

```
1 C:\Users\kavan\PyCharmMiscProject\venv\Scripts\python.exe "C:/Users/kavan\  
2 PyCharmMiscProject\PYTHON INTERNSHIP\task 5.py"  
3 ===== INITIAL DATA PREVIEW =====  
4          PassengerId  Survived  Pclass   ...  Fare Embarked Cabin  
5 0             1         0      3   ...  7.2500      S   NaN  
6 1             2         1      1   ...  71.2833      C   C85  
7 2             3         1      3   ...  7.9250      S   NaN  
8 3             4         1      1   ...  53.1000      S   C123  
9 4             5         0      3   ...  8.0500      S   NaN  
10 [5 rows x 11 columns]  
11  
12 ===== DATASET INFO =====  
13 <class 'pandas.core.frame.DataFrame'>  
14 RangeIndex: 20 entries, 0 to 19  
15 Data columns (total 11 columns):  
16 #   Column    Non-Null Count   Dtype  
17 ---  --  --  
18 0   PassengerId    20 non-null   int64  
19 1   Survived       20 non-null   int64  
20 2   Pclass          20 non-null   int64  
21 3   Name            20 non-null   object  
22 4   Sex              20 non-null   object  
23 5   Age             17 non-null   float64  
24 6   SibSp           20 non-null   int64  
25 7   Parch           20 non-null   int64  
26 8   Fare            20 non-null   float64
```

```
27 9 Embarked 20 non-null object
28 10 Cabin 4 non-null object
29 dtypes: float64(2), int64(5), object(4)
30 memory usage: 1.8+ KB
31 None
32
33 ===== MISSING VALUES (BEFORE CLEANING) =====
34 PassengerId 0
35 Survived 0
36 Pclass 0
37 Name 0
38 Sex 0
39 Age 3
40 SibSp 0
41 Parch 0
42 Fare 0
43 Embarked 0
44 Cabin 16
45 dtype: int64
```

C:\Users\Kavan\PyCharmMiscProject\PYTHON INTERNSHIP\task 5.py:31: FutureWarning: A value is trying to be set on a copy of a DataFrame or Series through chained assignment using an inplace method.

The behavior will change in pandas 3.0. This inplace method will never work because the intermediate object on which we are setting values always behaves as a copy.

For example, when doing 'df[col].method(value, inplace=True)', try using 'df.method({col: value}, inplace=True)' or df[col] = df[col].method(value) instead, to perform

```
49 the operation inplace on the original object.  
50  
51 df['Age'].fillna(df['Age'].median(), inplace=True)  
52 C:\Users\kavan\PyCharmMiscProject\PYTHON INTERNSHIP\task 5.py:34: FutureWarning: A  
53 value is trying to be set on a copy of a DataFrame or Series through chained  
54 assignment using an inplace method.  
55 The behavior will change in pandas 3.0. This inplace method will never work because  
56 the intermediate object on which we are setting values always behaves as a copy.  
57  
58 For example, when doing 'df[col].method(value, inplace=True)', try using 'df.method  
59 ({col: value}, inplace=True)' or df[col] = df[col].method(value) instead, to perform  
60 the operation inplace on the original object.  
61  
62 df['Embarked'].fillna(df['Embarked'].mode()[0], inplace=True)  
63 Duplicates removed: 0  
64  
65 ===== MISSING VALUES (AFTER CLEANING) =====  
66 PassengerId 0  
67 Survived 0  
68 Pclass 0  
69 Name 0  
70 Sex 0  
71 Age 0  
72 SibSp 0
```

```
71 Parch      0
72 Fare       0
73 Embarked   0
74 Age_Group  0
75 dtype: int64
76
77 ===== FINAL DATASET INFO =====
78 <class 'pandas.core.frame.DataFrame'>
79 RangeIndex: 20 entries, 0 to 19
80 Data columns (total 11 columns):
81 #   Column      Non-Null Count   Dtype  
82 ---  --          --          --      
83 0   PassengerId 20 non-null     int64  
84 1   Survived    20 non-null     category
85 2   Pclass      20 non-null     category
86 3   Name        20 non-null     object  
87 4   Sex         20 non-null     object  
88 5   Age         20 non-null     float64
89 6   SibSp      20 non-null     int64  
90 7   Parch      20 non-null     int64  
91 8   Fare        20 non-null     float64
92 9   Embarked   20 non-null     object  
93 10  Age_Group  20 non-null     object  
94 dtypes: category(2), float64(2), int64(3), object(4)
95 memory usage: 1.8+ KB
96 None
97
```

```
98 ===== FINAL DATA PREVIEW =====
99   PassengerId Survived Pclass ...
100  0            1          0    3    ...
101  1            2          1    1    ...
102  2            3          1    3    ...
103  3            4          1    1    ...
104  4            5          0    3    ...
105
106 [5 rows x 11 columns]
107
108 ✎ Cleaned dataset saved as 'cleaned_data.csv'
109
110 Process finished with exit code 0
111
```