

DATA ANALYTICS LIFECYCLE:

DATA DISCOVERY AND PREPARATION

Aim:

To Perform data discovery and Exploratory analysis on a real-world dataset.

PROGRAM CODE:

```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn import SimpleImputer.
```

```
df = pd.read_csv('titanic.csv')
print(df.head())
print(df.isnull().sum())
print(df.describe())
sns.heatmap(df.isnull(), cmap=False,
            plt.show()
numerical_cols = ['Age', 'Fare']
df[numerical_cols] = imputer.fit_transform(df[numerical_cols])
df['Embarked'].fillna(df['Embarked'].mode()[0],
                      inplace=True)
print(df.isnull().sum())
```



Shot on OnePlus

Order and label = 16

Output: 16 rows - 8 columns - 179 rows total

PassengerId Embarked Survived Pclass

1st class 3rd class 0 0 0 3 0 0 0

(C) 0 0 1 1 1 1 1 1

3 1 3

(S) 0 1 1 1 1 1 1 1

4 1 3

(B) 0 0 0 0 0 0 0 3

Embarked ticket to 1st class = 119

PassengerId 0

(C) 0 0 0 0 0 0 0 0

Survived 0

(C) 0 0 0 0 0 0 0 0

Pclass 0

(C) 0 0 0 0 0 0 0 0

Age 0 0 0 0 0 0 0 0

Name 0

(C) 0 0 0 0 0 0 0 0

Sex 0

(C) 0 0 0 0 0 0 0 0

Embarked 0

(C) 0 0 0 0 0 0 0 0

SibSp 0

(C) 0 0 0 0 0 0 0 0

Parch 0

(C) 0 0 0 0 0 0 0 0

Ticket 0

(C) 0 0 0 0 0 0 0 0

Fare 0

(C) 0 0 0 0 0 0 0 0

for passenger stub set out

Training Data shape: (712, 8)

and second graph print format

Testing Data shape: (179, 8)

and



Shot on OnePlus

$X = \{f\} \left[\begin{array}{l} \text{'Rock'}, \text{'Age'}, \text{'Fore'}, \text{'Stbsf'}, \text{'Porch'}, \\ \text{'Emberked'}, \text{'Sex'} \end{array} \right]$

$y = d_1 [\text{'Survived'}]$

```
x = pd.get_dummies(x, drop_first=True)
```

x_train , x_test , y_train = train-test split

c x,y, test size = 0.2,

random state = 42)

Print Lf "Training Data shape: { x-

and some
rogue
eidos
train-shapey")

```
print(f"Testing Data shape: {x-test,  
shape}")
```

Thurs. data discov.

exploratory analysis has been

~~Result.~~

Thus data discovery and exploratory analysis has been

