import pandas as pd

import numpy as np

import seaborn as sns

import matplotlib.pyplot as plt

df = pd.read\_csv('titanic\_dataset.csv')

print('titanic dataset is successfully loaded....\n')

print('Information of dataset:\n',df.info)

print('Shape of dataset (row \* column):',df.shape)

print('Columns name:',df.columns)

print('Total emlements in dataset:',df.size)

print('Datatype of attributes (columns):',df.dtypes)

print('First 5 rows:\n',df.head().T)

print('last 5 rows:\n',df.tail().T)

print('Any 5 rows:\n',df.sample(5).T)

print ('missing values')

print(df.isnull().sum())

df['Age'].fillna(df['Age'].median(),inplace=True)

print('null values are:\n',df.isnull().sum())

fig,axes=plt.subplots(1,2)

fig.suptitle('Boxplot of 1 variable')

sns.boxplot(data=df,x='Age',ax=axes[0])

sns.boxplot(data=df,x='Fare',ax=axes[1])

plt.show()

fig,axes=plt.subplots(2,2)

fig.suptitle('Boxplot of 2 variables')

sns.boxplot(data=df,x='Survived',y='Age',hue='Survived',ax=axes[0,0])

sns.boxplot(data=df,x='Survived',y='Fare',hue='Survived',ax=axes[0,1])

sns.boxplot(data=df,x='Sex',y='Age',hue='Sex',ax=axes[1,0])

sns.boxplot(data=df,x='Sex',y='Fare',hue='Sex',ax=axes[1,1])

plt.show()

fig,axes=plt.subplots(1,2)

fig.suptitle('Boxplot of 3 variables')

sns.boxplot(data=df,x='Sex',y='Age',hue='Survived',ax=axes[0])

sns.boxplot(data=df,x='Sex',y='Fare',hue='Survived',ax=axes[1])

plt.show()