```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

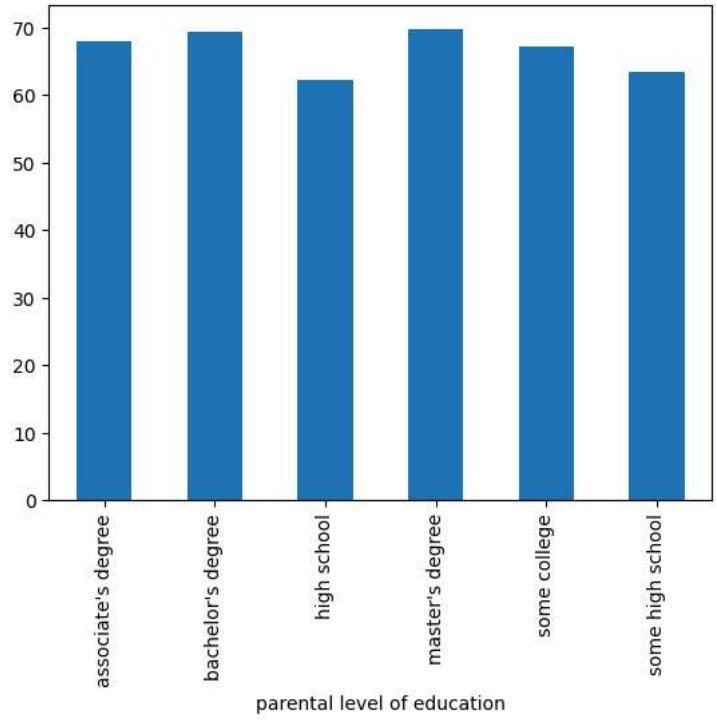
```
df = pd.read_csv("/content/StudentsPerformance.csv")
df.head(3)
```

	gender	race/ethnicity	parental level of education	lunch	test preparation course	math score	reading score	writing score
0	female	group B	bachelor's degree	standard	none	72	72	74
1	female	group C	some college	standard	completed	69	90	88
2	female	group B	master's degree	standard	none	90	95	93

Next steps: Generate code with df View recommended plots

```
df.groupby('parental level of education')['math score'].mean().plot(kind='bar')
```

```
<Axes: xlabel='parental level of education'>
```



Observations :

1. Master's degree parental level of education child get more Math Score on an average..

2. High School parental level of education child get less Math Score on an average

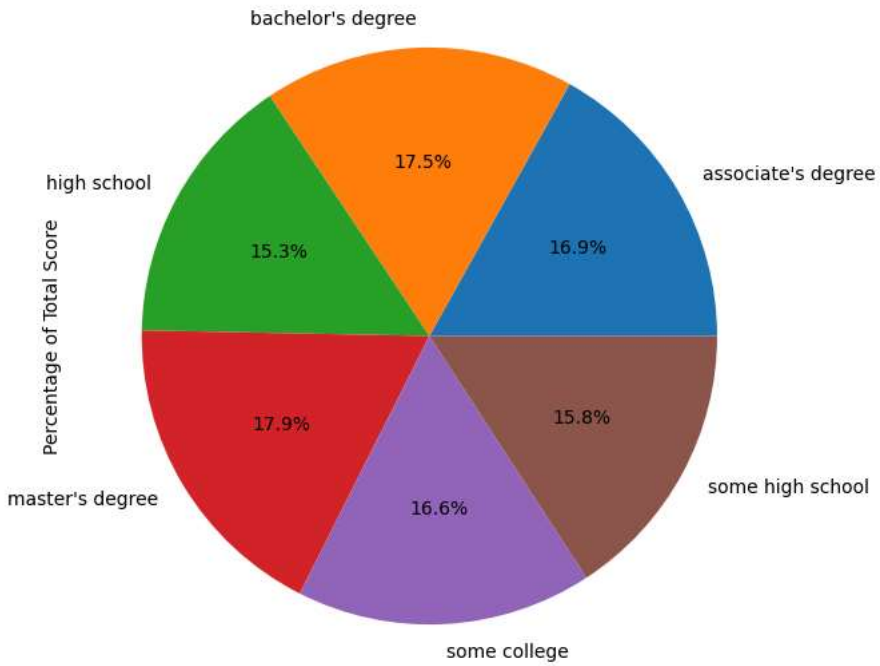
```
df["Percentage of Total Score"] = (df["math score"]+df["reading score"]+df["writing score"])/300*100
df.head()
```

	gender	race/ethnicity	parental level of education	lunch	test preparation course	math score	reading score	writing score
0	female	group B	bachelor's degree	standard	none	72	72	74
1	female	group C	some college	standard	completed	69	90	88
2	female	group B	master's	standard	none	90	95	93

Next steps: Generate code with df View recommended plots

