

StudentiAnaliza

August 1, 2025

```
[1]: import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

```
[2]: studenti = pd.read_csv("../..../datasets/Students.csv")
```

```
[3]: studenti.head()
```

```
[3]:   gender race/ethnicity parental level of education      lunch \
0  female      group B      bachelor's degree      standard
1  female      group C          some college      standard
2  female      group B      master's degree      standard
3   male      group A      associate's degree  free/reduced
4   male      group C          some college      standard

   test preparation course  math score  reading score  writing score
0                none         72         72         74
1          completed         69         90         88
2                none         90         95         93
3                none         47         57         44
4                none         76         78         75
```

```
[4]: studenti.describe()
```

```
[4]:   math score  reading score  writing score
count  1000.00000    1000.00000    1000.00000
mean    66.08900     69.16900     68.05400
std     15.16308     14.600192    15.195657
min      0.00000     17.00000     10.00000
25%     57.00000     59.00000     57.75000
50%     66.00000     70.00000     69.00000
75%     77.00000     79.00000     79.00000
max    100.00000    100.00000    100.00000
```

```
[5]: studenti.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1000 entries, 0 to 999
Data columns (total 8 columns):
```

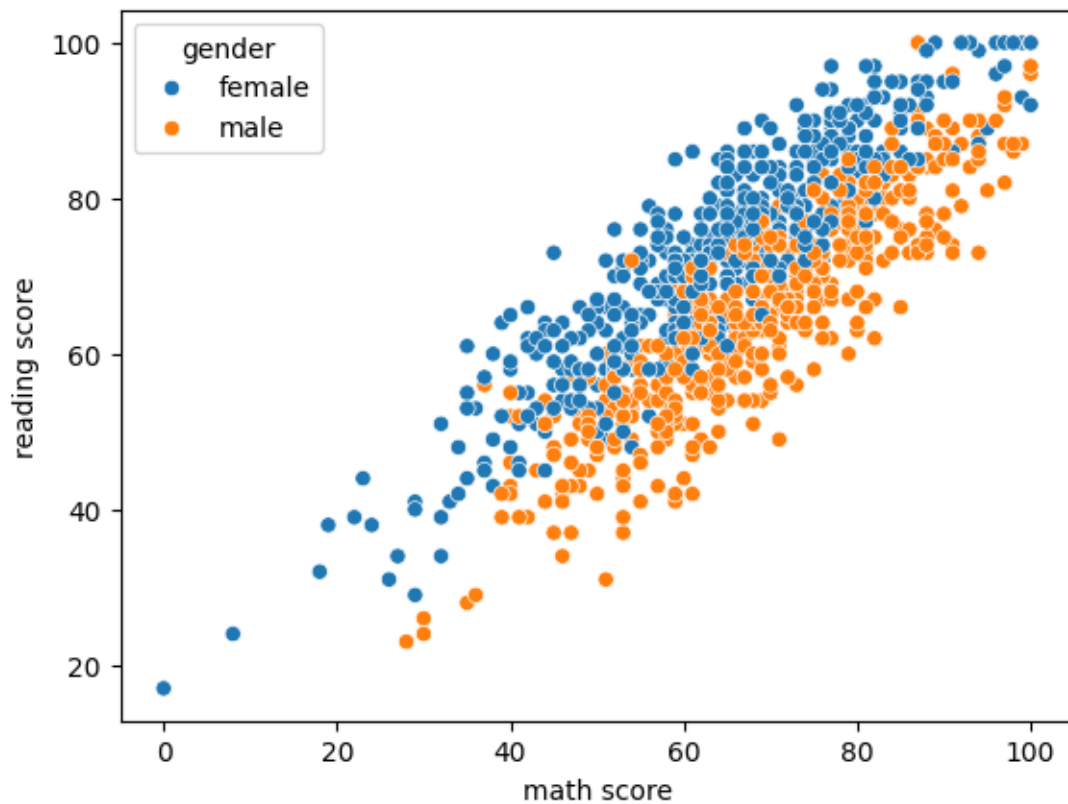
#	Column	Non-Null Count	Dtype
0	gender	1000 non-null	object
1	race/ethnicity	1000 non-null	object
2	parental level of education	1000 non-null	object
3	lunch	1000 non-null	object
4	test preparation course	1000 non-null	object
5	math score	1000 non-null	int64
6	reading score	1000 non-null	int64
7	writing score	1000 non-null	int64

dtypes: int64(3), object(5)

memory usage: 62.6+ KB

```
[6]: sns.scatterplot(data=studenti, x = "math score", y = "reading score", hue = "gender")
```

```
[6]: <Axes: xlabel='math score', ylabel='reading score'>
```



```
[7]: studenti["gender"].value_counts()
```

```
[7]: gender
     female    518
     male      482
     Name: count, dtype: int64
```

```
[8]: studenti["race/ethnicity"].value_counts()
```

```
[8]: race/ethnicity
     group C    319
     group D    262
     group B    190
     group E    140
     group A     89
     Name: count, dtype: int64
```

```
[9]: studenti.lunch.value_counts()
```

```
[9]: lunch
     standard    645
     free/reduced 355
     Name: count, dtype: int64
```

