MACHINE LEARNING (CSI0702)

PRACTICAL-2
(PANDAS)

Submitted By:

Harsh Saradva (20BCM024)

Submitted To:

Meenaxi Tank

1.

```
import pandas as pd
import numpy as np

ser = pd.Series()
data = np.array(['g', 'e', 'e', 'k', 's'])
ser = pd.Series(data)
print(ser)
```

output:

```
0    g
1    e
2    e
3    k
4    s
dtype: object

[Done] exited with code=0 in 1.117 seconds
```

2.

```
import pandas as pd
df = pd.DataFrame()
list = ['my', 'name', 'is', 'harsh',
            'and', 'my', 'roll', 'no', 'is', '24']
df = pd.DataFrame(list)
print(df)
```

output:

```
[Running] python -u "s:\TERM 7\ML\PRACTICALS\LAB-2\b.py"
       0
0     my
1   name
2     is
3  harsh
4    and
5     my
6   roll
7     no
8     is
9     24

[Done] exited with code=0 in 0.673 seconds
```

3.

```
import pandas as pd
df = pd.read_csv('S:\TERM 7\ML\PRACTICALS\LAB-2\data.csv')
print(df.to_string())
```

output:

```
[Running] python -u "s:\TERM 7\ML\PRACTICALS\LAB-2\c.py"
    rollno      name  marks  total marks
0        1      vats     77          100
1        2   devdeep     81          100
2        3       jay     49          100
3        4      nosh     68          100
4        5    nikhil     38          100
5        6     nehil     97          100
6        7   devansh     46          100
7        8      brax     18          100
8        9     jinay     24          100
9       10    bhavya     49          100
10      11    nandan     77          100
11      12    lovish     39          100
12      13     tirth     48          100
13      14   vishwam     36          100
14      15        om     79          100
15      16       yug     18          100
16      17  pururava     37          100

[Done] exited with code=0 in 0.771 seconds
```

4.

```
import pandas
mydataset = {  'cars': ["BMW", "Volvo", "Ford"],  'passings': [3, 7, 2]}
myvar = pandas.DataFrame(mydataset)
print(myvar)
```

output:

```
[Running] python -u "s:\TERM 7\ML\PRACTICALS\LAB-2\tempCodeRunnerFile.py"
    cars  passings
0    BMW         3
1  Volvo         7
2   Ford         2

[Done] exited with code=0 in 0.692 seconds
```

5.

```
import pandas as pd
a = [1, 7, 2]
myvar = pd.Series(a, index = ["x", "y", "z"])
print(myvar)
```

output:

```
[Running] python -u "s:\TERM 7\ML\PRACTICALS\LAB-2\e.py"
x    1
y    7
z    2
dtype: int64

[Done] exited with code=0 in 0.661 seconds
```

6.

```
import pandas as pd
calories = {"day1": 420, "day2": 380, "day3": 390}
myvar = pd.Series(calories, index = ["day1", "day2"])
print(myvar)
```

output:

```
[Running] python -u "s:\TERM 7\ML\PRACTICALS\LAB-2\f.py"
day1    420
day2    380
dtype: int64

[Done] exited with code=0 in 0.634 seconds
```

7.

```
import pandas as pd
f = ['FB', '2001-08-02', 90, 3.2]
s = pd.Series(f, index = ['name', 'date', 'shares', 'price'])
print(s)
print(s['price'])
print(s[['name', 'date']])
```

output:

```
[Running] python -u "s:\TERM 7\ML\PRACTICALS\LAB-2\g.py"
name              FB
date      2001-08-02
shares            90
price            3.2
dtype: object
3.2
name            FB
date    2001-08-02
dtype: object

[Done] exited with code=0 in 0.66 seconds
```

8.

```
import pandas as pd
data = {'name': ['AA', 'IBM', 'GOOG'], 'date': ['2001-12-01', '2012-02-10',
'2010-04-09'], 'shares': [100, 30, 90],
         'price': [12.3, 10.3, 32.2]}
df = pd.DataFrame(data)
df['owner'] = ['Rohan', 'Unknown', 'Unknown']
df.index = ['one', 'two', 'three']
print(df)
df = df.set_index(['name'])
print(df)
print(df['shares'])
del df['owner']
print(df)
```

output:

```
[Running] python -u "s:\TERM 7\ML\PRACTICALS\LAB-2\h.py"
       name        date  shares  price    owner
one      AA  2001-12-01     100   12.3    Rohan
two     IBM  2012-02-10      30   10.3  Unknown
three  GOOG  2010-04-09      90   32.2  Unknown
              date  shares  price    owner
name
AA    2001-12-01     100   12.3    Rohan
IBM   2012-02-10      30   10.3  Unknown
GOOG  2010-04-09      90   32.2  Unknown
name
AA      100
IBM      30
GOOG     90
Name: shares, dtype: int64
              date  shares  price
name
AA    2001-12-01     100   12.3
IBM   2012-02-10      30   10.3
GOOG  2010-04-09      90   32.2

[Done] exited with code=0 in 0.665 seconds
```

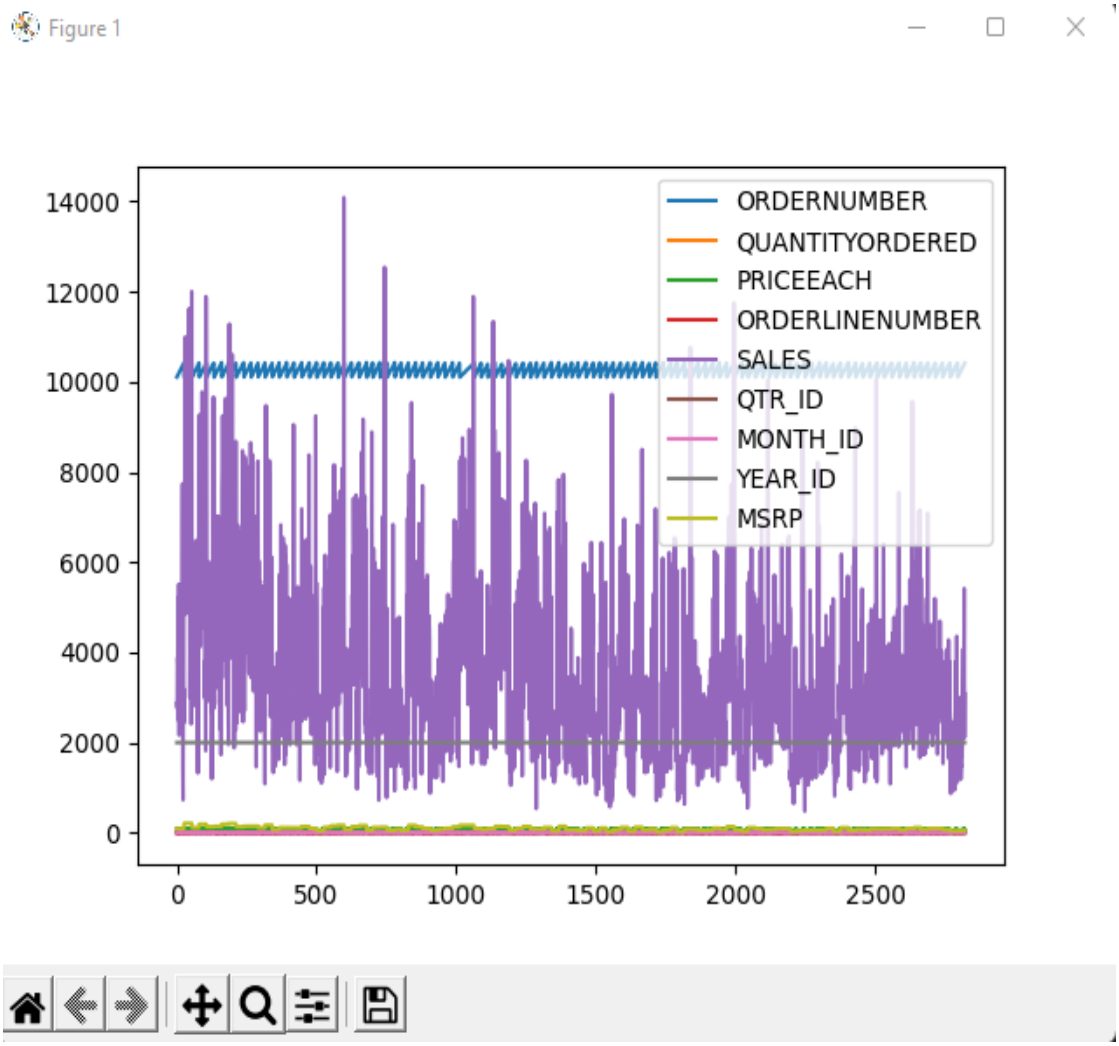
9.

```
import pandas as pd
import matplotlib.pyplot as plt

sales = pd.read_csv('S:\TERM 7\ML\PRACTICALS\LAB-2\sales_data_sample.csv',
encoding='Latin-1')

sales.plot()
plt.show()
```

output:



10.

```
import pandas as pd
import matplotlib.pyplot as plt


sales = pd.read_csv('S:\TERM 7\ML\PRACTICALS\LAB-2\sales_data_sample.csv',
encoding='Latin-1')
print(sales.head())
print(sales.tail())
```

output:

```
[Running] python -u "s:\TERM 7\ML\PRACTICALS\LAB-2\j.py"
   ORDERNUMBER  QUANTITYORDERED  ...  CONTACTFIRSTNAME  DEALSIZE
0        10107               30  ...              Kwai     Small
1        10121               34  ...              Paul     Small
2        10134               41  ...            Daniel    Medium
3        10145               45  ...             Julie    Medium
4        10159               49  ...             Julie    Medium

[5 rows x 25 columns]
      ORDERNUMBER  QUANTITYORDERED  ...  CONTACTFIRSTNAME  DEALSIZE
2818        10350               20  ...             Diego     Small
2819        10373               29  ...            Pirkko    Medium
2820        10386               43  ...             Diego    Medium
2821        10397               34  ...           Annette     Small
2822        10414               47  ...              Juri    Medium

[5 rows x 25 columns]

[Done] exited with code=0 in 1.28 seconds
```

11.

```
import pandas as pd
import matplotlib.pyplot as plt
sales = pd.read_csv('S:\TERM 7\ML\PRACTICALS\LAB-2\sales_data_sample.csv',
encoding='Latin-1')
sorted_sales = sales.sort_values(by=["CONTACTFIRSTNAME"], )
print(sorted_sales.head())
filterd_sales = sales[sales['CONTACTFIRSTNAME'].str.contains('Wendy')]
print(filterd_sales.head())
```

output:

```
[Running] python -u "s:\TERM 7\ML\PRACTICALS\LAB-2\k.py"
      ORDERNUMBER  QUANTITYORDERED  ...  CONTACTFIRSTNAME  DEALSIZE
1414        10420               36  ...            Adrian     Small
40          10270               21  ...            Adrian    Medium
2555        10361               24  ...            Adrian     Small
1715        10420               39  ...            Adrian    Medium
2764        10361               35  ...            Adrian    Medium

[5 rows x 25 columns]
     ORDERNUMBER  QUANTITYORDERED  ...  CONTACTFIRSTNAME  DEALSIZE
143        10217               48  ...             Wendy     Large
146        10259               26  ...             Wendy    Medium
380        10259               46  ...             Wendy    Medium
542        10288               20  ...             Wendy     Small
591        10288               32  ...             Wendy    Medium

[5 rows x 25 columns]

[Done] exited with code=0 in 1.308 seconds
```

12.

```
import pandas as pd
data = {
  "calories": [420, 380, 390],
  "duration": [50, 40, 45]
}
df = pd.DataFrame(data, index = ["day1", "day2", "day3"])
print(df.loc["day2"])
```

output:

```
[Running] python -u "s:\TERM 7\ML\PRACTICALS\LAB-2\1.py"
calories    380
duration     40
Name: day2, dtype: int64

[Done] exited with code=0 in 0.605 seconds
```

13.

```
import pandas as pd
df = pd.read_csv('S:\TERM 7\ML\PRACTICALS\LAB-2\data.csv')
print(df.info())
```

output:

```
[Running] python -u "s:\TERM 7\ML\PRACTICALS\LAB-2\m.py"
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 17 entries, 0 to 16
Data columns (total 4 columns):
 #   Column       Non-Null Count  Dtype
---  ------       --------------  -----
 0   rollno       17 non-null     int64
 1   name         17 non-null     object
 2   marks        17 non-null     int64
 3   total marks  17 non-null     int64
dtypes: int64(3), object(1)
memory usage: 672.0+ bytes
None

[Done] exited with code=0 in 0.632 seconds
```

14.

```
import pandas as pd
df = pd.read_csv('S:\TERM 7\ML\PRACTICALS\LAB-2\data.csv')
print(df.head(10))
```

output:

```
[Running] python -u "s:\TERM 7\ML\PRACTICALS\LAB-2\n.py"
   rollno     name  marks  total marks
0       1     vats     77          100
1       2  devdeep     81          100
2       3      jay     49          100
3       4     nosh     68          100
4       5   nikhil     38          100
5       6    nehil     97          100
6       7  devansh     46          100
7       8     brax     18          100
8       9    jinay     24          100
9      10   bhavya     49          100

[Done] exited with code=0 in 0.634 seconds
```

15.

```
import pandas as pd
import numpy as np
a = pd.Series([2, 1, 1, np.nan, 3])
print(a)
```

output:

```
[Running] python -u "s:\TERM 7\ML\PRACTICALS\LAB-2\o.py"
0    2.0
1    1.0
2    1.0
3    NaN
4    3.0
dtype: float64

[Done] exited with code=0 in 0.614 seconds
```