```
plt.figure(figsize=(8,7))
df.groupby('parental level of education')['Percentage of Total Score'].mean().plot(kind='pie',autopct='%1.1f%%')
```

```
<Axes: ylabel='Percentage of Total Score'>
```



Observations:

1. Master's Degree -> parental level of education have more total_score.

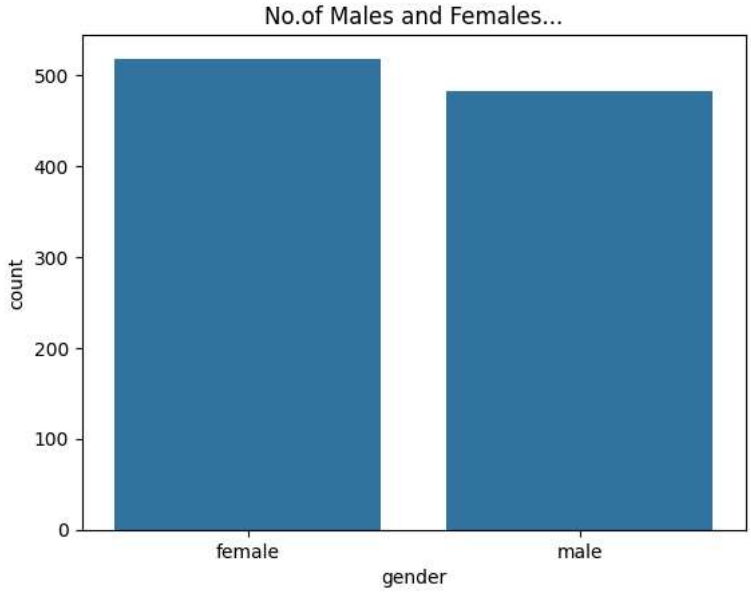
2. High School -> parental level of education have least total_score.

```
df.value_counts('gender')
```

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

```
plt.title("No.of Males and Females...")
sns.countplot(x=df["gender"])
```

```
<Axes: title={'center': 'No.of Males and Females...'}, xlabel='gender',
ylabel='count'>
```



1. Here , There is more no.of females than male

