_____DATA FRAME FILE HANDLING ASSIGNMENTS_____ :

1. ASSIGNMENT:[READ CSV FILE USING READ_CSV()]

In [2]:
```
import pandas as pd
df = pd.read_csv("Students.csv")
print(df.head())
```

```
    name  age  marks
0  ramiz   21     88
1   aman   22     76
2   neha   23     95
3   zara   21     92
```

2. ASSIGNMENT :[EXPORT DATAFRAME TO CSV USING TO CSV()]

In [8]:
```
import pandas as pd

date = {
       'name': ['ramiz', 'neha', 'koko'],
       'age': [21, 22, 23],

        'course': ['ai', 'web dev', 'ml']
}

df =pd.DataFrame(date)
df.to_csv("Fuck_hole.csv", index=False)
df_read =pd.read_csv("Fuck_hole.csv")
print(df_read)
```

```
    name  age   course
0  ramiz   21       ai
1   neha   22  web dev
2   koko   23       ml
```

3 ASSIGNMENT [HANDLE CUSTOM SEPARATOR IN CSV FILE(PIPE SEPARATED)]

In [19]:
```
import pandas as pd

date = {

        'name': ['ramiz', 'neha', 'sahil', 'suman'],

         'age': [21, 22, 23, 24],


         'course': ['ai', 'ml', 'web dev', 'data S']

}

df =pd.DataFrame(date)

print("\nThe Created the Csv file: ")

df.to_csv("Student_2.csv",sep='|', index=False)

print("\n Showin the csv file:")

df_read =pd.read_csv("Student_2.csv", sep='|')


print(df_read.head())
```

The Created the Csv file:

```
 Showin the csv file:
    name  age   course
0  ramiz   21       ai
1   neha   22       ml
2  sahil   23  web dev
3  suman   24   data S
```

4 ASSIGNMENT [MIISSING VALUE CHICK WITH (ISNULL) .SUM()]

In [35]:
```
import pandas as pd

df.to_csv("Student.csv", index=False)

print(" Data Info ")

print(df.info())

print("Data Summary ")

print(df.describe())

print("Showin the number of rows \n")

print(df.shape[0])

print("Showin the number of columns  \n")

print(df.shape[1])

print("showin the null value in the data libary :\n")


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

```
 Data Info
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4 entries, 0 to 3
Data columns (total 3 columns):
 #   Column  Non-Null Count  Dtype
---  ------  --------------  -----
 0   name    4 non-null      object
 1   age     4 non-null      int64
 2   course  4 non-null      object
dtypes: int64(1), object(2)
memory usage: 228.0+ bytes
None
Data Summary
             age
count   4.000000
mean   22.500000
std     1.290994
min    21.000000
25%    21.750000
50%    22.500000
75%    23.250000
max    24.000000
Showin the number of rows

4
Showin the number of columns

3
showin the null value in the data libary :

name      0
age       0
course    0
dtype: int64
```

5 ASSIGNMENT [MULTIPLE CSV MERGE]

In [29]:
```
import pandas as pd

df1 =pd.read_csv("Student_part1.csv")

df2 =pd.read_csv("Student_part2.csv")

final_df =pd.concat([df1, df2])

print(final_df)

final_df.to_csv("All_Students.csv", index=False)
```

```
     name  age  course
0   ramiz   21      ai
1    neha   22      ml
2  charli   23  data S
3     roy   24     web
0   suman   21      ai
1    Aman   22      ml
2   rohon   23  data S
3    gavy   24     web
```

6 ASSIGNMENT [EXCEL READ/WRITE]

In [32]:
```
import pandas as pd

data = {
      'name': ['ramiz', 'zara', 'suman', 'koko'],
       'age': [21, 22, 23, 24],
       'course': ['ai', 'ml', 'Data S', 'web Dev']
}
df =pd.DataFrame(date)
df.to_excel("Student_data1.xlsx", index=False)
df_read =pd.read_excel("Student_data1.xlsx")
print("Data Summary :")
print(df_read.describe())
```

```
Data Summary :
             age
count   4.000000
mean   22.500000
std     1.290994
min    21.000000
25%    21.750000
50%    22.500000
75%    23.250000
max    24.000000
```

7. Assignment 7 -csv (Export data processing)

In [5]:
```
import pandas as pd



date = {

       'name': ['ramiz', 'neha', 'sahil', 'koko'],

       'age': [21, 22, 23, 24],

       'course': ['ai', 'data s', 'web dev', 'game dev']
}

df =pd.DataFrame(date)

new_row = {

       'name':'arjun',

        'age': 21,

        'course': 'ai'

}
df = pd.concat([df, pd.DataFrame([new_row])], ignore_index=True)
df =df.drop('course', axis=1)
```

```
df.to_csv("Student_3.csv",index=False)

df_read =pd.read_csv("Student_3.csv")

print(df_read.head())
```

```
    name  age
0  ramiz   21
1   neha   22
2  sahil   23
3   koko   24
4  arjun   21
```

----->>> Final File Handling Mini Project

In [4]:

```
import pandas as pd
import numpy as np
date = {

      'name': ['ramiz', 'neha', 'sahil', 'koko'],

      'age': [21, 22, 23, None ],

      'course': ['ai', 'data S', 'web Dev',None],

      'marks': [88,  90,  None, None ],

      'city': ['calofoniya', 'kolkata', 'delhi',None]
}

df = pd.DataFrame(date)

print(df)

df.to_csv("Student_project.csv", index=False)


print(df)




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

print("\n Show in the data  Five Rows :")

print(df.head())

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

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


course_cor = df['course'].value_counts()

print("\n Number of Studens per course:")

print(course_cor)

course_ave =df.groupby('course') ['marks'].mean()


print("\nAverage marks per course:")

print(course_ave)


top_Student=df[df['marks']>= 90]

print("\n Top Scorers:")

print(top_Student[['name', 'marks']])

df.to_csv("Cleaned_Student_data.csv", index=False)

df_read=pd.read_csv("Cleaned_Student_data.csv")


print(df_read.head())
```

```
    name   age   course  marks        city
0  ramiz  21.0       ai   88.0  calofoniya
1   neha  22.0   data S   90.0     kolkata
2  sahil  23.0  web Dev    NaN       delhi
3   koko   NaN     None    NaN        None
    name   age   course  marks        city
0  ramiz  21.0       ai   88.0  calofoniya
1   neha  22.0   data S   90.0     kolkata
2  sahil  23.0  web Dev    NaN       delhi
3   koko   NaN     None    NaN        None

 Show in the data  Five Rows :
    name   age   course  marks        city
0  ramiz  21.0       ai   88.0  calofoniya
1   neha  22.0   data S   90.0     kolkata
2  sahil  23.0  web Dev    NaN       delhi
3   koko   NaN      NaN    NaN         NaN
    name    age  course  marks   city
0  False  False   False  False  False
1  False  False   False  False  False
2  False  False   False   True  False
3  False   True    True   True   True

 Number of Studens per course:
course
ai         1
data S     1
web Dev    1
Name: count, dtype: int64

Average marks per course:
course
ai         88.0
data S     90.0
web Dev    89.0
Name: marks, dtype: float64

 Top Scorers:
   name  marks
1  neha   90.0
    name   age   course  marks        city
0  ramiz  21.0       ai   88.0  calofoniya
1   neha  22.0   data S   90.0     kolkata
2  sahil  23.0  web Dev   89.0       delhi
3   koko   NaN      NaN   89.0         NaN
```

In [ ]: