

Assignment-2

Anjali vishwakarma-MT2310

```
In [60]: pip install pandas
```

```
Requirement already satisfied: pandas in c:\users\nano\appdata\local\programs\python\python312\lib\site-packages (2.2.3)
Note: you may need to restart the kernel to use updated packages.
```

```
[notice] A new release of pip is available: 24.1.1 -> 24.2
```

```
[notice] To update, run: python.exe -m pip install --upgrade pip
```

```
Requirement already satisfied: numpy>=1.26.0 in c:\users\nano\appdata\local\programs\python\python312\lib\site-packages (from pandas) (2.1.2)
```

```
Requirement already satisfied: python-dateutil>=2.8.2 in c:\users\nano\appdata\local\programs\python\python312\lib\site-packages (from pandas) (2.9.0.post0)
```

```
Requirement already satisfied: pytz>=2020.1 in c:\users\nano\appdata\local\programs\python\python312\lib\site-packages (from pandas) (2024.2)
```

```
Requirement already satisfied: tzdata>=2022.7 in c:\users\nano\appdata\local\programs\python\python312\lib\site-packages (from pandas) (2024.2)
```

```
Requirement already satisfied: six>=1.5 in c:\users\nano\appdata\local\programs\python\python312\lib\site-packages (from python-dateutil>=2.8.2->pandas) (1.16.0)
```

```
In [61]: pip install seaborn
```

```
Requirement already satisfied: seaborn in c:\users\nano\appdata\local\programs\python\python312\lib\site-packages (0.13.2)
Requirement already satisfied: numpy!=1.24.0,>=1.20 in c:\users\nano\appdata\local\programs\python\python312\lib\site-packages (from seaborn) (2.1.2)
Requirement already satisfied: pandas>=1.2 in c:\users\nano\appdata\local\programs\python\python312\lib\site-packages (from seaborn) (2.2.3)
Requirement already satisfied: matplotlib!=3.6.1,>=3.4 in c:\users\nano\appdata\local\programs\python\python312\lib\site-packages (from seaborn) (3.9.2)
Requirement already satisfied: contourpy>=1.0.1 in c:\users\nano\appdata\local\programs\python\python312\lib\site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (1.3.0)
Requirement already satisfied: cycler>=0.10 in c:\users\nano\appdata\local\programs\python\python312\lib\site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in c:\users\nano\appdata\local\programs\python\python312\lib\site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (4.54.1)
Requirement already satisfied: kiwisolver>=1.3.1 in c:\users\nano\appdata\local\programs\python\python312\lib\site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (1.4.7)
Requirement already satisfied: packaging>=20.0 in c:\users\nano\appdata\local\programs\python\python312\lib\site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (24.1)
Requirement already satisfied: pillow>=8 in c:\users\nano\appdata\local\programs\python\python312\lib\site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (1.1.0)
Requirement already satisfied: pyparsing>=2.3.1 in c:\users\nano\appdata\local\programs\python\python312\lib\site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (3.2.0)
Requirement already satisfied: python-dateutil>=2.7 in c:\users\nano\appdata\local\programs\python\python312\lib\site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (2.9.0.post0)
Requirement already satisfied: pytz>=2020.1 in c:\users\nano\appdata\local\programs\python\python312\lib\site-packages (from pandas>=1.2->seaborn) (2024.2)
Requirement already satisfied: tzdata>=2022.7 in c:\users\nano\appdata\local\programs\python\python312\lib\site-packages (from pandas>=1.2->seaborn) (2024.2)
Requirement already satisfied: six>=1.5 in c:\users\nano\appdata\local\programs\python\python312\lib\site-packages (from python-dateutil>=2.7->matplotlib!=3.6.1,>=3.4->seaborn) (1.16.0)
Note: you may need to restart the kernel to use updated packages.
```

```
[notice] A new release of pip is available: 24.1.1 -> 24.2
[notice] To update, run: python.exe -m pip install --upgrade pip
```

```
In [62]: pip install scipy
```

```
Requirement already satisfied: scipy in c:\users\nano\appdata\local\programs\python\python312\lib\site-packages (1.14.1)
Requirement already satisfied: numpy<2.3,>=1.23.5 in c:\users\nano\appdata\local\programs\python\python312\lib\site-packages (from scipy) (2.1.2)
Note: you may need to restart the kernel to use updated packages.
```

```
[notice] A new release of pip is available: 24.1.1 -> 24.2
[notice] To update, run: python.exe -m pip install --upgrade pip
```

```
In [63]: #import python Libaries
import numpy as np
import pandas as pd
```

```
In [64]: df = pd.read_csv(r"C:\Users\nano\Downloads\StudentPerformanceFactors.csv", encoding='ISO-8859-1')
```

```
In [14]: # (r"C:\Users\nano\Downloads\StudentPerformanceFactors.csv", encoding='ISO-8859-1')
```

```
In [65]: df.columns.values
```

```
Out[65]: array(['Hours_Studied', 'Attendance', 'Parental_Involvement',
       'Access_to_Resources', 'Extracurricular_Activities', 'Sleep_Hours',
       'Previous_Scores', 'Motivation_Level', 'Internet_Access',
       'Tutoring_Sessions', 'Family_Income', 'Teacher_Quality',
       'School_Type', 'Peer_Influence', 'Physical_Activity',
       'Learning_Disabilities', 'Parental_Education_Level',
       'Distance_from_Home', 'Gender', 'Exam_Score'], dtype=object)
```

```
In [66]: df.dtypes
```

```
Out[66]: Hours_Studied          int64
Attendance            int64
Parental_Involvement    object
Access_to_Resources      object
Extracurricular_Activities  object
Sleep_Hours            int64
Previous_Scores         int64
Motivation_Level        object
Internet_Access         object
Tutoring_Sessions        int64
Family_Income           object
Teacher_Quality          object
School_Type             object
Peer_Influence           object
Physical_Activity        int64
Learning_Disabilities    object
Parental_Education_Level  object
Distance_from_Home        object
Gender                  object
Exam_Score              int64
dtype: object
```

```
In [67]: df.shape
```

```
Out[67]: (6607, 20)
```

```
In [68]: df.shape[0]
```

```
Out[68]: 6607
```

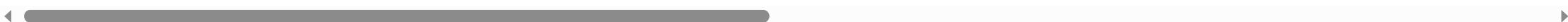
```
In [69]: df.shape[1]
```

```
Out[69]: 20
```

```
In [70]: df.head(5)
```

```
Out[70]:
```

	Hours_Studied	Attendance	Parental_Involvement	Access_to_Resources	Extracurricular_Activities	Sleep_Hours	Previous_Scores	Motivation_Level	Internet_Access	Tutorin
0	23	84	Low	High	No	7	73	Low	Yes	
1	19	64	Low	Medium	No	8	59	Low	Yes	
2	24	98	Medium	Medium	Yes	7	91	Medium	Yes	
3	29	89	Low	Medium	Yes	8	98	Medium	Yes	
4	19	92	Medium	Medium	Yes	6	65	Medium	Yes	



```
In [71]: df.tail(5)
```

```
Out[71]:
```

	Hours_Studied	Attendance	Parental_Involvement	Access_to_Resources	Extracurricular_Activities	Sleep_Hours	Previous_Scores	Motivation_Level	Internet_Access	Tutorin
6602	25	69	High	Medium	No	7	76	Medium	Yes	
6603	23	76	High	Medium	No	8	81	Medium	Yes	
6604	20	90	Medium	Low	Yes	6	65	Low	Yes	
6605	10	86	High	High	Yes	6	91	High	Yes	
6606	15	67	Medium	Low	Yes	9	94	Medium	Yes	



```
In [72]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 6607 entries, 0 to 6606
Data columns (total 20 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   Hours_Studied    6607 non-null   int64  
 1   Attendance        6607 non-null   int64  
 2   Parental_Involvement 6607 non-null  object  
 3   Access_to_Resources 6607 non-null  object  
 4   Extracurricular_Activities 6607 non-null  object  
 5   Sleep_Hours       6607 non-null   int64  
 6   Previous_Scores   6607 non-null   int64  
 7   Motivation_Level  6607 non-null   object  
 8   Internet_Access   6607 non-null   object  
 9   Tutoring_Sessions 6607 non-null   int64  
 10  Family_Income     6607 non-null   object  
 11  Teacher_Quality   6529 non-null   object  
 12  School_Type       6607 non-null   object  
 13  Peer_Influence    6607 non-null   object  
 14  Physical_Activity 6607 non-null   int64  
 15  Learning_Disabilities 6607 non-null  object  
 16  Parental_Education_Level 6517 non-null  object  
 17  Distance_from_Home 6540 non-null   object  
 18  Gender            6607 non-null   object  
 19  Exam_Score        6607 non-null   int64  
dtypes: int64(7), object(13)
memory usage: 1.0+ MB
```

```
In [73]: df.Attendance.count()
```

```
Out[73]: np.int64(6607)
```

```
In [74]: df.Attendance.value_counts()
```

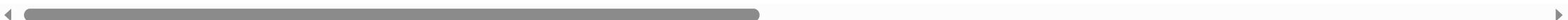
```
Out[74]: Attendance  
67    190  
98    187  
76    185  
77    184  
64    182  
94    180  
91    175  
84    175  
79    175  
82    173  
69    170  
68    170  
80    169  
96    168  
73    168  
81    168  
93    167  
72    167  
74    165  
78    165  
61    164  
95    163  
71    162  
89    162  
97    161  
70    161  
65    158  
83    157  
90    156  
63    155  
88    155  
92    154  
99    154  
62    152  
87    151  
86    151  
75    149  
85    146  
66    145  
60     87
```

```
100      81  
Name: count, dtype: int64
```

```
In [75]: df.isnull()
```

	Hours_Studied	Attendance	Parental_Involvement	Access_to_Resources	Extracurricular_Activities	Sleep_Hours	Previous_Scores	Motivation_Level	Internet_Access	Tut
0	False	False		False	False	False	False	False	False	False
1	False	False		False	False	False	False	False	False	False
2	False	False		False	False	False	False	False	False	False
3	False	False		False	False	False	False	False	False	False
4	False	False		False	False	False	False	False	False	False
...
6602	False	False		False	False	False	False	False	False	False
6603	False	False		False	False	False	False	False	False	False
6604	False	False		False	False	False	False	False	False	False
6605	False	False		False	False	False	False	False	False	False
6606	False	False		False	False	False	False	False	False	False

6607 rows × 20 columns



```
In [76]: df.notnull()
```

Out[76]:

	Hours_Studied	Attendance	Parental_Involvement	Access_to_Resources	Extracurricular_Activities	Sleep_Hours	Previous_Scores	Motivation_Level	Internet_Access	Tutoring
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
...
6602	True	True		True		True	True	True	True	True
6603	True	True		True		True	True	True	True	True
6604	True	True		True		True	True	True	True	True
6605	True	True		True		True	True	True	True	True
6606	True	True		True		True	True	True	True	True

6607 rows × 20 columns

In [77]: `df[df.notnull()]`

Out[77]:

	Hours_Studied	Attendance	Parental_Involvement	Access_to_Resources	Extracurricular_Activities	Sleep_Hours	Previous_Scores	Motivation_Level	Internet_Access	Tutoring
0	23	84	Low	High	No	7	73	Low	Yes	
1	19	64	Low	Medium	No	8	59	Low	Yes	
2	24	98	Medium	Medium	Yes	7	91	Medium	Yes	
3	29	89	Low	Medium	Yes	8	98	Medium	Yes	
4	19	92	Medium	Medium	Yes	6	65	Medium	Yes	
...
6602	25	69	High	Medium	No	7	76	Medium	Yes	
6603	23	76	High	Medium	No	8	81	Medium	Yes	
6604	20	90	Medium	Low	Yes	6	65	Low	Yes	
6605	10	86	High	High	Yes	6	91	High	Yes	
6606	15	67	Medium	Low	Yes	9	94	Medium	Yes	

6607 rows × 20 columns

In [78]: `df[df.isnull()]`

Out[78]:

	Hours_Studied	Attendance	Parental_Involvement	Access_to_Resources	Extracurricular_Activities	Sleep_Hours	Previous_Scores	Motivation_Level	Internet_Access	Tutoring
0	NaN	NaN		NaN	NaN	NaN	NaN	NaN	NaN	NaN
1	NaN	NaN		NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	NaN	NaN		NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	NaN	NaN		NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	NaN	NaN		NaN	NaN	NaN	NaN	NaN	NaN	NaN
...
6602	NaN	NaN		NaN	NaN	NaN	NaN	NaN	NaN	NaN
6603	NaN	NaN		NaN	NaN	NaN	NaN	NaN	NaN	NaN
6604	NaN	NaN		NaN	NaN	NaN	NaN	NaN	NaN	NaN
6605	NaN	NaN		NaN	NaN	NaN	NaN	NaN	NaN	NaN
6606	NaN	NaN		NaN	NaN	NaN	NaN	NaN	NaN	NaN

6607 rows × 20 columns



In [79]:

df.describe()

Out[79]:

	Hours_Studied	Attendance	Sleep_Hours	Previous_Scores	Tutoring_Sessions	Physical_Activity	Exam_Score
count	6607.000000	6607.000000	6607.000000	6607.000000	6607.000000	6607.000000	6607.000000
mean	19.975329	79.977448	7.02906	75.070531	1.493719	2.967610	67.235659
std	5.990594	11.547475	1.46812	14.399784	1.230570	1.031231	3.890456
min	1.000000	60.000000	4.00000	50.000000	0.000000	0.000000	55.000000
25%	16.000000	70.000000	6.00000	63.000000	1.000000	2.000000	65.000000
50%	20.000000	80.000000	7.00000	75.000000	1.000000	3.000000	67.000000
75%	24.000000	90.000000	8.00000	88.000000	2.000000	4.000000	69.000000
max	44.000000	100.000000	10.00000	100.000000	8.000000	6.000000	101.000000

In [81]: `df.sum(numeric_only=True)`

Out[81]:

Hours_Studied	131977
Attendance	528411
Sleep_Hours	46441
Previous_Scores	495991
Tutoring_Sessions	9869
Physical_Activity	19607
Exam_Score	444226
dtype:	int64

In [82]: `df.mean(numeric_only=True)`

Out[82]:

Hours_Studied	19.975329
Attendance	79.977448
Sleep_Hours	7.02906
Previous_Scores	75.070531
Tutoring_Sessions	1.493719
Physical_Activity	2.967610
Exam_Score	67.235659
dtype:	float64

In [83]: `df.corr(numeric_only=True)`

```
Out[83]:
```

	Hours_Studied	Attendance	Sleep_Hours	Previous_Scores	Tutoring_Sessions	Physical_Activity	Exam_Score
Hours_Studied	1.000000	-0.009908	0.010977	0.024846	-0.014282	0.004624	0.445455
Attendance	-0.009908	1.000000	-0.015918	-0.020186	0.014324	-0.022435	0.581072
Sleep_Hours	0.010977	-0.015918	1.000000	-0.021750	-0.012216	-0.000378	-0.017022
Previous_Scores	0.024846	-0.020186	-0.021750	1.000000	-0.013122	-0.011274	0.175079
Tutoring_Sessions	-0.014282	0.014324	-0.012216	-0.013122	1.000000	0.017733	0.156525
Physical_Activity	0.004624	-0.022435	-0.000378	-0.011274	0.017733	1.000000	0.027824
Exam_Score	0.445455	0.581072	-0.017022	0.175079	0.156525	0.027824	1.000000

```
In [84]:
```

```
df.cov(numeric_only=True)
```

```
Out[84]:
```

	Hours_Studied	Attendance	Sleep_Hours	Previous_Scores	Tutoring_Sessions	Physical_Activity	Exam_Score
Hours_Studied	35.887221	-0.685388	0.096539	2.143278	-0.105287	0.028568	10.381837
Attendance	-0.685388	133.344178	-0.269856	-3.356568	0.203537	-0.267155	26.104619
Sleep_Hours	0.096539	-0.269856	2.155377	-0.459816	-0.022070	-0.000572	-0.097222
Previous_Scores	2.143278	-3.356568	-0.459816	207.353789	-0.232527	-0.167409	9.808232
Tutoring_Sessions	-0.105287	0.203537	-0.022070	-0.232527	1.514304	0.022503	0.749361
Physical_Activity	0.028568	-0.267155	-0.000572	-0.167409	0.022503	1.063438	0.111630
Exam_Score	10.381837	26.104619	-0.097222	9.808232	0.749361	0.111630	15.135646

```
In [86]:
```

```
df.Sleep_Hours.mean()
```

```
Out[86]:
```

```
np.float64(7.029060087785682)
```

```
In [87]:
```

```
df.Exam_Score.sum()
```

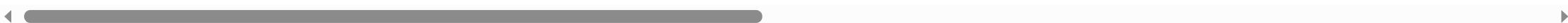
```
Out[87]:
```

```
np.int64(444226)
```

```
In [88]: df.fillna(0)
```

	Hours_Studied	Attendance	Parental_Involvement	Access_to_Resources	Extracurricular_Activities	Sleep_Hours	Previous_Scores	Motivation_Level	Internet_Access	Tutoring
0	23	84	Low	High	No	7	73	Low	Yes	
1	19	64	Low	Medium	No	8	59	Low	Yes	
2	24	98	Medium	Medium	Yes	7	91	Medium	Yes	
3	29	89	Low	Medium	Yes	8	98	Medium	Yes	
4	19	92	Medium	Medium	Yes	6	65	Medium	Yes	
...
6602	25	69	High	Medium	No	7	76	Medium	Yes	
6603	23	76	High	Medium	No	8	81	Medium	Yes	
6604	20	90	Medium	Low	Yes	6	65	Low	Yes	
6605	10	86	High	High	Yes	6	91	High	Yes	
6606	15	67	Medium	Low	Yes	9	94	Medium	Yes	

6607 rows × 20 columns

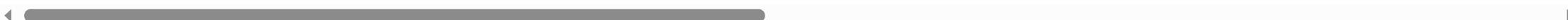


```
In [89]: df.isnull()
```

Out[89]:

	Hours_Studied	Attendance	Parental_Involvement	Access_to_Resources	Extracurricular_Activities	Sleep_Hours	Previous_Scores	Motivation_Level	Internet_Access	Tutoring
0	False	False		False	False	False	False	False	False	False
1	False	False		False	False	False	False	False	False	False
2	False	False		False	False	False	False	False	False	False
3	False	False		False	False	False	False	False	False	False
4	False	False		False	False	False	False	False	False	False
...
6602	False	False		False	False	False	False	False	False	False
6603	False	False		False	False	False	False	False	False	False
6604	False	False		False	False	False	False	False	False	False
6605	False	False		False	False	False	False	False	False	False
6606	False	False		False	False	False	False	False	False	False

6607 rows × 20 columns



In [90]:

df.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 6607 entries, 0 to 6606
Data columns (total 20 columns):
 #   Column           Non-Null Count  Dtype  
 ---  -- 
 0   Hours_Studied    6607 non-null   int64  
 1   Attendance        6607 non-null   int64  
 2   Parental_Involvement 6607 non-null   object  
 3   Access_to_Resources 6607 non-null   object  
 4   Extracurricular_Activities 6607 non-null   object  
 5   Sleep_Hours       6607 non-null   int64  
 6   Previous_Scores   6607 non-null   int64  
 7   Motivation_Level  6607 non-null   object  
 8   Internet_Access   6607 non-null   object  
 9   Tutoring_Sessions 6607 non-null   int64  
 10  Family_Income     6607 non-null   object  
 11  Teacher_Quality   6529 non-null   object  
 12  School_Type       6607 non-null   object  
 13  Peer_Influence    6607 non-null   object  
 14  Physical_Activity 6607 non-null   int64  
 15  Learning_Disabilities 6607 non-null   object  
 16  Parental_Education_Level 6517 non-null   object  
 17  Distance_from_Home 6540 non-null   object  
 18  Gender            6607 non-null   object  
 19  Exam_Score        6607 non-null   int64  
dtypes: int64(7), object(13)
memory usage: 1.0+ MB
```

In [91]: df[df.isnull()]

Out[91]:

	Hours_Studied	Attendance	Parental_Involvement	Access_to_Resources	Extracurricular_Activities	Sleep_Hours	Previous_Scores	Motivation_Level	Internet_Access	Tutoring_Sessions
0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
1	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
...
6602	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
6603	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
6604	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
6605	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
6606	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

6607 rows × 20 columns



In [94]:

```
df1=df.drop(['Tutoring_Sessions'],axis=1)
df1
```

Out[94]:

	Hours_Studied	Attendance	Parental_Involvement	Access_to_Resources	Extracurricular_Activities	Sleep_Hours	Previous_Scores	Motivation_Level	Internet_Access	Fan
0	23	84	Low	High	No	7	73	Low	Yes	
1	19	64	Low	Medium	No	8	59	Low	Yes	
2	24	98	Medium	Medium	Yes	7	91	Medium	Yes	
3	29	89	Low	Medium	Yes	8	98	Medium	Yes	
4	19	92	Medium	Medium	Yes	6	65	Medium	Yes	
...	
6602	25	69	High	Medium	No	7	76	Medium	Yes	
6603	23	76	High	Medium	No	8	81	Medium	Yes	
6604	20	90	Medium	Low	Yes	6	65	Low	Yes	
6605	10	86	High	High	Yes	6	91	High	Yes	
6606	15	67	Medium	Low	Yes	9	94	Medium	Yes	

6607 rows × 19 columns

In [97]: `df1.drop(['Parental_Involvement', 'Motivation_Level'], axis=1)`

Out[97]:

	Hours_Studied	Attendance	Access_to_Resources	Extracurricular_Activities	Sleep_Hours	Previous_Scores	Internet_Access	Family_Income	Teacher_Quality	School_Type
0	23	84	High	No	7	73	Yes	Low	Medium	Public
1	19	64	Medium	No	8	59	Yes	Medium	Medium	Public
2	24	98	Medium	Yes	7	91	Yes	Medium	Medium	Public
3	29	89	Medium	Yes	8	98	Yes	Medium	Medium	Public
4	19	92	Medium	Yes	6	65	Yes	Medium	High	Public
...
6602	25	69	Medium	No	7	76	Yes	High	Medium	Public
6603	23	76	Medium	No	8	81	Yes	Low	High	Public
6604	20	90	Low	Yes	6	65	Yes	Low	Medium	Public
6605	10	86	High	Yes	6	91	Yes	Low	Medium	Private
6606	15	67	Low	Yes	9	94	Yes	Medium	Medium	Public

6607 rows × 17 columns

In [99]: `del df['Previous_Scores']`In [100...]: `df`

Out[100...]

	Hours_Studied	Attendance	Parental_Involvement	Access_to_Resources	Extracurricular_Activities	Sleep_Hours	Motivation_Level	Internet_Access	Tutoring_Sessions	Fatigue_Level
0	23	84	Low	High	No	7	Low	Yes	0	Medium
1	19	64	Low	Medium	No	8	Low	Yes	2	Medium
2	24	98	Medium	Medium	Yes	7	Medium	Yes	2	Low
3	29	89	Low	Medium	Yes	8	Medium	Yes	1	Low
4	19	92	Medium	Medium	Yes	6	Medium	Yes	3	Medium
...
6602	25	69	High	Medium	No	7	Medium	Yes	1	Low
6603	23	76	High	Medium	No	8	Medium	Yes	3	Low
6604	20	90	Medium	Low	Yes	6	Low	Yes	3	Medium
6605	10	86	High	High	Yes	6	High	Yes	2	Medium
6606	15	67	Medium	Low	Yes	9	Medium	Yes	0	Medium

6607 rows × 19 columns



In [101...]

df.T

Out[101...]	0	1	2	3	4	5	6	7	8	9	...	6597	6598	6599
Hours_Studied	23	19	24	29	19	19	29	25	17	23	...	16	9	30
Attendance	84	64	98	89	92	88	84	78	94	98	...	91	87	94
Parental_Involvement	Low	Low	Medium	Low	Medium	Medium	Medium	Low	Medium	Medium	...	High	Low	Medium
Access_to_Resources	High	Medium	Medium	Medium	Medium	Medium	Low	High	High	Medium	...	Medium	Medium	Low
Extracurricular_Activities	No	No	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	...	Yes	Yes	No
Sleep_Hours	7	8	7	8	6	8	7	6	6	8	...	6	6	5
Motivation_Level	Low	Low	Medium	Medium	Medium	Medium	Low	Medium	High	Medium	...	High	Medium	Low
Internet_Access	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	...	Yes	Yes	No
