

```
import pandas as pd
import numpy as np

df = pd.read_csv("Students.csv")

print(df, "\n")

print(df.describe(include='all'), "\n")

print(df['age'], "\n")

df.rename(columns={'age': 'student_age'}, inplace=True)
print(df.columns, "\n")

df['marks'] = pd.to_numeric(df['marks'], errors='coerce')

mean_marks = df['marks'].mean()
df['marks'] = df['marks'].fillna(mean_marks)

median_age = df['student_age'].median()
df['student_age'] = df['student_age'].fillna(median_age)

df = df.drop_duplicates()
print(df, "\n")

print(df.head(5), "\n")

print(df.tail(5), "\n")

print(df.sample(3), "\n")

df_cleaned = df.dropna()

print(df_cleaned, "\n")
df_cleaned.to_csv("Cleaned_Students.csv", index=False)
```

Outputs:

● PS D:\Internship\Day16> **python task.py**

```
    id  name  age  marks      city
0   1   Asha  21.0   85  Mysore
1   2   Ravi  22.0   90  Bangalore
2   3   John  NaN    78  Mangalore
3   4   Meena 21.0   abc   Delhi
4   5   Kiran  23.0   88  Chennai
5   6   Ravi  22.0   90  Bangalore
6   7   Pooja  NaN    95  Mysore
7   8   Arun  24.0   70  Hyderabad
8   3   John  NaN    78  Mangalore
```

```
          id  name  age  marks      city
count  9.000000    9  6.000000    9      9
unique  NaN        7  NaN        7      6
top     NaN        Ravi  NaN        90  Mysore
freq    NaN        2  NaN        2      2
mean   4.333333  NaN  22.166667  NaN  NaN
std    2.345208  NaN  1.169045  NaN  NaN
min   1.000000  NaN  21.000000  NaN  NaN
25%  3.000000  NaN  21.250000  NaN  NaN
50%  4.000000  NaN  22.000000  NaN  NaN
75%  6.000000  NaN  22.750000  NaN  NaN
max   8.000000  NaN  24.000000  NaN  NaN
```

```
0   21.0
1   22.0
2   NaN
3   21.0
4   23.0
5   22.0
6   NaN
7   24.0
8   NaN
```

Name: age, dtype: float64

Index(['id', 'name', 'student_age', 'marks', 'city'], dtype='object')

```
    id  name  student_age  marks      city
0   1   Asha  21.0  85.00  Mysore
1   2   Ravi  22.0  90.00  Bangalore
2   3   John  22.0  78.00  Mangalore
3   4   Meena 21.0  84.25  Delhi
4   5   Kiran  23.0  88.00  Chennai
5   6   Ravi  22.0  90.00  Bangalore
6   7   Pooja  22.0  95.00  Mysore
7   8   Arun  24.0  70.00  Hyderabad
```

```
    id  name  student_age  marks      city
0   1   Asha  21.0  85.00  Mysore
1   2   Ravi  22.0  90.00  Bangalore
2   3   John  22.0  78.00  Mangalore
3   4   Meena 21.0  84.25  Delhi
4   5   Kiran  23.0  88.00  Chennai
```

```
    id  name  student_age  marks      city
3   4   Meena 21.0  84.25  Delhi
4   5   Kiran  23.0  88.00  Chennai
5   6   Ravi  22.0  90.00  Bangalore
6   7   Pooja  22.0  95.00  Mysore
7   8   Arun  24.0  70.00  Hyderabad
```

	id	name	student_age	marks	city
2	3	John	22.0	78.0	Mangalore
4	5	Kiran	23.0	88.0	Chennai
0	1	Asha	21.0	85.0	Mysore

	id	name	student_age	marks	city
0	1	Asha	21.0	85.00	Mysore
1	2	Ravi	22.0	90.00	Bangalore
2	3	John	22.0	78.00	Mangalore
3	4	Meena	21.0	84.25	Delhi
4	5	Kiran	23.0	88.00	Chennai
5	6	Ravi	22.0	90.00	Bangalore
6	7	Pooja	22.0	95.00	Mysore
7	8	Arun	24.0	70.00	Hyderabad

Day16 > Cleaned_Students.csv > data

```

1 id,name,student_age,marks,city
2 1,Asha,21.0,85.0,Mysore
3 2,Ravi,22.0,90.0,Bangalore
4 3,John,22.0,78.0,Mangalore
5 4,Meena,21.0,84.25,Delhi
6 5,Kiran,23.0,88.0,Chennai
7 6,Ravi,22.0,90.0,Bangalore
8 7,Pooja,22.0,95.0,Mysore
9 8,Arun,24.0,70.0,Hyderabad
10

```