

```

# Data Cleaning and Preprocessing - Titanic Dataset
# Beginner-level code with basic steps

import pandas as pd
import numpy as np
from sklearn.preprocessing import StandardScaler

# Load sample Titanic data
data = {
    'survived': [1, 0, 1, 0, 1],
    'sex': ['female', 'male', 'female', 'male', 'female'],
    'age': [22, 35, np.nan, 28, 19],
    'fare': [7.25, 71.83, 8.05, 8.46, np.nan],
    'embarked': ['S', 'C', 'S', np.nan, 'Q'],
    'class': ['Third', 'First', 'Third', 'Second', 'Third'],
    'deck': ['C', np.nan, np.nan, 'E', np.nan]
}
df = pd.DataFrame(data)

# Handle missing values
df['age'].fillna(df['age'].mean(), inplace=True)
df['embarked'].fillna(df['embarked'].mode()[0], inplace=True)
df.drop(columns=['deck'], inplace=True)

# Encode categorical columns
df['sex'] = df['sex'].map({'male': 0, 'female': 1})
df['embarked'] = df['embarked'].map({'S': 0, 'C': 1, 'Q': 2})
df['class'] = df['class'].map({'Third': 3, 'Second': 2, 'First': 1})

# Normalize numerical data
scaler = StandardScaler()
df[['age', 'fare']] = scaler.fit_transform(df[['age', 'fare']])

# Remove outliers in age column
df = df[(df['age'] > -2) & (df['age'] < 2)]

# Save cleaned dataset
df.to_csv("cleaned_titanic.csv", index=False)

```