

```
In [1]: import pandas as pd
```

```
In [2]: pd.__version__
```

```
Out[2]: '1.3.4'
```

Series Data Structure

```
In [3]: l=[1,2,3,6.0,-3,'Data Values']
print(l)
```

```
[1, 2, 3, 6.0, -3, 'Data Values']
```

```
In [4]: s1=pd.Series(l) #For creating series 1st columns represents index and 2nd column data val
print(s1)      # Its a 1D indexed array
```

```
0          1
1          2
2          3
3         6.0
4         -3
5    Data Values
dtype: object
```

```
In [5]: type(s1)
```

```
Out[5]: pandas.core.series.Series
```

```
In [6]: s2=pd.Series([1,2,3,4])
print(s2)
```

```
0    1
1    2
2    3
3    4
dtype: int64
```

```
In [7]: empty_series=pd.Series([]) # Creating empty series
print(empty_series)
```

```
Series([], dtype: float64)
```

```
C:\Users\hp\AppData\Local\Temp\ipykernel_12824/514017011.py:1: DeprecationWarning: The default dtype for empty Series will be 'object' instead of 'float64' in a future version. Specify a dtype explicitly to silence this warning.
```

```
empty_series=pd.Series([]) # Creating empty series
```

```
In [8]: s3=pd.Series([1,2,3,4],index=['a','b','c','d']) # Creating our own index but the index va
print(s3)
```

```
a    1
b    2
c    3
d    4
```

```
In [9]: s3=pd.Series([1,2,3,4],index=['a','b','c','d'],dtype=float) # Creating our own index but print(s3)
```

```
a    1.0  
b    2.0  
c    3.0  
d    4.0  
dtype: float64
```

```
In [10]: s3=pd.Series([1,2,3,4],index=['a','b','c','d'],name='Data values') # Creating our own ind print(s3)
```

```
a    1  
b    2  
c    3  
d    4  
Name: Data values, dtype: int64
```

```
In [11]: scalar_s=pd.Series(0.5)  
print(scalar_s)
```

```
0    0.5  
dtype: float64
```

```
In [12]: scalar_s=pd.Series(0.5,index=[1,2,3,4]) #Creating series of the scalar values  
print(scalar_s)
```

```
1    0.5  
2    0.5  
3    0.5  
4    0.5  
dtype: float64
```

```
In [13]: d=pd.Series({'a':1,'b':2})  
print(d)
```

```
a    1  
b    2  
dtype: int64
```

```
In [14]: d=pd.Series({'a':1,'b':2},dtype=float)  
print(d)
```

```
a    1.0  
b    2.0  
dtype: float64
```

```
In [15]: scalar_s[2]
```

```
Out[15]: 0.5
```

```
In [16]: s3[0:2]
```

```
Out[16]: a    1  
b    2  
Name: Data values, dtype: int64
```

```
s3[0:3]
```

```
Out[17]: a    1  
          b    2  
          c    3  
          Name: Data values, dtype: int64
```

```
In [18]: max(s3)
```

```
Out[18]: 4
```

```
In [19]: min(s3)
```

```
Out[19]: 1
```

```
In [20]: s3[s3>1] #Values greater than 1 in s3
```

```
Out[20]: b    2  
          c    3  
          d    4  
          Name: Data values, dtype: int64
```

```
In [21]: s4=pd.Series([1,2,3,4])  
print(s4)
```

```
0    1  
1    2  
2    3  
3    4  
dtype: int64
```

```
In [22]: s2+s4
```

```
Out[22]: 0    2  
          1    4  
          2    6  
          3    8  
          dtype: int64
```

```
In [23]: s2-s4 #These are for equal data values
```

```
Out[23]: 0    0  
          1    0  
          2    0  
          3    0  
          dtype: int64
```

```
In [24]: s2*s4 #These are for equal data values
```

```
Out[24]: 0    1  
          1    4  
          2    9  
          3   16  
          dtype: int64
```

```
In [25]: s6=pd.Series([1,2,3])  
print(s6)
```

```
0    1  
1    2  
2    3  
dtype: int64
```

```
In [26]: s4+s6 # Adding different indexes of data values and NaN stands for not a number
```

```
Out[26]: 0    2.0  
1    4.0  
2    6.0  
3    NaN  
dtype: float64
```

Data Frame: Its a 2D,size mutable,heterogenous data structure with labeled axes

```
In [27]: empt=pd.DataFrame()  
print(empt)
```

```
Empty DataFrame  
Columns: []  
Index: []
```

```
In [28]: l=['a','b','c']  
print(l)
```

```
['a', 'b', 'c']
```

```
In [29]: df1=pd.DataFrame(l)  
print(df1)
```

```
          0  
0    a  
1    b  
2    c
```

```
In [30]: df1
```

```
Out[30]: 0  
         0  a  
        1  b  
        2  c
```

```
In [31]: l1=[[1,2,3,4],[5,6,7,8],[4,5,28,5]]  
print(l1)
```

```
[[1, 2, 3, 4], [5, 6, 7, 8], [4, 5, 28, 5]]
```

```
In [32]: df2=pd.DataFrame(l1)  
print(df2)
```

```
          0  1  2  3
```

```
1 5 6 7 8  
2 4 5 28 5
```

In [33]:

```
df2
```

Out[33]:

| | 0 | 1 | 2 | 3 |
|---|---|---|----|---|
| 0 | 1 | 2 | 3 | 4 |
| 1 | 5 | 6 | 7 | 8 |
| 2 | 4 | 5 | 28 | 5 |

In [34]:

```
d1={'ID':[11,22,33,44]}  
d1
```

Out[34]:

```
{'ID': [11, 22, 33, 44]}
```

In [35]:

```
df3=pd.DataFrame(d1)  
df3
```

Out[35]:

| | ID |
|---|----|
| 0 | 11 |
| 1 | 22 |
| 2 | 33 |
| 3 | 44 |

In [36]:

```
ls_dict=[{'a':1,'b':2},{'c':3,'d':4}]  
ls_dict
```

Out[36]:

```
[{'a': 1, 'b': 2}, {'c': 3, 'd': 4}]
```

In [37]:

```
d4=pd.DataFrame(ls_dict)  
d4
```

Out[37]:

| | a | b | c | d |
|---|-----|-----|-----|-----|
| 0 | 1.0 | 2.0 | NaN | NaN |
| 1 | NaN | NaN | 3.0 | 4.0 |

In [38]:

```
ls_dict1=[{'a':1,'b':2},{'a':3,'b':4}]  
d5=pd.DataFrame(ls_dict1)  
d5
```

Out[38]:

| | a | b |
|---|---|---|
| 0 | 1 | 2 |
| 1 | 3 | 4 |

How to read CSV file

```
In [39]: pd.read_csv("E:\\New folder\\student_results.csv")
```

```
Out[39]:
```

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

```
In [40]:
```

```
import os
```

```
In [41]:
```

```
print(os.getcwd())
```

C:\\Users\\hp\\NUMPY, PANDAS,ML,DEEP LEARNING TUTORIAL

How to write CSV Files

```
In [42]:
```

```
df=pd.read_csv("E:\\New folder\\student_results.csv")  
df
```

```
Out[42]:
```

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [43]:

```
type(df)
```

Out[43]:

pandas.core.frame.DataFrame

In [44]:

```
df.columns
```

Out[44]:

```
Index(['Date', 'Student ID', 'Class', 'Study hrs', 'Sleeping hrs', 'Social Media usage hrs', 'Mobile Games hrs', 'Percanage'],
      dtype='object')
```

In [45]:

```
df=pd.read_csv("E:\\New folder\\student_results.csv",nrows=1) #Printing data of specific row
```

Out[45]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |

In [46]:

```
df=pd.read_csv("E:\\New folder\\student_results.csv",nrows=3) #Printing data of specific rows
```

Out[46]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |

In [47]:

```
df=pd.read_csv("E:\\New folder\\student_results.csv",usecols=[0,1,2]) #Printing data of specific columns
```

Out[47]:

| | Date | Student ID | Class |
|---|------------|------------|-------|
| 0 | 01-12-2021 | 1001 | 10 |
| 1 | 02-12-2021 | 1002 | 10 |
| 2 | 03-12-2021 | 1003 | 10 |
| 3 | 04-12-2021 | 1004 | 10 |

Date Student ID Class

| | | | |
|---|------------|------|----|
| 4 | 05-12-2021 | 1005 | 11 |
| 5 | 06-12-2021 | 1006 | 11 |
| 6 | 07-12-2021 | 1007 | 12 |
| 7 | 08-12-2021 | 1008 | 12 |
| 8 | 09-12-2021 | 1009 | 12 |
| 9 | 10-12-2021 | 1010 | 12 |

In [48]:

```
df=pd.read_csv("E:\\New folder\\student_results.csv",skiprows=3) #Skiping a specific row  
df
```

Out[48]:

| | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
|---|------------|------|----|----|---|---|---|----|
| 0 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 1 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 2 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 3 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 4 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 5 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 6 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [49]:

```
df=pd.read_csv("E:\\New folder\\student_results.csv",skiprows=[2]) #Skiping a specific row  
df
```

Out[49]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 2 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 3 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 4 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 5 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 6 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 7 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 8 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [50]:

```
df=pd.read_csv("E:\\New folder\\student_results.csv")  
df
```

Out[50]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [51]:

```
df=pd.read_csv("E:\\New folder\\student_results.csv",header=1) #A header necessarily stored in the first row
df
```

Out[51]:

| | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
|---|------------|------|----|----|---|---|---|----|
| 0 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 1 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 2 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 3 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 4 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 5 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 6 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 7 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 8 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [52]:

```
df=pd.read_csv("E:\\New folder\\student_results.csv",header=None)
df
```

Out[52]:

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
| 1 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 2 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |

| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|----|------------|------|----|----|---|---|---|----|
| 3 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 4 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 5 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 6 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 7 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 8 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 9 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 10 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [53]:

```
df=pd.read_csv("E:\\New folder\\student_results.csv",header=None,prefix='Columns')# Given
df
```

Out[53]:

| Columns0 | Columns1 | Columns2 | Columns3 | Columns4 | Columns5 | Columns6 | Columns7 | |
|----------|------------|------------|----------|-----------|--------------|------------------------|------------------|-----------|
| 0 | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
| 1 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 2 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 3 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 4 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 5 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 6 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 7 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 8 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 9 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 10 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [54]:

```
df=pd.read_csv("E:\\New folder\\student_results.csv",header=None,prefix='Col')
df
```

Out[54]:

| Col0 | Col1 | Col2 | Col3 | Col4 | Col5 | Col6 | Col7 | |
|------|------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |

| Col0 | Col1 | Col2 | Col3 | Col4 | Col5 | Col6 | Col7 |
|------|------------|------|------|------|------|------|------|
| 1 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 |
| 2 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 |
| 3 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 |
| 4 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 |
| 5 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 |
| 6 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 |
| 7 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 |
| 8 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 |
| 9 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 |
| 10 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 |

In [55]:

```
df=pd.read_csv("E:\\New folder\\student_results.csv",names=['a','b','c','d','e','f','g','h'])
df
```

Out[55]:

| a | b | c | d | e | f | g | h |
|----|------------|------------|-------|-----------|--------------|------------------------|------------------|
| 0 | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs |
| 1 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 |
| 2 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 |
| 3 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 |
| 4 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 |
| 5 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 |
| 6 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 |
| 7 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 |
| 8 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 |
| 9 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 |
| 10 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 |

In [56]:

```
df=pd.read_csv("E:\\New folder\\student_results.csv")
```

Out[56]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [57]:

df.head() #Prints Top 5 rows

Out[57]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |

In [58]:

df.head(2)

Out[58]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |

In [59]:

df.head(7)

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |

In [60]: `df.tail() # Default prints last 5 rows`

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [61]: `df.tail(1)`

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [62]: `df.tail(4)`

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |

| Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage | |
|------|------------|-------|-----------|--------------|------------------------|------------------|-----------|----|
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [63]:

```
df=pd.read_csv("E:\\New folder\\student_results.csv")
df
```

Out[63]:

| Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage | |
|------|------------|-------|-----------|--------------|------------------------|------------------|-----------|----|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [64]:

```
df=pd.read_csv("E:\\New folder\\student_results.csv",dtype={'Class':'float64'}) # Changing the data type of Class column
df
```

Out[64]:

| Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage | |
|------|------------|-------|-----------|--------------|------------------------|------------------|-----------|----|
| 0 | 01-12-2021 | 1001 | 10.0 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10.0 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10.0 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10.0 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11.0 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11.0 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12.0 | 4 | 6 | 0 | 0 | 75 |

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 7 | 08-12-2021 | 1008 | 12.0 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12.0 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12.0 | 6 | 9 | 1 | 0 | 85 |

In [65]:

```
df=pd.read_csv("E:\\New folder\\student_results.csv",dtype={'Class':'float64','Social Med
df
```

Out[65]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10.0 | 2 | 9 | 3.0 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10.0 | 6 | 8 | 2.0 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10.0 | 3 | 8 | 2.0 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10.0 | 0 | 8 | 1.0 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11.0 | 4 | 7 | 2.0 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11.0 | 10 | 7 | 0.0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12.0 | 4 | 6 | 0.0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12.0 | 10 | 6 | 2.0 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12.0 | 2 | 8 | 2.0 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12.0 | 6 | 9 | 1.0 | 0 | 85 |

In [66]:

```
df=pd.read_csv("E:\\New folder\\student_results.csv")
df
```

Out[66]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [67]:

```
df=pd.read_csv("E:\\New folder\\student_results.csv",true_values=['Yes']) # Changing Stri
df
```

Out[67]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [68]:

```
df=pd.read_csv("E:\\New folder\\student_results.csv")
df
```

Out[68]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [69]:

```
df=pd.read_csv("E:\\New folder\\student_results.csv",true_values=['Yes'],false_values=['N
df
```

Out[69]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

Handling missing values

In [70]:

```
df=pd.read_csv("E:\\New folder\\student_results.csv")
df
```

Out[70]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|--|------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
|--|------|------------|-------|-----------|--------------|------------------------|------------------|-----------|

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [71]:

```
df=pd.read_csv("E:\\New folder\\student_results.csv") #Missing values are considered as N
df
```

Out[71]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [72]:

```
df=pd.read_csv("E:\\New folder\\student_results.csv") #Missing values are considered as N
df
```

Out[72]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [73]:

```
df=pd.read_csv("E:\\New folder\\student_results.csv")
df
```

Out[73]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [74]:

```
df=pd.read_csv("E:\\New folder\\student_results.csv",na_values='Hi how r u?') # MAKING A
```

Out[74]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [75]:

```
df=pd.read_csv("E:\\New folder\\student_results.csv",na_values=['Hi how r u?','Hello']) #
```

Out[75]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [76]:

```
df=pd.read_csv("E:\\New folder\\student_results.csv")
```

Out[76]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [77]:

```
df=pd.read_csv("E:\\New folder\\student_results.csv",na_values={'Class':'Hello'}) #MAking df
```

Out[77]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [78]:

```
df=pd.read_csv("E:\\New folder\\student_results.csv")
df
```

Out[78]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [79]:

```
df=pd.read_csv("E:\\New folder\\student_results.csv",keep_default_na=False) #Keeping defa
df
```

Out[79]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [80]:

```
df=pd.read_csv("E:\\New folder\\student_results.csv",na_filter=False) #Detect missing val
df
```

Out[80]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [81]:

```
df.isnull() # Means there are no null values
```

Out[81]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|-------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | False | False | False | False | False | False | False | False |
| 1 | False | False | False | False | False | False | False | False |
| 2 | False | False | False | False | False | False | False | False |
| 3 | False | False | False | False | False | False | False | False |
| 4 | False | False | False | False | False | False | False | False |
| 5 | False | False | False | False | False | False | False | False |
| 6 | False | False | False | False | False | False | False | False |
| 7 | False | False | False | False | False | False | False | False |
| 8 | False | False | False | False | False | False | False | False |
| 9 | False | False | False | False | False | False | False | False |

In [82]:

```
df=pd.read_csv("E:\\New folder\\student_results.csv")
df
```

Out[82]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percantage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [83]:

```
df.isnull()
```

Out[83]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percantage |
|---|-------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | False | False | False | False | False | False | False | False |
| 1 | False | False | False | False | False | False | False | False |
| 2 | False | False | False | False | False | False | False | False |
| 3 | False | False | False | False | False | False | False | False |
| 4 | False | False | False | False | False | False | False | False |
| 5 | False | False | False | False | False | False | False | False |
| 6 | False | False | False | False | False | False | False | False |
| 7 | False | False | False | False | False | False | False | False |
| 8 | False | False | False | False | False | False | False | False |
| 9 | False | False | False | False | False | False | False | False |

In [84]:

```
df.isnull().sum() #Counting null values in every column
```

Out[84]:

| | |
|------------------------|---|
| Date | 0 |
| Student ID | 0 |
| Class | 0 |
| Study hrs | 0 |
| Sleeping hrs | 0 |
| Social Media usage hrs | 0 |
| Mobile Games hrs | 0 |
| Percantage | 0 |

dtype: int64

In [85]:

```
df.isnull().sum().sum() #Counting total null values
```

Out[85]:

0

```
In [86]: df=pd.read_csv("E:\\New folder\\student_results.csv")
df
```

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

```
In [87]: df.notnull() #Opposite to isnull
```

