```
# -*- coding: utf-8 -*-
"""Practical_2.ipynb

Automatically generated by Colaboratory.

Original file is located at
    https://colab.research.google.com/drive/1vXs-4Vf2_ANJooe6uL0gz8lyJ_oQ8b47
"""

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sea

df=pd.read_csv('prac_1.csv')

df

df.shape

df.describe

df.info

df.isnull()

df.isna()

df.isnull().sum()

# Data Cleaning
#1. Finding Duplicated Values
df.duplicated()

#2. Filling Null Values
df['Age']=df['Age'].fillna(df['Age'].mode()[0])
df['Embarked']=df['Embarked'].fillna(df['Embarked'].mode()[0])
df['Cabin']=df['Cabin'].fillna(df['Cabin'].mode()[0])

df.isnull().sum()

# Data Reduction
df

subset=df[['PassengerId','Age','Sex','Ticket']]
subset.head()

df['Sex']

#Capitilizing First Letter In UpprCase of Sex Column
df['Sex']=df['Sex'].str.title()

df['Sex']

#Capilitizing All Letters in UpperCase of Name Column
df['Name']=df['Name'].str.upper()

df['Name']

#Capilitizing All Letters in LowerCase of Cabin Column
df['Cabin']=df['Cabin'].str.lower()

df['Cabin']

#Data Integration
df1=pd.read_csv('prac_3.csv')

df1

output=pd.merge(df,df1,left_index=True,right_index=True)

output.head()

reduced=output[['Sex','Name','Score1']]
```

```
reduced

#Data Transformation
#1. Hist
df.hist()

reduced.hist()
```