```
df.describe()
```

	math score	reading score	writing score	Percentage of Total Score
count	1000.00000	1000.000000	1000.000000	1000.000000
mean	66.08900	69.169000	68.054000	67.770667
std	15.16308	14.600192	15.195657	14.257326
min	0.00000	17.000000	10.000000	9.000000
25%	57.00000	59.000000	57.750000	58.333333
50%	66.00000	70.000000	69.000000	68.333333
75%	77.00000	79.000000	79.000000	77.666667
max	100.00000	100.000000	100.000000	100.000000

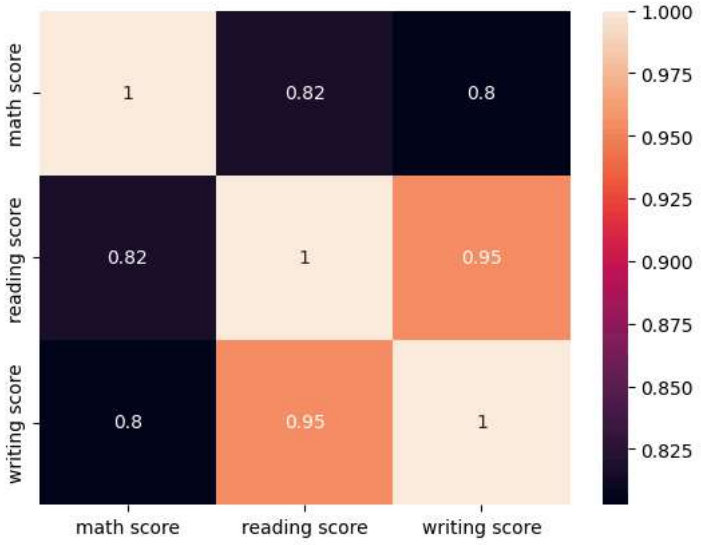
```
df1 = df[["math score","reading score","writing score"]].corr()
df1
```

	math score	reading score	writing score
math score	1.000000	0.817580	0.802642
reading score	0.817580	1.000000	0.954598
writing score	0.802642	0.954598	1.000000

Next steps: Generate code with df1 View recommended plots

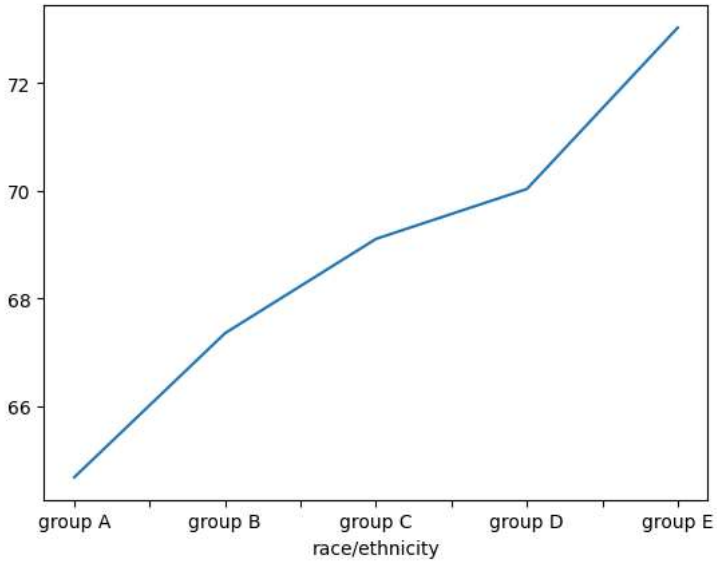
```
sns.heatmap(df1,annot=True)
```

```
<Axes: >
```




```
df.groupby('race/ethnicity')['reading score'].mean().plot(kind='line')
```

```
<Axes: xlabel='race/ethnicity'>
```



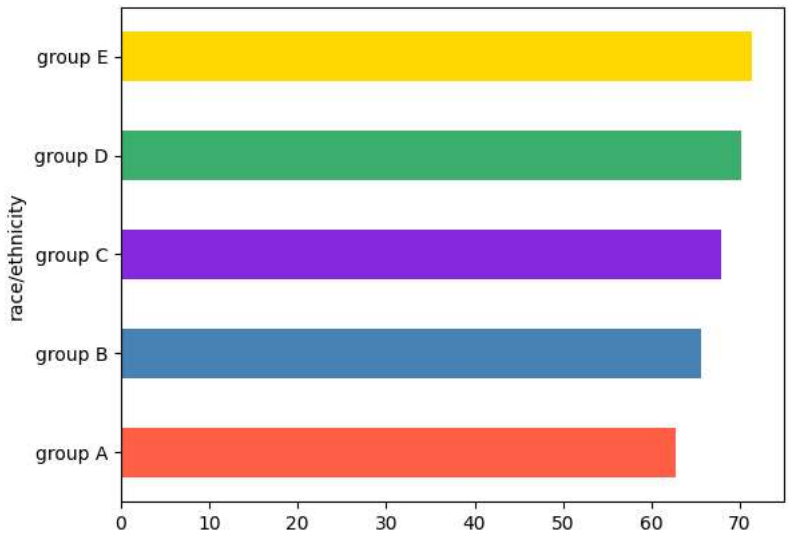
Observations :

1.Here, In Group-E have high Reading Score..

2.In Group-A have low reading score..

```
colors = ['#FF6347', '#4682B4', '#8A2BE2', '#3CB371', '#FFD700']
df.groupby('race/ethnicity')['writing score'].mean().plot(kind='barh',color=colors)
```

```
<Axes: ylabel='race/ethnicity'>
```

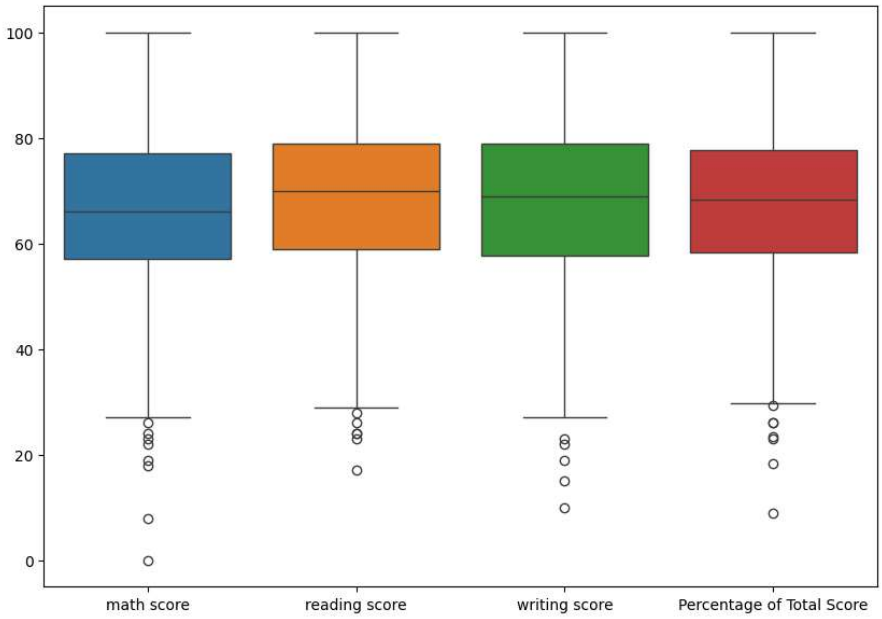


Observations :

Here, In Group-E have high Writing Score & Group-A have low writing score

```
plt.figure(figsize=(10,7))
sns.boxplot(df)
```

```
<Axes: >
```

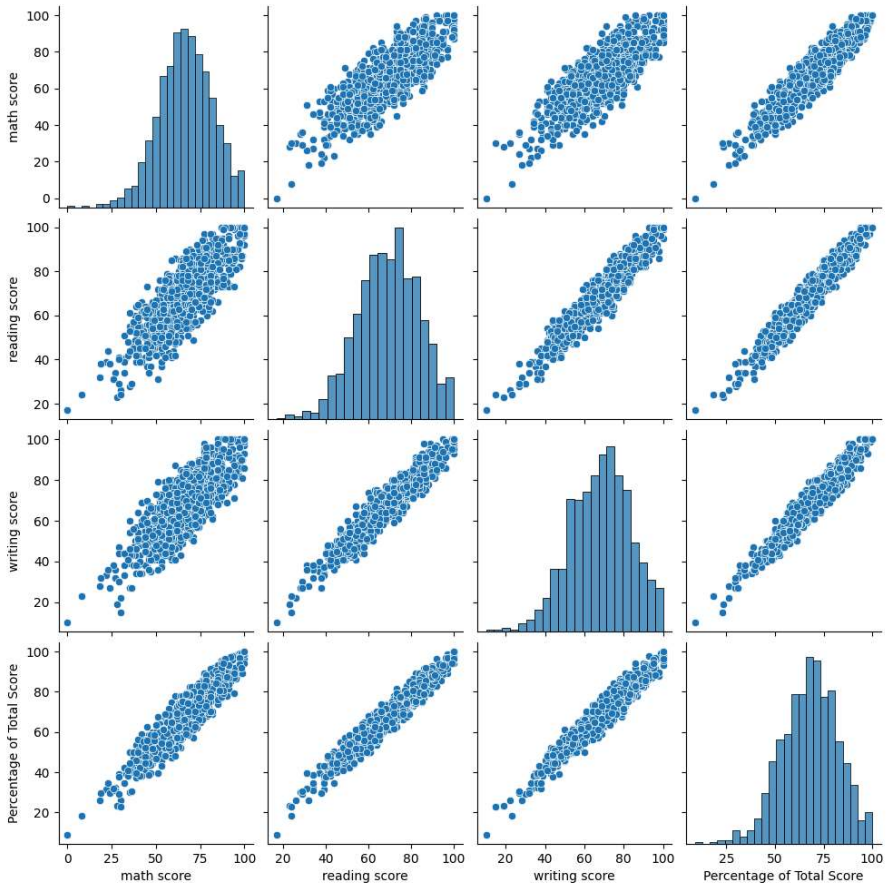


Observations :

These are the outliers of different Columns..