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | True | True | True | True | True | True | True | True |
| 1 | True | True | True | True | True | True | True | True |
| 2 | True | True | True | True | True | True | True | True |
| 3 | True | True | True | True | True | True | True | True |
| 4 | True | True | True | True | True | True | True | True |
| 5 | True | True | True | True | True | True | True | True |
| 6 | True | True | True | True | True | True | True | True |
| 7 | True | True | True | True | True | True | True | True |
| 8 | True | True | True | True | True | True | True | True |
| 9 | True | True | True | True | True | True | True | True |

```
In [88]: df.notnull().sum()
```

```
Out[88]: Date          10
          Student ID  10
          Class         10
          Study hrs    10
          Sleeping hrs 10
          Social Media usage hrs 10
```

```
Percanage          10  
dtype: int64
```

```
In [89]: df.notnull().sum().sum()
```

```
Out[89]: 80
```

```
In [90]: #SERIES  
import numpy as np  
sr=pd.Series([1,2,3,np.nan,4,np.NAN])  
sr
```

```
Out[90]: 0    1.0  
1    2.0  
2    3.0  
3    NaN  
4    4.0  
5    NaN  
dtype: float64
```

```
In [91]: sr.isnull()
```

```
Out[91]: 0    False  
1    False  
2    False  
3    True  
4    False  
5    True  
dtype: bool
```

```
In [92]: sr.isnull().sum()
```

```
Out[92]: 2
```

```
In [93]: df=pd.read_csv("E:\\New folder\\student_results.csv")  
df
```

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |

| Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 |

In [94]:

```
df.dropna() # Drops rows which contains null values
```

Out[94]:

| Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 |

In [95]:

```
df.dropna(axis =1) # Droping null values columns Default for axis is 0 which means row
```

Out[95]:

| Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 75 |

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [96]: `df.dropna(how='any') #Removes null rows`

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [97]: `df.dropna(how='all') #Works only when all values are null`

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [98]:

```
df.dropna(thresh =1) #Printing rows which has atleast 1 NAN Value
```

Out[98]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [99]:

```
df.dropna(thresh =1, axis=1)
```

Out[99]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [100]:

```
df=pd.read_csv("E:\\New folder\\student_results.csv")
df
```

Out[100]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [101]:

```
df.dropna(subset=['Class']) #Dropping rows by selecting columns....removing those rows having null values in Class column
```

Out[101]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [102]:

```
df.dropna(subset=['Percanage'])
```

Out[102]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [103]:

```
df.dropna(inplace=True) #Dropping all NAN Values rows...update exsisting Data Frame
df
```

Out[103]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [104]:

```
df=pd.read_csv("E:\\New folder\\student_results.csv")
df
```

Out[104]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [105]:

```
df.fillna(0) #Filling NAN values by scalar
```

Out[105]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [106]:

```
df.fillna(0) #Filling NAN values by scalar
```

Out[106]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [107]:

```
df.fillna({'Class':'NA','Percentege':'Hello'})
```

Out[107]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [108]:

```
df.fillna(method='ffill') #Replacing NAN values with previous values
```

Out[108]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [109]:

```
df.fillna(method='pad') # Same as ffill
```

Out[109]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|--|------|------------|-------|-----------|--------------|------------------------|------------------|------------|
|--|------|------------|-------|-----------|--------------|------------------------|------------------|------------|

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [110]:

```
df.fillna(method='bfill') #Filling backward values
```

Out[110]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [111]:

```
df.fillna(method='pad',axis=0)
```

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [112...]

```
df.fillna(method='pad',axis=1)
```

Out[112...]

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentge |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [113...]

```
df.fillna(0,limit=1) #Fill only one NAN with 0
```

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [114...]

```
df.fillna(method='ffill',limit=1) #Fill only one NAN with method
```

Out[114...]

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [115...]

```
df.fillna(5,inplace=True)
df
```

Out[115...]

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [116...]

```
df=pd.read_csv("E:\\New folder\\student_results.csv")
df
```

Out[116...]

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [117...]

```
df.replace(to_replace=10,value=100) #To replace any value as per our need
```

Out[117...]

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 100 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 100 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 100 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 100 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 100 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 100 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [118...]

```
df.replace(10,100) # We can use directly we can change NaN value
```

Out[118...]

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 100 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 100 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 100 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 100 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 100 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 100 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [119...]

```
df.replace([1001,1002,1003,1004],0) #Changing specified column values as per our need
```

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percantage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 0 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 0 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 0 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 0 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [120...]

```
df.replace([1001,1002,1003,1004],[1,2,3,4]) #Changing specified column values as per our requirement
```

Out[120...]

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percantage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 2 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 3 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 4 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [121...]

```
df.replace({'Percantage':80},'none') #Changing for specified column specified value
```

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | none |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | none |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [122...]

```
df.replace({'Mobile Games hrs':0}, 'none') #Changing for specified column specified value
```

Out[122...]

In [123...]

df

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [124...]

```
df=pd.read_csv("E:\\New folder\\chord-progressions.csv")
df
```

Out[124...]

| | 1st chord | 2nd chord | 3rd chord | 4th chord | Progression Quality |
|---|-----------|-----------|-----------|-------------|---------------------|
| 6 | 4 | 1 | 5 | Alternative | NaN |
| 4 | 4 | 1 | 5 | Catchy | NaN |
| 1 | 1 | 1 | 1 | Didgeridoo | NaN |
| 1 | 6 | 4 | 5 | Dreadful | NaN |
| 1 | 6 | 2 | 5 | Dreadful | NaN |
| 1 | 6 | 2 | 4 | Endless | NaN |
| 1 | 3 | 4 | 6 | Energetic | NaN |
| 1 | 5 | 1 | 4 | Folk | NaN |
| 1 | 6 | 1 | 4 | Folk | NaN |
| 6 | 5 | 4 | 3 | Flamenco | NaN |
| 6 | 5 | 6 | 5 | Flamenco | NaN |
| 1 | 4 | 3 | 6 | Grunge | NaN |
| 2 | 5 | 1 | 6 | Jazz | NaN |
| 1 | 4 | 5 | 4 | Love | NaN |
| 1 | 4 | 1 | 5 | Memories | NaN |
| 1 | 5 | 6 | 4 | Pop | NaN |
| 1 | 6 | 3 | 7 | Pop | NaN |
| 4 | 1 | 4 | 5 | Rebellious | NaN |
| 1 | 1 | 1 | 5 | Sad | NaN |

| 1st chord | 2nd chord | 3rd chord | 4th chord | Progression | Quality |
|-----------|-----------|-----------|-----------|-------------|---------|
| 1 | 5 | 4 | 4 | Sad | NaN |
| 1 | 4 | 5 | 4 | Sad | NaN |
| 5 | 4 | 1 | 1 | Sweet | NaN |
| 1 | 4 | 1 | 4 | Simple | NaN |
| 1 | 5 | 5 | 1 | Simple | NaN |
| 1 | 4 | 1 | 4 | Wildside | NaN |
| 1 | 1 | 4 | 6 | Wistful | NaN |
| 2 | 1 | 5 | 7 | Moody | NaN |
| 2 | 1 | 7 | 6 | Moody | NaN |

In [125...]

```
df.replace(' [A-Za-z]', 0, regex=True) #Replacing all strings to 0 by regex) by default its
```

Out[125...]

| 1st chord | 2nd chord | 3rd chord | 4th chord | Progression | Quality |
|-----------|-----------|-----------|-----------|-------------|---------|
| 6 | 4 | 1 | 5 | 0 | NaN |
| 4 | 4 | 1 | 5 | 0 | NaN |
| 1 | 1 | 1 | 1 | 0 | NaN |
| 1 | 6 | 4 | 5 | 0 | NaN |
| 1 | 6 | 2 | 5 | 0 | NaN |
| 1 | 6 | 2 | 4 | 0 | NaN |
| 1 | 3 | 4 | 6 | 0 | NaN |
| 1 | 5 | 1 | 4 | 0 | NaN |
| 1 | 6 | 1 | 4 | 0 | NaN |
| 6 | 5 | 4 | 3 | 0 | NaN |
| 6 | 5 | 6 | 5 | 0 | NaN |
| 1 | 4 | 3 | 6 | 0 | NaN |
| 2 | 5 | 1 | 6 | 0 | NaN |
| 1 | 4 | 5 | 4 | 0 | NaN |
| 1 | 4 | 1 | 5 | 0 | NaN |
| 1 | 5 | 6 | 4 | 0 | NaN |
| 1 | 6 | 3 | 7 | 0 | NaN |
| 4 | 1 | 4 | 5 | 0 | NaN |
| 1 | 4 | 5 | 5 | 0 | NaN |
| 1 | 5 | 4 | 4 | 0 | NaN |
| 1 | 4 | 5 | 4 | 0 | NaN |
| 5 | 4 | 1 | 1 | 0 | NaN |
| 1 | 4 | 1 | 4 | 0 | NaN |
| 1 | 5 | 5 | 1 | 0 | NaN |
| 1 | 4 | 1 | 4 | 0 | NaN |
| 1 | 1 | 4 | 6 | 0 | NaN |