```
[10]: studenti["parental level of education"].value_counts()
```

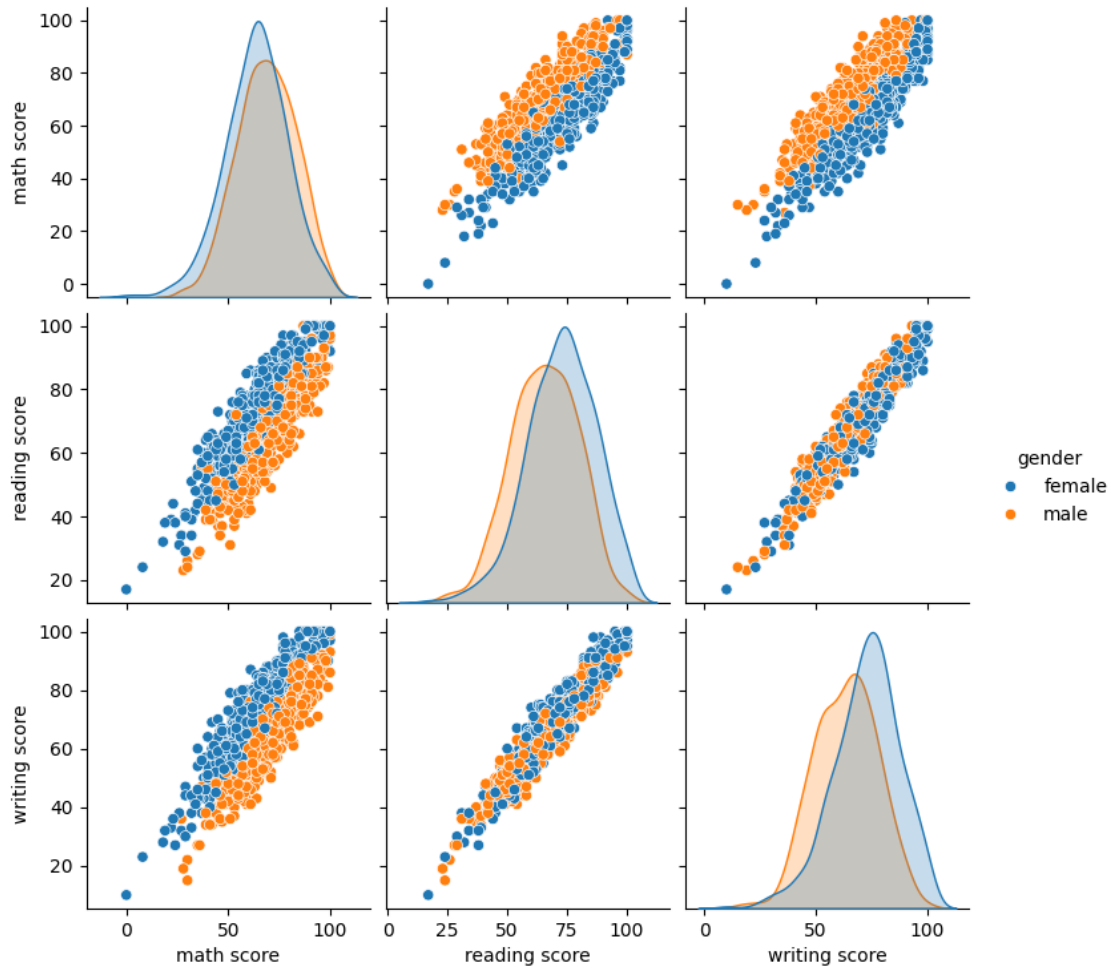
```
[10]: parental level of education
     some college    226
     associate's degree 222
     high school    196
     some high school 179
     bachelor's degree 118
     master's degree   59
     Name: count, dtype: int64
```

```
[11]: studenti["test preparation course"].value_counts()
```

```
[11]: test preparation course
     none        642
     completed   358
     Name: count, dtype: int64
```

```
[12]: sns.pairplot(data=studenti, hue="gender")
```

```
[12]: <seaborn.axisgrid.PairGrid at 0x7f386cee1210>
```



```
[13]: studenti_spol = studenti.groupby("gender")
```

```
[14]: studenti_spol["math score"].mean()
```

```
[14]: gender
female    63.633205
male      68.728216
Name: math score, dtype: float64
```

```
[15]: studenti_spol[["math score", "reading score", "writing score"]].mean()
```

```
[15]:      math score  reading score  writing score
gender
female    63.633205      72.608108      72.467181
male      68.728216      65.473029      63.311203
```

```
[16]: studenti_spol["lunch"].value_counts()
```

```
[16]: gender  lunch
      female  standard      329
           free/reduced    189
      male    standard      316
           free/reduced    166
      Name: count, dtype: int64
```

```
[17]: studenti_rasa_roditelji = studenti.groupby(["race/ethnicity", "parental level_
      ↪of education"])
```

```
[18]: studenti_rasa_roditelji ["math score"].mean()
```

```
[18]: race/ethnicity  parental level of education
      group A
           associate's degree      61.000000
           bachelor's degree      67.166667
           high school            60.444444
           master's degree       57.666667
           some college          63.888889
           some high school      58.916667
      group B
           associate's degree      66.097561
           bachelor's degree      69.300000
           high school            59.791667
           master's degree       67.166667
           some college          63.189189
           some high school      61.815789
      group C
           associate's degree      66.730769
           bachelor's degree      68.150000
           high school            60.906250
           master's degree       67.052632
           some college          65.130435
           some high school      60.551020
      group D
           associate's degree      67.600000
           bachelor's degree      67.571429
           high school            62.863636
           master's degree       72.521739
           some college          68.731343
           some high school      66.760000
      group E
           associate's degree      74.897436
           bachelor's degree      76.555556
           high school            70.772727
           master's degree       74.625000
           some college          73.828571
           some high school      72.111111
      Name: math score, dtype: float64
```