Mids1

pip install seaborn

import pandas as pd

import seaborn as sns

from sklearn.preprocessing import LabelEncoder, StandardScaler

df = pd.read\_csv('Downloads/Titanic-Dataset - Titanic-Dataset.csv')

df.head()

df.info()

df.isnull().sum()

df.describe(include='all')

df.drop(columns=['Cabin'],axis=1,inplace=True)

df.drop(columns=['PassengerId','Name', 'Ticket'],axis=1,inplace=True)

df.sample(5)

df.describe()

df['Family']=df['SibSp']+df['Parch']+1

df.head()

df.fillna({"Age": df["Age"].median()}, inplace=True)

df.fillna({"Embarked": df["Embarked"].mode()[0]}, inplace=True)

le = LabelEncoder()

df['Sex'] = le.fit\_transform(df['Sex'])

df = pd.get\_dummies(df, columns=['Embarked'], drop\_first=True)

scaler = StandardScaler()

df[['Age', 'Fare']] = scaler.fit\_transform(df[['Age', 'Fare']])

corm = df.corr(numeric\_only=True)

print('Correlation Matrix:')

print(corm['Survived'].sort\_values(ascending=False))

print('\n')

from sklearn.ensemble import RandomForestClassifier

model = RandomForestClassifier()

x = df.drop('Survived', axis=1)

y = df['Survived']

model.fit(x , y)

feature\_importances = pd.Series(model.feature\_importances\_, index=x.columns)

print('Feature Importances:')

print(feature\_importances.sort\_values(ascending=False))