| 1st chord | 2nd chord | 3rd chord | 4th chord | Progression | Quality |
|-----------|-----------|-----------|-----------|-------------|---------|
| 2 | 1 | 5 | 7 | 0 | NaN |
| 2 | 1 | 7 | 6 | 0 | NaN |

```
In [126]: df.replace(6,100,inplace=True) #Putting 100 in place of 6  
df
```

| Out[126...] | 1st chord | 2nd chord | 3rd chord | 4th chord | Progression | Quality |
|-------------|-----------|-----------|-----------|-------------|-------------|---------|
| 6 | 4 | 1 | 5 | Alternative | | NaN |
| 4 | 4 | 1 | 5 | Catchy | | NaN |
| 1 | 1 | 1 | 1 | Didgeridoo | | NaN |
| 1 | 100 | 4 | 5 | Dreadful | | NaN |
| 1 | 100 | 2 | 5 | Dreadful | | NaN |
| 1 | 100 | 2 | 4 | Endless | | NaN |
| 1 | 3 | 4 | 100 | Energetic | | NaN |
| 1 | 5 | 1 | 4 | Folk | | NaN |
| 1 | 100 | 1 | 4 | Folk | | NaN |
| 6 | 5 | 4 | 3 | Flamenco | | NaN |
| 6 | 5 | 100 | 5 | Flamenco | | NaN |
| 1 | 4 | 3 | 100 | Grunge | | NaN |
| 2 | 5 | 1 | 100 | Jazz | | NaN |
| 1 | 4 | 5 | 4 | Love | | NaN |
| 1 | 4 | 1 | 5 | Memories | | NaN |
| 1 | 5 | 100 | 4 | Pop | | NaN |
| 1 | 100 | 3 | 7 | Pop | | NaN |
| 4 | 1 | 4 | 5 | Rebellious | | NaN |
| 1 | 4 | 5 | 5 | Sad | | NaN |
| 1 | 5 | 4 | 4 | Sad | | NaN |
| 1 | 4 | 5 | 4 | Sad | | NaN |
| 5 | 4 | 1 | 1 | Sweet | | NaN |
| 1 | 4 | 1 | 4 | Simple | | NaN |
| 1 | 5 | 5 | 1 | Simple | | NaN |
| 1 | 4 | 1 | 4 | Wildside | | NaN |
| 1 | 1 | 4 | 100 | Wistful | | NaN |
| 2 | 1 | 5 | 7 | Moody | | NaN |
| 2 | 1 | 7 | 100 | Moody | | NaN |

```
In [127]: df=pd.read_csv("E:\\New folder\\student_results.csv")  
df
```

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [128...]

```
df.interpolate() #It fills NaN values by predicting the values to be filled and doesn't g
```

Out[128...]

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [129...]

```
df=pd.read_csv("E:\\New folder\\student_results.csv",parse_dates=['Date'])
df
```

Out[129...]

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 2021-01-12 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 2021-02-12 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 2021-03-12 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 2021-04-12 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 2021-05-12 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 2021-06-12 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 2021-07-12 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 2021-08-12 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 2021-09-12 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 2021-10-12 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [130...]

```
type(df.Date[0])
```

Out[130...]

pandas._libs.tslibs.timestamps.Timestamp

In [131...]

```
df=pd.read_csv("E:\\New folder\\student_results.csv",parse_dates=['Date'],index_col='Date')
df
```

Out[131...]

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|--|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| | 2021-01-12 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| | 2021-02-12 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| | 2021-03-12 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| | 2021-04-12 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| | 2021-05-12 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| | 2021-06-12 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| | 2021-07-12 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| | 2021-08-12 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| | 2021-09-12 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |

| | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percantere |
|------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| Date | | | | | | | |

| | | | | | | | |
|------------|------|----|---|---|---|---|----|
| 2021-10-12 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |
|------------|------|----|---|---|---|---|----|

In [132]: df.interpolate(method='time') #Filled missing values

| | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percantere |
|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| Date | | | | | | | |
| 2021-01-12 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 2021-02-12 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2021-03-12 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 2021-04-12 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 2021-05-12 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 2021-06-12 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 2021-07-12 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 2021-08-12 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 2021-09-12 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 2021-10-12 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [133]: type(df.Class[1])

Out[133]: numpy.int64

In [134]: df.interpolate(method='polynomial',order=2) #Changes NaN values with order2

| | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percantere |
|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| Date | | | | | | | |
| 2021-01-12 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 2021-02-12 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2021-03-12 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 2021-04-12 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |

| Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 2021-05-12 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 2021-06-12 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 2021-07-12 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 2021-08-12 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 2021-09-12 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 2021-10-12 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [135...]

```
df.interpolate(method='polynomial',order=3) #Changes NaN values with order2
```

Out[135...]

| Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 2021-01-12 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 2021-02-12 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2021-03-12 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 2021-04-12 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 2021-05-12 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 2021-06-12 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 2021-07-12 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 2021-08-12 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 2021-09-12 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 2021-10-12 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [136...]

```
df.interpolate(method='spline',order=2) #Spline means a rectangular key filling
```

Out[136...]

| Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 2021-01-12 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |

| Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 2021-02-12 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2021-03-12 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 2021-04-12 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 2021-05-12 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 2021-06-12 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 2021-07-12 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 2021-08-12 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 2021-09-12 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 2021-10-12 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [137...]

```
df.interpolate(method='nearest') #Filling NaN by using nearest values b/w 2 values
```

Out[137...]

| Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 2021-01-12 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 2021-02-12 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2021-03-12 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 2021-04-12 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 2021-05-12 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 2021-06-12 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 2021-07-12 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 2021-08-12 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 2021-09-12 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 2021-10-12 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [138...]

```
df=pd.read_csv("E:\\New folder\\student_results.csv") #Changing date to column
```

Out[138...]

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [139...]

`df.interpolate()`

Out[139...]

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [140...]

`df.dtypes #To use axis parameter in interpolate none of the column should be of object type`

Date

object

```

Student ID          int64
Class              int64
Study hrs          int64
Sleeping hrs       int64
Social Media usage hrs  int64
Mobile Games hrs   int64
Percanage          int64
dtype: object

```

In [141]: df.interpolate(limit=1) #Default method is linear filling only 1 NaN values

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [142]: df.interpolate(limit=1,limit_direction='backward') #Default is forward and backward fills

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [143]:

```
df.interpolate(limit=1,limit_direction='both') #Filling both from upper and bottom
```

Out[143]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [144]:

```
df.interpolate(limit_area='inside') #Only fill NaNs surrounded by valid values (interpolate)
```

Out[144]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [145]:

```
df.interpolate(limit_area='outside') #Only fill NaNs outside valid values (extrapolate),
```

Out[145]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [146]:

df

Out[146]:

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [147...]

```
df.interpolate(inplace=True) #Here default method is linear inplace updates ur data frame
df
```

Out[147...]

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [148...]

df

Out[148...]

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [149...]

```
df.loc[0] #loc provides values of specific row with coulmn name
```

Out[149...]

```
Date          01-12-2021
Student ID    1001
Class          10
Study hrs      2
Sleeping hrs   9
Social Media usage hrs  3
Mobile Games hrs  5
Percanage       50
Name: 0, dtype: object
```

In [150...]

```
df.loc[4]
```

Out[150...]

```
Date          05-12-2021
Student ID    1005
Class          11
Study hrs      4
Sleeping hrs   7
Social Media usage hrs  2
Mobile Games hrs  0
Percanage       80
Name: 4, dtype: object
```

In [151...]

```
df.loc[[0,4]] #list of labels
```

Out[151...]

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |

In [152...]

```
df.loc[4,'Class']
```

Out[152...]

```
11
```

In [153...]

```
df.loc[0:3]
```

Out[153...]

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |

In [154...]

df.loc[0:3, 'Percanage']

Out[154...]

```
0    50
1    80
2    60
3    70
Name: Percanage, dtype: int64
```

In [155...]

df.loc[[False, False, True, True, True, True, True, True, True, True]]#it follows indexing...jha p

Out[155...]

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [156...]

df.loc[df['Class']<11]

Out[156...]

