

• To read the files located and  
convert it into a DataFrame and  
explore it.

### Program code

```

# Import numpy as np
import numpy as np

# Import pandas as pd
import pandas as pd

# Import matplotlib.pyplot as plt
import matplotlib.pyplot as plt

# Import Sklearn. Preprocessing and Label
from sklearn.preprocessing import LabelEncoder
from sklearn import preprocessing

# Import Seaborn as sns
import seaborn as sns

# Load dataset
balance = sns.load_dataset('balance')

# Print (shape info)
print(balance.shape)

# Print (Balance. head())
print(balance.head(5))

# Print (Balance. sum())
print(balance.sum())

# Print (Balance. count())
print(balance.count())

# Print (Balance. mean())
print(balance.mean())

# Print (Balance. std())
print(balance.std())

# Print (Balance. min())
print(balance.min())

# Print (Balance. max())
print(balance.max())

```

• loc [Crossing - deck, rudder, deck]  
= 1 unknown,  
0 known

### Categorical ([Unknown])

Balance [deck] - frame [deck], 0.0000

Output: 2020-01-22T09:28:09.000Z

	survived	pclass	Age	sibsp	parch	fare	class
0	0	3	22.0	1	0	7.2500	Third
1	1	1	38.0	1	0	71.2513	First
2	1	3	26.0	0	0	7.9250	Third
3	1	3	35.0	1	0	53.1000	First
4	0	1	35.0	0	0	8.0500	Third

Missing values per column

total survived: 0  
Pclass: 0

(Sex: 0 female, 0 male)

age: 177

sibsp: 0

Parch: 0

Duplicates removed: 0

NaNs: 0



Shot on OnePlus

df = lab\_df\_ENCODER()

titanic [ "Sea - survived"] = titanic["survived"]

slabx = standard (scaled)

titanic [ "fare - scaled "]- scale - fit transform  
(titanic [ "fare" ]))

sns . pairplot (titanic, vars = [ "Pclass", "sex",  
"age", "sibsp" ])

plt . subtitle ('pairplot of selected features')

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

sns . heatmap (cor, annot = True, cmap =  
(colman,  
font = 'serif' ))

plt . title ('Correlation Heatmap')

plt . show ()

Thus the data preprocessing and

cleaning (8.21%) , 3962000 entries  
using titanic dataset has

been (8.85%) 3962000 entries  
executed successfully.