```
sns.pairplot(df)
```

```
<seaborn.axisgrid.PairGrid at 0x7fcf2bff06a0>
```

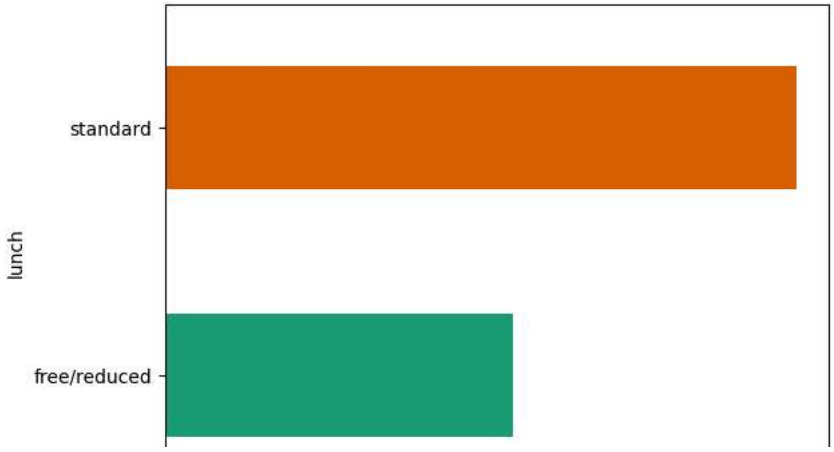



```
df.tail()
```

	gender	race/ethnicity	parental level of education	lunch	test preparation course	math score	reading score	writin scor
995	female	group E	master's degree	standard	completed	88	99	9
996	male	group C	high school	free/reduced	none	62	55	5
997	female	group C	high	free/reduced	completed	59	71	6

```
df.groupby('lunch').size().plot(kind='barh', color=sns.palettes.mpl_palette('Dark2'))
```

```
<Axes: ylabel='lunch'>
```

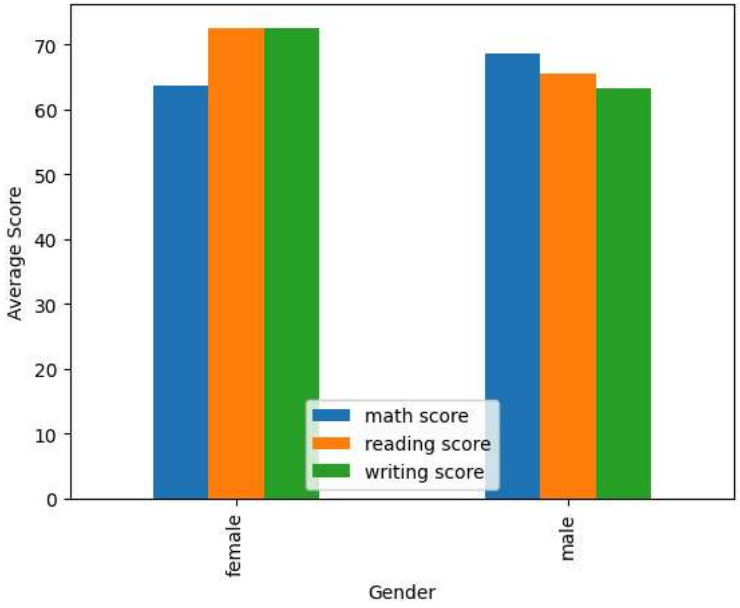


Observations :

Here , Most of the students are having the Standard price lunch at school..

```
df.groupby('gender')[['math score', 'reading score', 'writing score']].mean().plot(kind='bar')
plt.xlabel('Gender')
plt.ylabel('Average Score')
```

```
Text(0, 0.5, 'Average Score')
```



1. In females if we observe the reading & writing scores are same..

2. Whereas in males the math score is highest..

```
Start coding or generate with AI.
```