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |

```
In [157... df.loc[df['Class']<11,['Percantage']] #prints only percent value which is less than 11
```

```
Out[157... Percantage
```

| | |
|---|----|
| 0 | 50 |
| 1 | 80 |
| 2 | 60 |
| 3 | 70 |

```
In [158... df.iloc[[0,1]]#integer location
```

```
Out[158...      Date  Student ID  Class  Study hrs  Sleeping hrs  Social Media usage hrs  Mobile Games hrs  Percantage
```

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percantage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |

```
In [159... df.iloc[:4, 0]
```

```
Out[159... 0 01-12-2021  
1 02-12-2021  
2 03-12-2021  
3 04-12-2021  
Name: Date, dtype: object
```

```
In [160... #PANDAS groupby function is used to split the data into groups based on some criteria
```

```
#Splitting object  
#Applying a function  
#Combining the result
```

```
In [161... df=pd.read_csv("E:\\New folder\\student_results.csv")  
df
```

```
Out[161...      Date  Student ID  Class  Study hrs  Sleeping hrs  Social Media usage hrs  Mobile Games hrs  Percantage
```

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percantage |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |

| Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentage |
|--------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 7 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [162]:

```
gr1=df.groupby(by='Student ID')
gr1
```

Out[162]:

```
<pandas.core.groupby.generic.DataFrameGroupBy object at 0x000001ACE9B5FCD0>
```

In [163]:

```
gr1.groups #creates groups
```

Out[163]:

```
{1001: [0], 1002: [1], 1003: [2], 1004: [3], 1005: [4], 1006: [5], 1007: [6], 1008: [7], 1009: [8], 1010: [9]}
```

In [164]:

```
df.groupby(['Student ID','Class']).groups
```

Out[164]:

```
{(1001, 10): [0], (1002, 10): [1], (1003, 10): [2], (1004, 10): [3], (1005, 11): [4], (1006, 11): [5], (1007, 12): [6], (1008, 12): [7], (1009, 12): [8], (1010, 12): [9]}
```

In [165]:

```
for sid,values in gr1:
    print(sid)
    print(values)
```

1001

| | Date | Student ID | Class | Study hrs | Sleeping hrs | \ |
|---|------------|------------|-------|-----------|--------------|---|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | |

| | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------------------|------------------|------------|
| 0 | 3 | 5 | 50 |

1002

| | Date | Student ID | Class | Study hrs | Sleeping hrs | \ |
|---|------------|------------|-------|-----------|--------------|---|
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | |

| | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------------------|------------------|------------|
| 1 | 2 | 0 | 80 |

1003

| | Date | Student ID | Class | Study hrs | Sleeping hrs | \ |
|---|------------|------------|-------|-----------|--------------|---|
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | |

| | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------------------|------------------|------------|
| 2 | 2 | 4 | 60 |

1004

| | Date | Student ID | Class | Study hrs | Sleeping hrs | \ |
|---|------------|------------|-------|-----------|--------------|---|
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | |

| | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------------------|------------------|------------|
| 3 | 1 | 5 | 70 |

1005

| | Date | Student ID | Class | Study hrs | Sleeping hrs | \ |
|---|------------|------------|-------|-----------|--------------|---|
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | |

| | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------------------|------------------|------------|
| 4 | 2 | 0 | 80 |

```

1006
      Date Student ID Class Study hrs Sleeping hrs \
5 06-12-2021       1006    11      10            7

      Social Media usage hrs Mobile Games hrs Percentege
5                               0           0            90

1007
      Date Student ID Class Study hrs Sleeping hrs \
6 07-12-2021       1007    12      4             6

      Social Media usage hrs Mobile Games hrs Percentege
6                               0           0            75

1008
      Date Student ID Class Study hrs Sleeping hrs \
7 08-12-2021       1008    12      10            6

      Social Media usage hrs Mobile Games hrs Percentege
7                               2           0            85

1009
      Date Student ID Class Study hrs Sleeping hrs \
8 09-12-2021       1009    12      2             8

      Social Media usage hrs Mobile Games hrs Percentege
8                               2           4            60

1010
      Date Student ID Class Study hrs Sleeping hrs \
9 10-12-2021       1010    12      6             9

      Social Media usage hrs Mobile Games hrs Percentege
9                               1           0            85

```

In [166...]

```
list(gr1)
```

Out[166...]

```

[(1001,
      Date Student ID Class Study hrs Sleeping hrs \
0 01-12-2021       1001    10      2            9

      Social Media usage hrs Mobile Games hrs Percentege
0                               3           5            50 ),

(1002,
      Date Student ID Class Study hrs Sleeping hrs \
1 02-12-2021       1002    10      6            8

      Social Media usage hrs Mobile Games hrs Percentege
1                               2           0            80 ),

(1003,
      Date Student ID Class Study hrs Sleeping hrs \
2 03-12-2021       1003    10      3            8

      Social Media usage hrs Mobile Games hrs Percentege
2                               2           4            60 ),

(1004,
      Date Student ID Class Study hrs Sleeping hrs \
3 04-12-2021       1004    10      0            8

      Social Media usage hrs Mobile Games hrs Percentege
3                               1           5            70 ),

(1005,
      Date Student ID Class Study hrs Sleeping hrs \
4 05-12-2021       1005    11      4            7

      Social Media usage hrs Mobile Games hrs Percentege
4                               2           0            80 ),

```

| | Date | Student ID | Class | Study hrs | Sleeping hrs | \ |
|--------|------------------------|------------------|------------|-----------|--------------|---|
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | |
| | Social Media usage hrs | Mobile Games hrs | Percantege | | | |
| 5 | 0 | 0 | 90 | , | | |
| (1007, | | | | | | |
| | Date | Student ID | Class | Study hrs | Sleeping hrs | \ |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | |
| | Social Media usage hrs | Mobile Games hrs | Percantege | | | |
| 6 | 0 | 0 | 75 | , | | |
| (1008, | | | | | | |
| | Date | Student ID | Class | Study hrs | Sleeping hrs | \ |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | |
| | Social Media usage hrs | Mobile Games hrs | Percantege | | | |
| 7 | 2 | 0 | 85 | , | | |
| (1009, | | | | | | |
| | Date | Student ID | Class | Study hrs | Sleeping hrs | \ |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | |
| | Social Media usage hrs | Mobile Games hrs | Percantege | | | |
| 8 | 2 | 4 | 60 | , | | |
| (1010, | | | | | | |
| | Date | Student ID | Class | Study hrs | Sleeping hrs | \ |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | |
| | Social Media usage hrs | Mobile Games hrs | Percantege | | | |
| 9 | 1 | 0 | 85 |)] | | |

In [167...]

```
dict(list(gr1))
```

Out[167...]

| | Date | Student ID | Class | Study hrs | Sleeping hrs | \ |
|-------|------------------------|------------------|------------|-----------|--------------|---|
| 0001: | 01-12-2021 | 1001 | 10 | 2 | 9 | |
| | Social Media usage hrs | Mobile Games hrs | Percantege | | | |
| 0 | 3 | 5 | 50 | , | | |
| 1002: | 02-12-2021 | 1002 | 10 | 6 | 8 | |
| | Social Media usage hrs | Mobile Games hrs | Percantege | | | |
| 1 | 2 | 0 | 80 | , | | |
| 1003: | 03-12-2021 | 1003 | 10 | 3 | 8 | |
| | Social Media usage hrs | Mobile Games hrs | Percantege | | | |
| 2 | 2 | 4 | 60 | , | | |
| 1004: | 04-12-2021 | 1004 | 10 | 0 | 8 | |
| | Social Media usage hrs | Mobile Games hrs | Percantege | | | |
| 3 | 1 | 5 | 70 | , | | |
| 1005: | 05-12-2021 | 1005 | 11 | 4 | 7 | |
| | Social Media usage hrs | Mobile Games hrs | Percantege | | | |
| 4 | 2 | 0 | 80 | , | | |
| 1006: | 06-12-2021 | 1006 | 11 | 10 | 7 | |
| | Social Media usage hrs | Mobile Games hrs | Percantege | | | |
| 5 | 0 | 0 | 90 | , | | |
| 1007: | 07-12-2021 | 1007 | 12 | 4 | 6 | |

```

Social Media usage hrs  Mobile Games hrs  Percanntege
6                      0                  0      75 ,
1008:                 Date  Student ID  Class  Study hrs  Sleeping hrs \
7 08-12-2021           1008        12       10            6

Social Media usage hrs  Mobile Games hrs  Percanntege
7                      2                  0      85 ,
1009:                 Date  Student ID  Class  Study hrs  Sleeping hrs \
8 09-12-2021           1009        12       2             8

Social Media usage hrs  Mobile Games hrs  Percanntege
8                      2                  4      60 ,
1010:                 Date  Student ID  Class  Study hrs  Sleeping hrs \
9 10-12-2021           1010        12       6             9

Social Media usage hrs  Mobile Games hrs  Percanntege
9                      1                  0      85 }

```

In [168...]

```
gr=df.groupby('Class').get_group(10)#Splitting according to class 10
gr
```

Out[168...]

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanntege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |

In [169...]

```
gr=df.groupby('Percanntege').get_group(80)#Splitting
gr
```

Out[169...]