Tutoring_Sessions	0	2	2	1	3	3	1	1	0	0	...	0	1	3
Family_Income	Low	Medium	Medium	Medium	Medium	Medium	Low	High	Medium	High	...	High	High	High
Teacher_Quality	Medium	Medium	Medium	Medium	High	Medium	Medium	High	Low	High	...	High	Medium	Medium
School_Type	Public	Public	Public	Public	Public	Public	Private	Public	Private	Public	...	Public	Public	Private
Peer_Influence	Positive	Negative	Neutral	Negative	Neutral	Positive	Neutral	Negative	Neutral	Positive	...	Negative	Neutral	Neutral
Physical_Activity	3	4	4	4	4	3	2	2	1	5	...	2	2	2
Learning_Disabilities	No	No	No	No	No	No	No	No	No	No	...	No	No	No
Parental_Education_Level	High School	College	Postgraduate	High School	College	Postgraduate	High School	High School	College	High School	...	Postgraduate	High School	Postgraduate
Distance_from_Home	Near	Moderate	Near	Moderate	Near	Near	Moderate	Far	Near	Moderate	...	Near	Near	Moderate
Gender	Male	Female	Male	Male	Female	Male	Male	Male	Male	Male	...	Female	Female	Female
Exam_Score	67	61	74	71	70	71	67	66	69	72	...	70	64	70

19 rows × 6607 columns

In [102... df.transpose()

Out[102...]

	0	1	2	3	4	5	6	7	8	9	...	6597	6598	6599
Hours_Studied	23	19	24	29	19	19	29	25	17	23	...	16	9	30
Attendance	84	64	98	89	92	88	84	78	94	98	...	91	87	94
Parental_Involvement	Low	Low	Medium	Low	Medium	Medium	Medium	Low	Medium	Medium	...	High	Low	Medium
Access_to_Resources	High	Medium	Medium	Medium	Medium	Medium	Low	High	High	Medium	...	Medium	Medium	Low
Extracurricular_Activities	No	No	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	...	Yes	Yes	No
Sleep_Hours	7	8	7	8	6	8	7	6	6	8	...	6	6	5
Motivation_Level	Low	Low	Medium	Medium	Medium	Medium	Low	Medium	High	Medium	...	High	Medium	Low
Internet_Access	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	...	Yes	Yes	No
Tutoring_Sessions	0	2	2	1	3	3	1	1	0	0	...	0	1	3
Family_Income	Low	Medium	Medium	Medium	Medium	Medium	Low	High	Medium	High	...	High	High	High
Teacher_Quality	Medium	Medium	Medium	Medium	High	Medium	Medium	High	Low	High	...	High	Medium	Medium
School_Type	Public	Public	Public	Public	Public	Public	Private	Public	Private	Public	...	Public	Public	Private
Peer_Influence	Positive	Negative	Neutral	Negative	Neutral	Positive	Neutral	Negative	Neutral	Positive	...	Negative	Neutral	Neutral
Physical_Activity	3	4	4	4	4	3	2	2	1	5	...	2	2	2
Learning_Disabilities	No	No	No	No	No	No	No	No	No	No	...	No	No	No
Parental_Education_Level	High School	College	Postgraduate	High School	College	Postgraduate	High School	High School	College	High School	...	Postgraduate	High School	Postgraduate
Distance_from_Home	Near	Moderate	Near	Moderate	Near	Near	Moderate	Far	Near	Moderate	...	Near	Near	Moderate
Gender	Male	Female	Male	Male	Female	Male	Male	Male	Male	Male	...	Female	Female	Female
Exam_Score	67	61	74	71	70	71	67	66	69	72	...	70	64	70

19 rows × 6607 columns

