Ex.1 Load the titanic dataset and convert it into a data frame.

AIM:

To perform basic programming and exploring data analysis on the titanic dataset using pandas, seaborn and sklearn.

Procedure/Algorithm:

Step 1: Load the titanic dataset into a datafram

Step 2: Display the first few new using head.

Step 3: Explore column data type and check for missing value using info() and shell() scan()

Step 4: Apply forward fill and back fill method to the age column.

Step 5: fill mining column values with unknown.

Step 6: Remove any duplicable head label Encoder.

Step 7: Encode the sex colum with label encoder.

Step 8: Scale column using standard Scaler.

Step 9: Create a pair plot for 'Pda'

Step 10: Display conclusion treat map.

Output:

First 5 ROW

Survived	pclass	sex	age	sibSp	Parch	fare	emb	class
0	3	M	22	1	0	7.25	S	Third
1	1	F	38	1	0	T.1283	C	First
1	3	F	26	0	0	7.925	S	Third
1	1	F	35	1	0	53.10	S	First
0	3	M	35	0	0	8.050	S	Third

who	adult male	deck	embark town	alive	alone
man	True	NAN	Southampton	NO	F
woman	False	C	Cherbourg	Yes	F
woman	False	NAN	Southampton	yes	T
woman	False	C	Southampton	Yes	F
man	True	NAN	Southampton	NO	T

Data info:

<class 'Pandas.core.frame.DataFrame'>

Range Index : 891 entries, 0 to 890

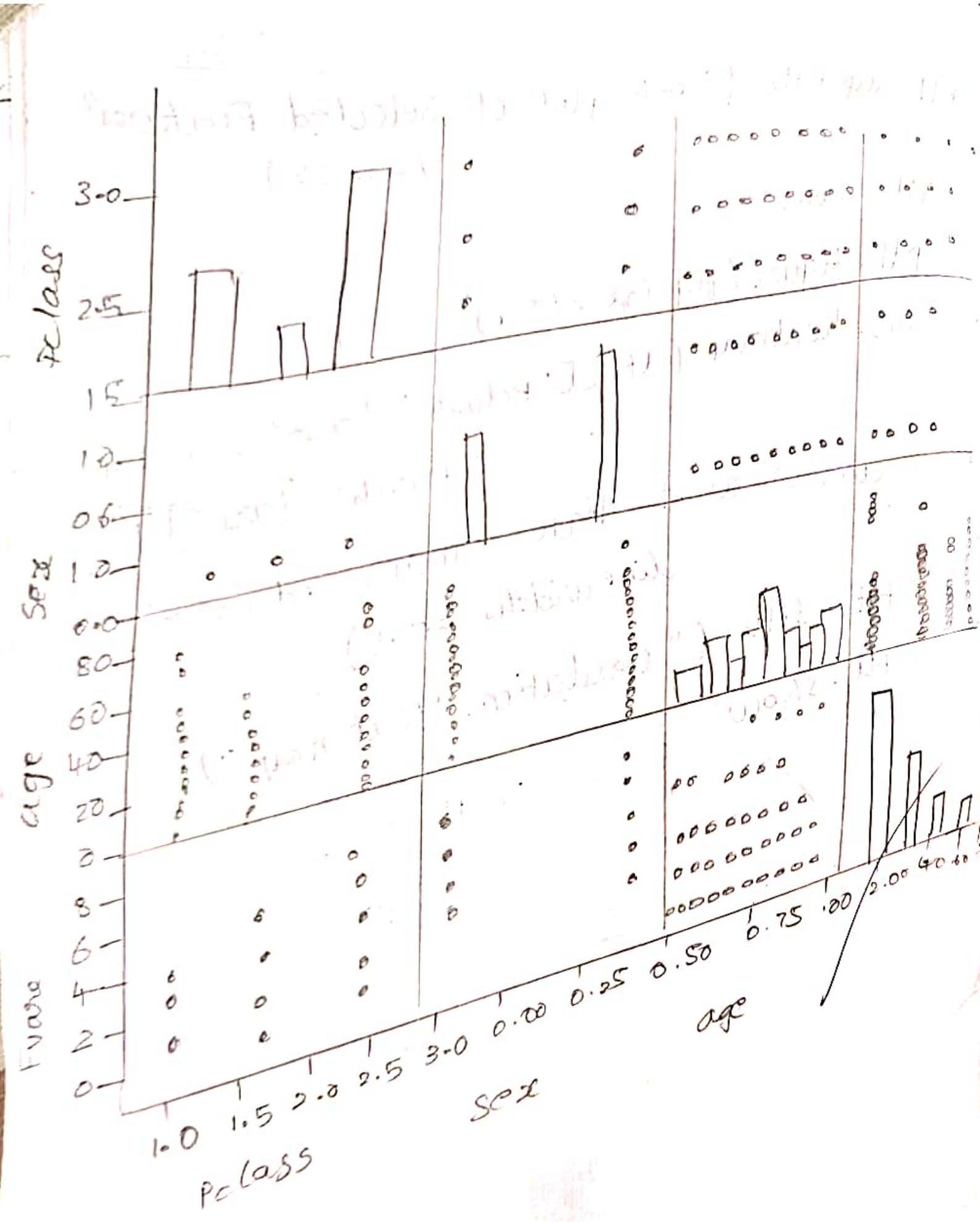
Data columns (total 15 columns

#	Columns

Code:

```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.preprocessing import label
                                  Standard Scale.

df = sns.load.dataset ('titatic')
Print (("First 5 rows"))
display (df.head())
Print ("\n Data Info :")
df.info()
df ['age'] = df ['age'].fillna (method
                                  = 'bfill')
df ['deck'] = df ['deck'].cat.add.categories
                ('Unknown') fillna
                        ('unknown', limit = 5)
df = df.dmp_duplicates()
le = label ==Encoder()
df ['sex'] = le.fit-perform (df ('sex').atype
                                  (str)

Scaler = standard scaler()
af ['fare'] = scaler.fit-transform (df['fare']
                                  fillna (0))
Sns.pairplot (df [['Pclass', 'sex', 'age',
                  'sibsp']).dropna()
```

```
PH. Sup title ("Pair plot of Selected Features".
                                    y=1.02)

Plt. Show()

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

Sns. heatmap (df [['Pclass', 'age', 'sibsp',
                   'Parch', 'fare']]

    Corr().annot = True . (map='coolwarm',
                    linewidths =0.5)

Plt. title ("correlation. Heat map")

plt. show
```


Pclass
sex
age
Fvara
3.0
2.5
1.5
1.0
0.0
80
60
40
20
8
6
4
2
0
1.0
1.5
2.0
2.5
3.0
0.00
0.25
0.50
0.75
1.00
2.00
4.00

RESULT:

Load the titanic dataset and convert it into a data frame is completed successfully.