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percanntege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-------------|
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |

In [170...]

```
gr.sum() #ADDS all data according to given percantage columns and criteria of splitting
```

Out[170...]

| | |
|------------------------|------------|
| Date | 02-12-2021 |
| Student ID | 2007 |
| Class | 21 |
| Study hrs | 10 |
| Sleeping hrs | 15 |
| Social Media usage hrs | 4 |
| Mobile Games hrs | 0 |
| Percanntege | 160 |
| dtype: object | |

In [171...]

```
gr.mean()
```

```
C:\Users\hp\AppData\Local\Temp/ipykernel_12824/273744718.py:1: FutureWarning: Dropping of  
nuisance columns in DataFrame reductions (with 'numeric_only=None') is deprecated; in a fu  
ture version this will raise TypeError. Select only valid columns before calling the redu  
ction.
```

```
    gr.mean()
```

```
Out[171... Student ID          1003.5  
Class             10.5  
Study hrs         5.0  
Sleeping hrs      7.5  
Social Media usage hrs  2.0  
Mobile Games hrs  0.0  
Percantage        80.0  
dtype: float64
```

```
In [172... gr.describe()
```

| | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percantage |
|--------------|------------|-----------|-----------|--------------|------------------------|------------------|------------|
| count | 2.00000 | 2.000000 | 2.000000 | 2.000000 | 2.0 | 2.0 | 2.0 |
| mean | 1003.50000 | 10.500000 | 5.000000 | 7.500000 | 2.0 | 0.0 | 80.0 |
| std | 2.12132 | 0.707107 | 1.414214 | 0.707107 | 0.0 | 0.0 | 0.0 |
| min | 1002.00000 | 10.000000 | 4.000000 | 7.000000 | 2.0 | 0.0 | 80.0 |
| 25% | 1002.75000 | 10.250000 | 4.500000 | 7.250000 | 2.0 | 0.0 | 80.0 |
| 50% | 1003.50000 | 10.500000 | 5.000000 | 7.500000 | 2.0 | 0.0 | 80.0 |
| 75% | 1004.25000 | 10.750000 | 5.500000 | 7.750000 | 2.0 | 0.0 | 80.0 |
| max | 1005.00000 | 11.000000 | 6.000000 | 8.000000 | 2.0 | 0.0 | 80.0 |

```
In [173... gr.agg(['sum', 'max', 'mean'])
```

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percantage |
|-------------|----------------------|------------|--------|-----------|--------------|------------------------|------------------|------------|
| sum | 02-12-202105-12-2021 | 2007.0 | 21.0 | 10.0 | 15.0 | 4.0 | 0.0 | 160.0 |
| max | 05-12-2021 | 1005.0 | 11.0 | 6.0 | 8.0 | 2.0 | 0.0 | 80.0 |
| mean | | Nan | 1003.5 | 10.5 | 7.5 | 2.0 | 0.0 | 80.0 |

```
In [174... #PANDAS MERGING FUNCTION CONNECTS COLUMNS OR INDEXES IN DATAFRAME BASED ON ONE OR MORE KE
```

```
In [175... df1=pd.DataFrame({'ID':[1,2,3,4],  
                           'Class':[9,10,11,12]})  
df1
```

```
Out[175...   ID  Class
```

| | | |
|----------|---|----|
| 0 | 1 | 9 |
| 1 | 2 | 10 |
| 2 | 3 | 11 |
| 3 | 4 | 12 |

```
In [176...]: df2=pd.DataFrame({'ID':[1,2,3,5],  
                           'Name':['A','B','C','D']})  
df2
```

```
Out[176...]:   ID  Name  
0    1    A  
1    2    B  
2    3    C  
3    5    D
```

```
In [177...]: pd.merge(df1,df2) #Default join is inner join
```

```
Out[177...]:   ID  Class  Name  
0    1      9    A  
1    2     10    B  
2    3     11    C
```

```
In [178...]: pd.merge(df1,df2,on='ID')#Prints for similar ID
```

```
Out[178...]:   ID  Class  Name  
0    1      9    A  
1    2     10    B  
2    3     11    C
```

```
In [179...]: pd.merge(df1,df2,on='ID',how='left')
```

```
Out[179...]:   ID  Class  Name  
0    1      9    A  
1    2     10    B  
2    3     11    C  
3    4     12    NaN
```

```
In [180...]: pd.merge(df1,df2,on='ID',how='right')
```

```
Out[180...]:   ID  Class  Name  
0    1    9.0    A  
1    2   10.0    B  
2    3   11.0    C  
3    5    NaN    D
```

```
In [181...]: pd.merge(df1,df2,on='ID',how='outer')
```

| | ID | Class | Name |
|---|----|-------|------|
| 0 | 1 | 9.0 | A |
| 1 | 2 | 10.0 | B |
| 2 | 3 | 11.0 | C |
| 3 | 4 | 12.0 | NaN |
| 4 | 5 | NaN | D |

```
In [182... pd.merge(df1,df2,on='ID',how='outer',indicator=True)
```

| | ID | Class | Name | _merge |
|---|----|-------|------|------------|
| 0 | 1 | 9.0 | A | both |
| 1 | 2 | 10.0 | B | both |
| 2 | 3 | 11.0 | C | both |
| 3 | 4 | 12.0 | NaN | left_only |
| 4 | 5 | NaN | D | right_only |

```
In [183... df1
```

| | ID | Class |
|---|----|-------|
| 0 | 1 | 9 |
| 1 | 2 | 10 |
| 2 | 3 | 11 |
| 3 | 4 | 12 |

```
In [184... df2
```

| | ID | Name |
|---|----|------|
| 0 | 1 | A |
| 1 | 2 | B |
| 2 | 3 | C |
| 3 | 5 | D |

```
In [185... df2=pd.DataFrame({'ID':[5,6,7,8],  
                           'Name':['A','B','C','D']})  
df2
```

| | ID | Name |
|---|----|------|
| 0 | 5 | A |
| 1 | 6 | B |
| 2 | 7 | C |
| 3 | 8 | D |

```
In [186... pd.merge(df1,df2, left_index=True, right_index=True)
```

| | ID_x | Class | ID_y | Name |
|---|------|-------|------|------|
| 0 | 1 | 9 | 5 | A |
| 1 | 2 | 10 | 6 | B |
| 2 | 3 | 11 | 7 | C |
| 3 | 4 | 12 | 8 | D |

```
In [187... pd.merge(df1,df2)#cuз both ID columns are different
```

```
Out[187... ID Class Name
```

```
In [188... df1
```

```
Out[188... ID Class
```

| 0 | 1 | 9 |
|---|---|----|
| 1 | 2 | 10 |
| 2 | 3 | 11 |
| 3 | 4 | 12 |

```
In [189... df2=pd.DataFrame({'ID':[1,2,3,4],  
                           'Class':[9,10,11,12]})  
df2
```

```
Out[189... ID Class
```

| 0 | 1 | 9 |
|---|---|----|
| 1 | 2 | 10 |
| 2 | 3 | 11 |
| 3 | 4 | 12 |

```
In [190... pd.merge(df1,df2, on='ID', suffixes=('ON X AXIS', 'ON Y AXIS'))#cuз we merged on ID but clas
```

```
Out[190... ID ClassON X AXIS ClassON Y AXIS
```

| 0 | 1 | 9 | 9 |
|---|---|----|----|
| 1 | 2 | 10 | 10 |
| 2 | 3 | 11 | 11 |
| 3 | 4 | 12 | 12 |

```
In [191... #CONCAT() funvtion is use to combine series,dataframe and panel objects
```

```
In [192... srl=pd.Series([0,1,2])  
srl
```

```
Out[192... 0 0  
1 1  
2 2  
dtype: int64
```

```
In [193... sr2=pd.Series([3,4,5,6,7])  
sr2
```

```
Out[193... 0 3  
1 4  
2 5  
3 6  
4 7  
dtype: int64
```

```
In [194... pd.concat([sr1,sr2])
```

```
Out[194... 0 0  
1 1  
2 2  
0 3  
1 4  
2 5  
3 6  
4 7  
dtype: int64
```

```
In [195... df1=pd.DataFrame({'ID':[1,2,3,4], 'Name':['A','B','C','D'],  
                           'Class':[5,6,7,8]})  
df1
```

```
Out[195...   ID Name Class  
0 1 A 5  
1 2 B 6  
2 3 C 7  
3 4 D 8
```

```
In [196... df2=pd.DataFrame({'ID':[5,6,7,8], 'Name':['E','F','G','H'],  
                           'Class':[9,10,11,12]})  
df2
```

```
Out[196...   ID Name Class  
0 5 E 9  
1 6 F 10  
2 7 G 11  
3 8 H 12
```

```
In [197... pd.concat([df1,df2],ignore_index=True)
```

```
Out[197...   ID Name Class  
0 1 A 5  
1 2 B 6  
2 3 C 7  
3 4 D 8  
4 5 E 9  
5 6 F 10  
6 7 G 11  
7 8 H 12
```

| ID | Name | Class |
|----|------|-------|
| 2 | 3 | C 7 |
| 3 | 4 | D 8 |
| 4 | 5 | E 9 |
| 5 | 6 | F 10 |
| 6 | 7 | G 11 |
| 7 | 8 | H 12 |

```
In [198... pd.concat([df1,df2],axis=1)
```

| ID | Name | Class | ID | Name | Class |
|----|------|-------|----|------|-------|
| 0 | 1 | A 5 | 5 | E 9 | |
| 1 | 2 | B 6 | 6 | F 10 | |
| 2 | 3 | C 7 | 7 | G 11 | |
| 3 | 4 | D 8 | 8 | H 12 | |

```
In [199... df2=pd.DataFrame({'ID':[5,6],'Name':['E','F'],
   'Class':[9,10]})  
df2
```

| ID | Name | Class |
|----|------|-------|
| 0 | 5 | E 9 |
| 1 | 6 | F 10 |

```
In [200... pd.concat([df1,df2],ignore_index=True)
```

| ID | Name | Class |
|----|------|-------|
| 0 | 1 | A 5 |
| 1 | 2 | B 6 |
| 2 | 3 | C 7 |
| 3 | 4 | D 8 |
| 4 | 5 | E 9 |
| 5 | 6 | F 10 |

```
In [201... pd.concat([df1,df2],axis=1)
```

| ID | Name | Class | ID | Name | Class |
|----|------|-------|-----|------|--------|
| 0 | 1 | A 5 | 5 | 5.0 | E 9.0 |
| 1 | 2 | B 6 | 6 | 6.0 | F 10.0 |
| 2 | 3 | C 7 | NaN | NaN | NaN |
| 3 | 4 | D 8 | NaN | NaN | NaN |

```
In [202... pd.concat([df1,df2],axis=1,join='inner')
```

| | ID | Name | Class | ID | Name | Class |
|---|----|------|-------|----|------|-------|
| 0 | 1 | A | 5 | 5 | E | 9 |
| 1 | 2 | B | 6 | 6 | F | 10 |

```
In [203... df1=pd.DataFrame({'ID':[1,2,3,4], 'Name':['A','B','C','D'],
                           'Class':[5,6,7,8]})  
df1
```

| | ID | Name | Class |
|---|----|------|-------|
| 0 | 1 | A | 5 |
| 1 | 2 | B | 6 |
| 2 | 3 | C | 7 |
| 3 | 4 | D | 8 |

```
In [204... df2=pd.DataFrame({'ID':[5,6,7,8], 'Name':['E','F','G','H'],
                           'Class':[9,10,11,12]})  
df2
```

| | ID | Name | Class |
|---|----|------|-------|
| 0 | 5 | E | 9 |
| 1 | 6 | F | 10 |
| 2 | 7 | G | 11 |
| 3 | 8 | H | 12 |

```
In [205... pd.concat([df1,df2],keys=['df1','df2'])#keys used to provide labelling/sperating datas
```

| | ID | Name | Class |
|-----|----|------|-------|
| df1 | 0 | 1 | A 5 |
| | 1 | 2 | B 6 |
| | 2 | 3 | C 7 |
| | 3 | 4 | D 8 |
| df2 | 0 | 5 | E 9 |
| | 1 | 6 | F 10 |
| | 2 | 7 | G 11 |
| | 3 | 8 | H 12 |

```
In [206... df2=pd.DataFrame({'Marks':[40,63,91,34]})  
df2
```

| | Marks |
|---|-------|
| 0 | 40 |

Marks

| | |
|----------|----|
| 1 | 63 |
| 2 | 91 |
| 3 | 34 |

In [207...]

```
pd.concat([df1,df2],ignore_index=True,sort=False)
```

Out[207...]

| | ID | Name | Class | Marks |
|---|-----|------|-------|-------|
| 0 | 1.0 | A | 5.0 | NaN |
| 1 | 2.0 | B | 6.0 | NaN |
| 2 | 3.0 | C | 7.0 | NaN |
| 3 | 4.0 | D | 8.0 | NaN |
| 4 | NaN | NaN | NaN | 40.0 |
| 5 | NaN | NaN | NaN | 63.0 |
| 6 | NaN | NaN | NaN | 91.0 |
| 7 | NaN | NaN | NaN | 34.0 |

In [208...]

#JOIN function is used for combining the columns of two potentially differently indexed

In [209...]

```
df1=pd.DataFrame({'A':[1,2,3],'B':[10,20,30]},index=['a','b','c'])
df2=pd.DataFrame({'C':[4,5],'D':[40,50]},index=['a','b'])
display(df1,df2)
```

| | A | B |
|----------|---|----|
| a | 1 | 10 |
| b | 2 | 20 |
| c | 3 | 30 |

| | C | D |
|----------|---|----|
| a | 4 | 40 |
| b | 5 | 50 |

In [210...]

```
df1.join(df2) #NOT Joined properly cuz indexing are different
```

Out[210...]

| | A | B | C | D |
|---|---|----|-----|------|
| a | 1 | 10 | 4.0 | 40.0 |
| b | 2 | 20 | 5.0 | 50.0 |
| c | 3 | 30 | NaN | NaN |

In [211...]

```
df1.join(df2)
```

Out[211...]

| | A | B | C | D |
|--|---|---|---|---|
|--|---|---|---|---|

Loading [MathJax]/extensions/Safe.js

| | A | B | C | D |
|---|---|----|-----|------|
| a | 1 | 10 | 4.0 | 40.0 |
| b | 2 | 20 | 5.0 | 50.0 |
| c | 3 | 30 | NaN | NaN |

In [212...]: df1.join(df2) #default how is left join

Out[212...]:

| | A | B | C | D |
|---|---|----|-----|------|
| a | 1 | 10 | 4.0 | 40.0 |
| b | 2 | 20 | 5.0 | 50.0 |
| c | 3 | 30 | NaN | NaN |

In [213...]: df1.join(df2, how='right')

Out[213...]:

| | A | B | C | D |
|---|---|----|---|----|
| a | 1 | 10 | 4 | 40 |
| b | 2 | 20 | 5 | 50 |

In [214...]: df1.join(df2, how='inner')

Out[214...]:

| | A | B | C | D |
|---|---|----|---|----|
| a | 1 | 10 | 4 | 40 |
| b | 2 | 20 | 5 | 50 |

In [215...]: df1.join(df2, how='outer')

Out[215...]:

| | A | B | C | D |
|---|---|----|-----|------|
| a | 1 | 10 | 4.0 | 40.0 |
| b | 2 | 20 | 5.0 | 50.0 |
| c | 3 | 30 | NaN | NaN |

In [216...]: #APPEND function is used to append rows of other dataframe to the end of the given dataframe

In [219...]: df1=pd.DataFrame({'A':[1,2,3], 'B':[10,20,30]})
df2=pd.DataFrame({'A':[4,5,6], 'B':[40,50,60]})
display(df1,df2)

| | A | B |
|---|---|----|
| 0 | 1 | 10 |
| 1 | 2 | 20 |
| 2 | 3 | 30 |

| | A | B |
|---|---|----|
| 0 | 4 | 40 |
| 1 | 5 | 50 |
| 2 | 6 | 60 |

```
In [221]: df1.append(df2, ignore_index=True)
```

| | A | B |
|---|---|----|
| 0 | 1 | 10 |
| 1 | 2 | 20 |
| 2 | 3 | 30 |
| 3 | 4 | 40 |
| 4 | 5 | 50 |
| 5 | 6 | 60 |

```
In [222]: #PANDAS PIVOT TABLE
#Pivot table in pandas is an excellent tool to summarize one or more numeric variable based on categories
#Pivot tables in pandas are popularly seen in MS Excel files
```

```
In [223]: df=pd.read_csv("E:\\New folder\\student_results.csv")
df
```

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentge |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|-----------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

```
In [224]: df.pivot_table(index='Date') #Index changed to date
```

Out[224...]

| | Class | Mobile Games hrs | Percanage | Sleeping hrs | Social Media usage hrs | Student ID | Study hrs |
|------------|-------|------------------|-----------|--------------|------------------------|------------|-----------|
| Date | | | | | | | |
| 01-12-2021 | 10 | 5 | 50 | 9 | 3 | 1001 | 2 |
| 02-12-2021 | 10 | 0 | 80 | 8 | 2 | 1002 | 6 |
| 03-12-2021 | 10 | 4 | 60 | 8 | 2 | 1003 | 3 |
| 04-12-2021 | 10 | 5 | 70 | 8 | 1 | 1004 | 0 |
| 05-12-2021 | 11 | 0 | 80 | 7 | 2 | 1005 | 4 |
| 06-12-2021 | 11 | 0 | 90 | 7 | 0 | 1006 | 10 |
| 07-12-2021 | 12 | 0 | 75 | 6 | 0 | 1007 | 4 |
| 08-12-2021 | 12 | 0 | 85 | 6 | 2 | 1008 | 10 |
| 09-12-2021 | 12 | 4 | 60 | 8 | 2 | 1009 | 2 |
| 10-12-2021 | 12 | 0 | 85 | 9 | 1 | 1010 | 6 |

In [226...]

```
df.pivot_table(index='Percanage') #Index is percentage and other columns comes in aggregate
```

Out[226...]

| | Class | Mobile Games hrs | Sleeping hrs | Social Media usage hrs | Student ID | Study hrs |
|-----------|-------|------------------|--------------|------------------------|------------|-----------|
| Percanage | | | | | | |
| 50 | 10.0 | 5 | 9.0 | 3.0 | 1001.0 | 2.0 |
| 60 | 11.0 | 4 | 8.0 | 2.0 | 1006.0 | 2.5 |
| 70 | 10.0 | 5 | 8.0 | 1.0 | 1004.0 | 0.0 |
| 75 | 12.0 | 0 | 6.0 | 0.0 | 1007.0 | 4.0 |
| 80 | 10.5 | 0 | 7.5 | 2.0 | 1003.5 | 5.0 |
| 85 | 12.0 | 0 | 7.5 | 1.5 | 1009.0 | 8.0 |
| 90 | 11.0 | 0 | 7.0 | 0.0 | 1006.0 | 10.0 |

In [231...]

```
df.pivot_table(index='Social Media usage hrs',columns='Date',aggfunc='count',
               fill_value='none',margins=True) #aggfunc is default mean
#fill_value to fill NaN with other values
```

Out[231...]

| | Class ... | | | | | | | | | | | | | | |
|--------------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|-----|------------|------------|------------|------------|
| Date | 01-12-2021 | 02-12-2021 | 03-12-2021 | 04-12-2021 | 05-12-2021 | 06-12-2021 | 07-12-2021 | 08-12-2021 | 09-12-2021 | 10-12-2021 | ... | 02-12-2021 | 03-12-2021 | 04-12-2021 | 05-12-2021 |
| Social Media usage | | | | | | | | | | | | | | | |

| | | | | | | | | | | | | | Class | ... | | | | |
|-------------------------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|-----|------------|------------|------------|------------|------|------|--|
| Date | 01-12-2021 | 02-12-2021 | 03-12-2021 | 04-12-2021 | 05-12-2021 | 06-12-2021 | 07-12-2021 | 08-12-2021 | 09-12-2021 | 10-12-2021 | ... | 02-12-2021 | 03-12-2021 | 04-12-2021 | 05-12-2021 | | | |
| Social Media usage hrs | | | | | | | | | | | | | | | | | | |
| 0 | none | none | none | none | none | 1.0 | 1.0 | none | none | none | ... | none | none | none | none | non | none | |
| 1 | none | none | none | 1.0 | none | none | none | none | none | 1.0 | ... | none | none | 1.0 | non | none | | |
| 2 | none | 1.0 | 1.0 | none | 1.0 | none | none | 1.0 | 1.0 | none | ... | 1.0 | 1.0 | none | 1. | none | | |
| 3 | 1.0 | none | ... | none | none | none | none | non | | |
| All | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | ... | 1 | 1 | 1 | 1 | 1 | | |

5 rows × 66 columns

In [232]...

```
#PANDAS MELT function is used to tranform or reshape data
```

In [233]...

```
df=pd.read_csv("E:\\New folder\\student_results.csv")
df
```

Out[233]...

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentage |
|----------|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [234]...

```
pd.melt(df)#it calls according to columns
```

Out[234]...

| | variable | value |
|----------|----------|------------|
| 0 | Date | 01-12-2021 |
| 1 | Date | 02-12-2021 |

| | variable | value |
|-----|------------|------------|
| 2 | Date | 03-12-2021 |
| 3 | Date | 04-12-2021 |
| 4 | Date | 05-12-2021 |
| ... | ... | ... |
| 75 | Percantege | 90 |
| 76 | Percantege | 75 |
| 77 | Percantege | 85 |
| 78 | Percantege | 60 |
| 79 | Percantege | 85 |

80 rows × 2 columns

In [235...]

```
pd.melt(df,id_vars=['Date'])#Column(s) to use as identifier variables.
```

Out[235...]

| | Date | variable | value |
|-----|------------|------------|-------|
| 0 | 01-12-2021 | Student ID | 1001 |
| 1 | 02-12-2021 | Student ID | 1002 |
| 2 | 03-12-2021 | Student ID | 1003 |
| 3 | 04-12-2021 | Student ID | 1004 |
| 4 | 05-12-2021 | Student ID | 1005 |
| ... | ... | ... | ... |
| 65 | 06-12-2021 | Percantege | 90 |
| 66 | 07-12-2021 | Percantege | 75 |
| 67 | 08-12-2021 | Percantege | 85 |
| 68 | 09-12-2021 | Percantege | 60 |
| 69 | 10-12-2021 | Percantege | 85 |

70 rows × 3 columns

In [236...]

```
#UNPIVOT turn a wide format(many columns) to a long format (few columns but many rows)
```

In [237...]

```
pd.melt(df,id_vars=['Date'],value_vars=['Percantege'])#percent ke hissab se values aayi h
```

Out[237...]

| | Date | variable | value |
|---|------------|------------|-------|
| 0 | 01-12-2021 | Percantege | 50 |
| 1 | 02-12-2021 | Percantege | 80 |
| 2 | 03-12-2021 | Percantege | 60 |
| 3 | 04-12-2021 | Percantege | 70 |
| 4 | 05-12-2021 | Percantege | 80 |
| 5 | 06-12-2021 | Percantege | 90 |

Date variable value

| | | | |
|---|------------|------------|----|
| 7 | 08-12-2021 | Percantege | 85 |
| 8 | 09-12-2021 | Percantege | 60 |
| 9 | 10-12-2021 | Percantege | 85 |

In [239...]

```
pd.melt(df,id_vars=['Date'],value_vars=['Percantege'],var_name='Percent Column',value_nam
```

Out[239...]

Date Percent Column Data

| | | | |
|---|------------|------------|----|
| 0 | 01-12-2021 | Percantege | 50 |
| 1 | 02-12-2021 | Percantege | 80 |
| 2 | 03-12-2021 | Percantege | 60 |
| 3 | 04-12-2021 | Percantege | 70 |
| 4 | 05-12-2021 | Percantege | 80 |
| 5 | 06-12-2021 | Percantege | 90 |
| 6 | 07-12-2021 | Percantege | 75 |
| 7 | 08-12-2021 | Percantege | 85 |
| 8 | 09-12-2021 | Percantege | 60 |
| 9 | 10-12-2021 | Percantege | 85 |

In [240...]

```
df=pd.read_csv("E:\\New folder\\student_results.csv")
df
```

Out[240...]

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percentege |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 01-12-2021 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 02-12-2021 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 03-12-2021 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 04-12-2021 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 05-12-2021 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 06-12-2021 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 07-12-2021 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 08-12-2021 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 09-12-2021 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 10-12-2021 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

In [242...]

```
type(df.Date[0])
```

Loading [MathJax]/extensions/Safe.js

```
Out[242... str
```

```
In [243... df=pd.read_csv("E:\\New folder\\student_results.csv",parse_dates=['Date'])  
df
```

```
Out[243... 
```

| | Date | Student ID | Class | Study hrs | Sleeping hrs | Social Media usage hrs | Mobile Games hrs | Percantere |
|---|------------|------------|-------|-----------|--------------|------------------------|------------------|------------|
| 0 | 2021-01-12 | 1001 | 10 | 2 | 9 | 3 | 5 | 50 |
| 1 | 2021-02-12 | 1002 | 10 | 6 | 8 | 2 | 0 | 80 |
| 2 | 2021-03-12 | 1003 | 10 | 3 | 8 | 2 | 4 | 60 |
| 3 | 2021-04-12 | 1004 | 10 | 0 | 8 | 1 | 5 | 70 |
| 4 | 2021-05-12 | 1005 | 11 | 4 | 7 | 2 | 0 | 80 |
| 5 | 2021-06-12 | 1006 | 11 | 10 | 7 | 0 | 0 | 90 |
| 6 | 2021-07-12 | 1007 | 12 | 4 | 6 | 0 | 0 | 75 |
| 7 | 2021-08-12 | 1008 | 12 | 10 | 6 | 2 | 0 | 85 |
| 8 | 2021-09-12 | 1009 | 12 | 2 | 8 | 2 | 4 | 60 |
| 9 | 2021-10-12 | 1010 | 12 | 6 | 9 | 1 | 0 | 85 |

```
In [244... 
```

```
type(df.Date[0])
```

```
Out[244... 
```

```
pandas._libs.tslibs.timestamps.Timestamp
```

```
In [ ]:
```