

```
import pandas as pd
df = pd.read_csv("StudentsPerformance.csv")
```

```
print(df.head())

   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
```

```
print(df.tail())

   gender race/ethnicity parental level of education      lunch \
995  female        group E          master's degree  standard
996   male        group C          high school  free/reduced
997  female        group C          high school  free/reduced
998  female        group D           some college  standard
999  female        group D           some college  free/reduced

   test preparation course  math score  reading score  writing score
995        completed            88         99            95
996             none            62         55            55
997        completed            59         71            65
998        completed            68         78            77
999             none            77         86            86
```

```
print(df.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
None
```

```
print(df.shape)

(1000, 8)
```

```
high_math = df[df["math score"] > 80]

# Female students only
female_students = df[df["gender"] == "female"]

# Students with reading score > 70 and writing score > 70
good_read_write = df[(df["reading score"] > 70) & (df["writing score"] > 70)]

# Selecting specific columns
subset = df[["gender", "math score", "reading score"]]
```

```
# Mean
print(df["math score"].mean())

# Median
print(df["math score"].median())

# Mode
print(df["math score"].mode()[0])
```

66.089
66.0
65

```
math_range = df["math score"].max() - df["math score"].min()
print("Range:", math_range)
```

```
# Variance
print("Variance:", df["math score"].var())
```

```
# Standard Deviation
print("Standard Deviation:", df["math score"].std())
```

```
Range: 100
Variance: 229.91899799799847
Standard Deviation: 15.163080096009468
```

```
print(df.describe().T)
```

	count	mean	std	min	25%	50%	75%	max
math score	1000.0	66.089	15.163080	0.0	57.00	66.0	77.0	100.0
reading score	1000.0	69.169	14.600192	17.0	59.00	70.0	79.0	100.0
writing score	1000.0	68.054	15.195657	10.0	57.75	69.0	79.0	100.0