```
In [103...]
```

```
df['new']=12  
df
```

```
Out[103...]
```

	Hours_Studied	Attendance	Parental_Involvement	Access_to_Resources	Extracurricular_Activities	Sleep_Hours	Motivation_Level	Internet_Access	Tutoring_Sessions	Fa
0	23	84	Low	High	No	7	Low	Yes	0	
1	19	64	Low	Medium	No	8	Low	Yes	2	
2	24	98	Medium	Medium	Yes	7	Medium	Yes	2	
3	29	89	Low	Medium	Yes	8	Medium	Yes	1	
4	19	92	Medium	Medium	Yes	6	Medium	Yes	3	
...	
6602	25	69	High	Medium	No	7	Medium	Yes	1	
6603	23	76	High	Medium	No	8	Medium	Yes	3	
6604	20	90	Medium	Low	Yes	6	Low	Yes	3	
6605	10	86	High	High	Yes	6	High	Yes	2	
6606	15	67	Medium	Low	Yes	9	Medium	Yes	0	

6607 rows × 20 columns



```
In [108...]
```

```
df['new_column_added_with85'] = 85  
df
```

Out[108...]

	Hours_Studied	Attendance	Parental_Involvement	Access_to_Resources	Extracurricular_Activities	Sleep_Hours	Motivation_Level	Internet_Access	Tutoring_Sessions	Fatigue_Level
0	23	84	Low	High	No	7	Low	Yes	0	Medium
1	19	64	Low	Medium	No	8	Low	Yes	2	High
2	24	98	Medium	Medium	Yes	7	Medium	Yes	2	Medium
3	29	89	Low	Medium	Yes	8	Medium	Yes	1	Medium
4	19	92	Medium	Medium	Yes	6	Medium	Yes	3	Medium
...
6602	25	69	High	Medium	No	7	Medium	Yes	1	Medium
6603	23	76	High	Medium	No	8	Medium	Yes	3	Medium
6604	20	90	Medium	Low	Yes	6	Low	Yes	3	Medium
6605	10	86	High	High	Yes	6	High	Yes	2	Medium
6606	15	67	Medium	Low	Yes	9	Medium	Yes	0	Medium

6607 rows × 23 columns



In []: