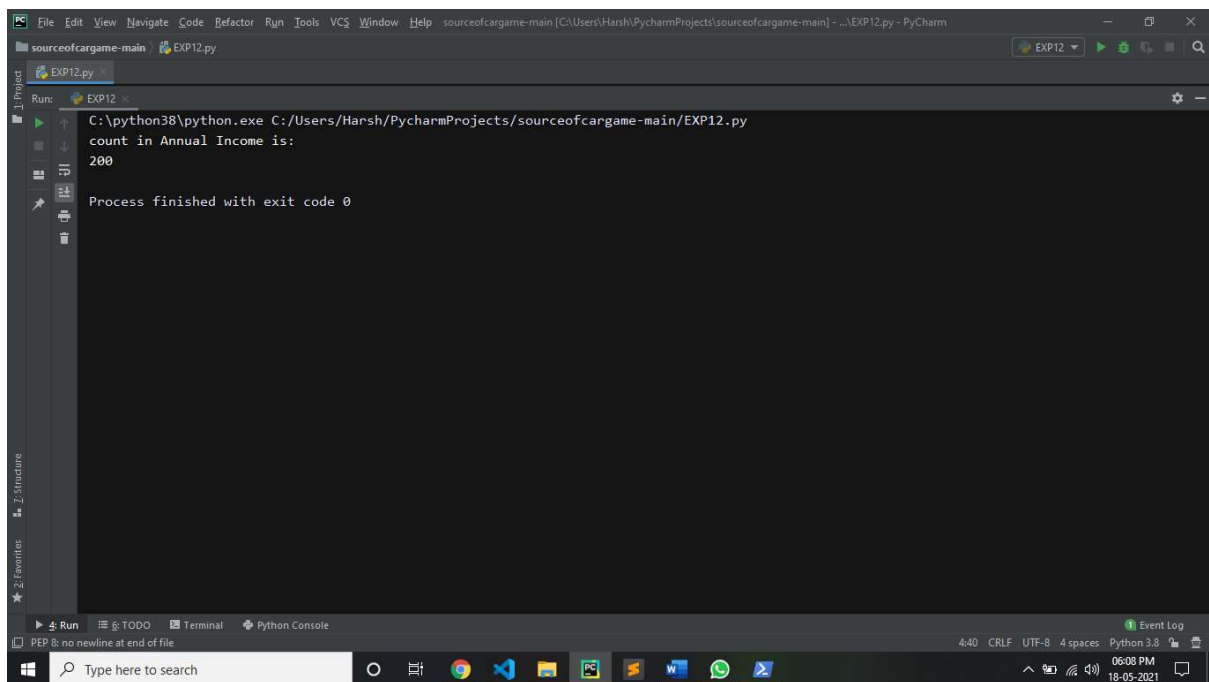
```
import pandas as pd
```

```
df = pd.read_csv('Mall_Customers.csv')
```

```
print("count in Annual Income is:")
```

```
print(df['Annual Income (k$)'].count())
```

OUTPUT:



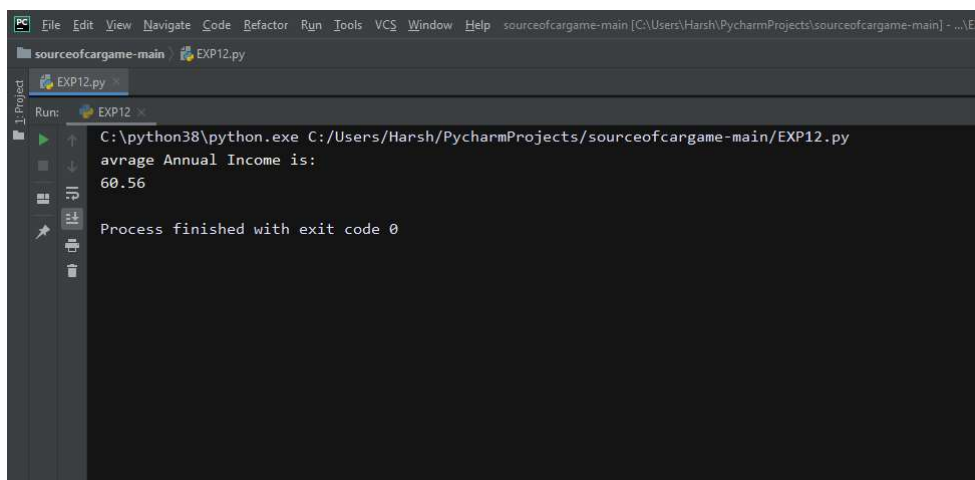
```
File Edit View Navigate Code Refactor Run Tools VCS Window Help sourceofcargame-main [C:/Users/Harsh/PycharmProjects/sourceofcargame-main] - ...EXP12.py - PyCharm
sourceofcargame-main EXP12.py
Run: EXP12 x
C:\python38\python.exe C:/Users/Harsh/PycharmProjects/sourceofcargame-main/EXP12.py
count in Annual Income is:
200
Process finished with exit code 0
4:40 CRLF UTF-8 4 spaces Python 3.8 18-05-2021
```

6: CALCULATE AVRAGE ANNUAL INCOME:

CODE:

```
import pandas as pd
df = pd.read_csv('Mall_Customers.csv')
print("avrage Annual Income is:")
print(df['Annual Income (k$)'].mean())
```

OUTPUT:



```
File Edit View Navigate Code Refactor Run Tools VCS Window Help sourceofcargame-main [C:/Users/Harsh/PycharmProjects/sourceofcargame-main] - ...EXP12.py
sourceofcargame-main EXP12.py
Run: EXP12 x
C:\python38\python.exe C:/Users/Harsh/PycharmProjects/sourceofcargame-main/EXP12.py
avrage Annual Income is:
60.56
Process finished with exit code 0
```


7: CALCULATE MEAN ANNUAL INCOME OF EACH MALE CUSTOMER:

CODE:

```
import pandas as pd

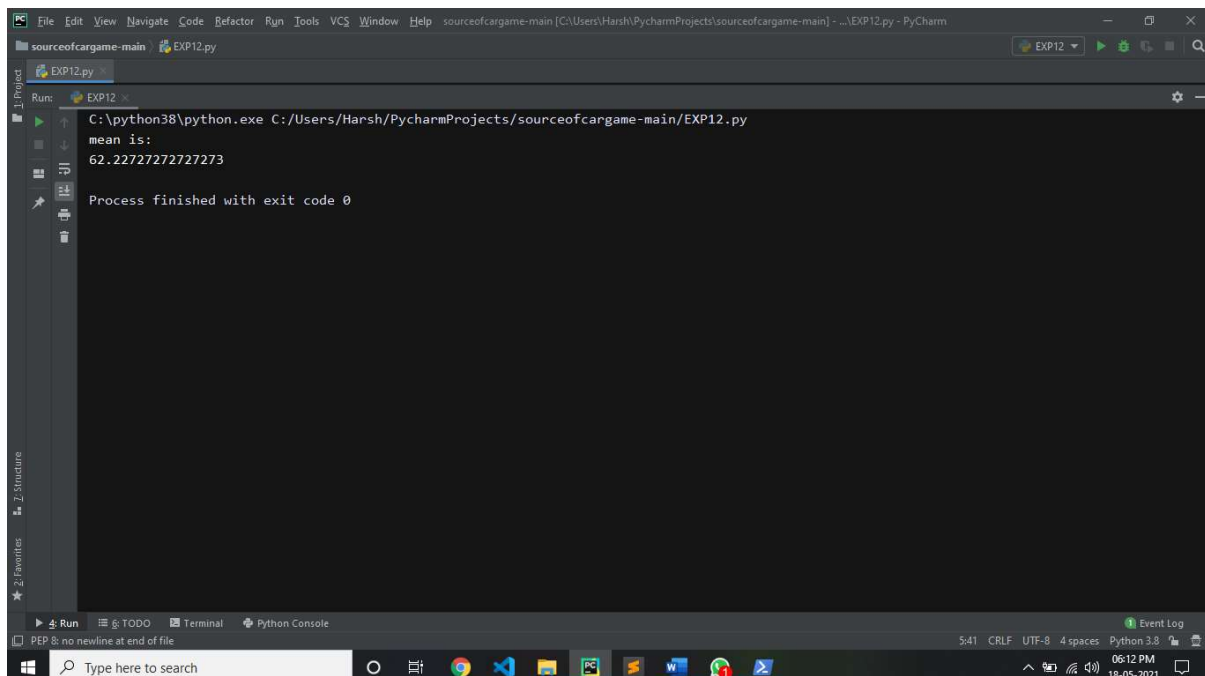
df = pd.read_csv('Mall_Customers.csv')

male = df[df['Gender'] == 'Male']

print('mean is:')

print(male['Annual Income (k$)'].mean())
```

OUTPUT:



```
Run: EXP12
C:\python38\python.exe C:/Users/Harsh/PycharmProjects/sourceofcargame-main/EXP12.py
mean is:
62.22727272727273
Process finished with exit code 0
```

8: SELECT ONLY THOSE ROW WHICH CONTAIN FEAMLE :

CODE:

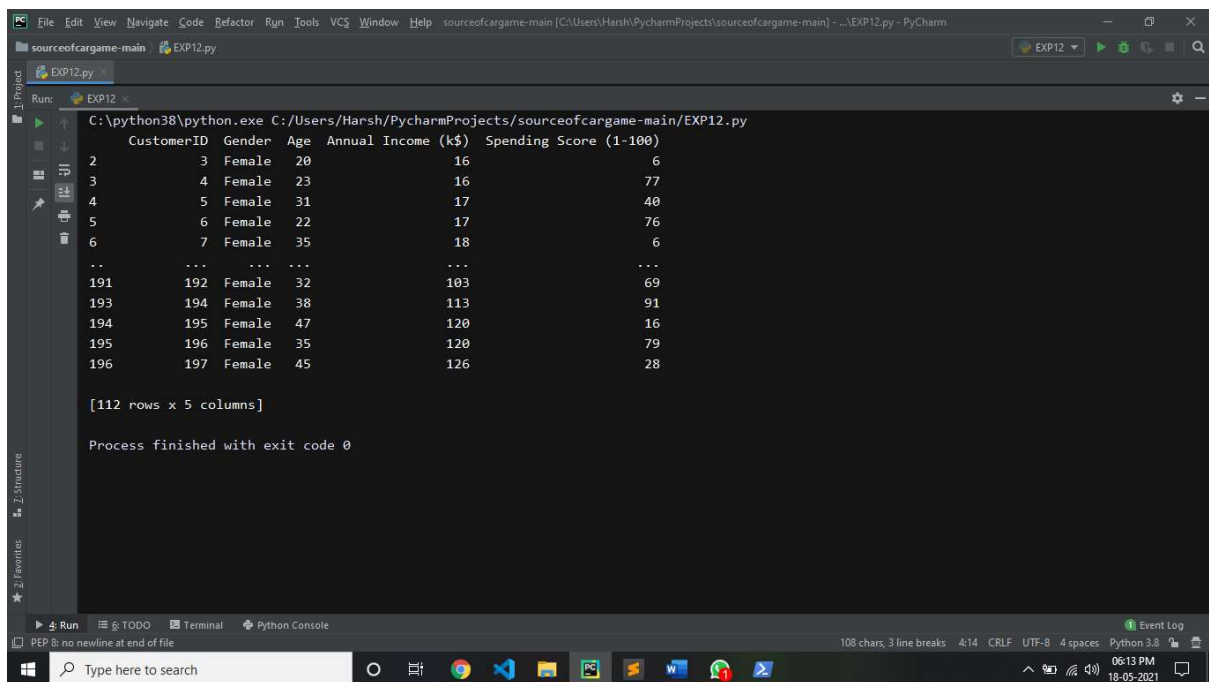
```
import pandas as pd

df = pd.read_csv('Mall_Customers.csv')

female = df[df['Gender'] == 'Female']

print(female)
```

OUTPUT:



```
C:\python38\python.exe C:/Users/Harsh/PycharmProjects/sourceofcargame-main/EXP12.py
CustomerID Gender Age Annual Income (k$) Spending Score (1-100)
2 3 Female 20 16 6
3 4 Female 23 16 77
4 5 Female 31 17 40
5 6 Female 22 17 76
6 7 Female 35 18 6
.. ..
191 192 Female 32 103 69
193 194 Female 38 113 91
194 195 Female 47 120 16
195 196 Female 35 120 79
196 197 Female 45 126 28

[112 rows x 5 columns]

Process finished with exit code 0
```

9: CREATE NEW DATAFRAME FROM ORIGINAL SORTED BY SPENDING SCORE:

CODE:

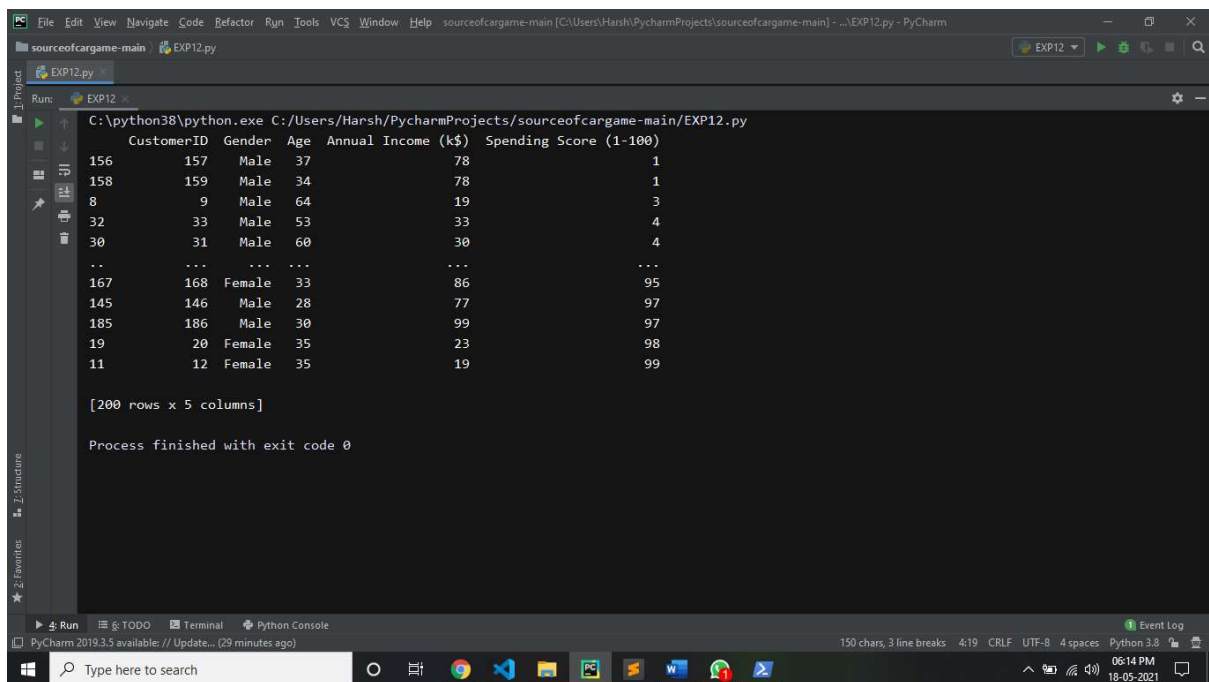
```
import pandas as pd

df = pd.read_csv('Mall_Customers.csv')

sorted_data = df.sort_values("Spending Score (1-100)", ascending = True)

print(sorted_data)
```

OUTPUT:



```
C:\python38\python.exe C:/Users/Harsh/PycharmProjects/sourceofcargame-main/EXP12.py
CustomerID Gender Age Annual Income (k$) Spending Score (1-100)
156 157 Male 37 78 1
158 159 Male 34 78 1
8 9 Male 64 19 3
32 33 Male 53 33 4
30 31 Male 60 30 4
... ..
167 168 Female 33 86 95
145 146 Male 28 77 97
185 186 Male 30 99 97
19 20 Female 35 23 98
11 12 Female 35 19 99

[200 rows x 5 columns]

Process finished with exit code 0
```