In [ ]:
```
EDA:
Understand the data
clean the data
Analysis of relationship between variables
```

In [59]:
```
import numpy as np
import pandas as pd
import seaborn as sns

import warnings
warnings.filterwarnings('ignore')
```

In [30]:
```
student = pd.read_csv('c://Users//Chitra//Desktop//Datasets//students.csv')
```

In [31]:
```
# Cleaning the headers
student.columns=student.columns.str.replace(" ","_")
student.columns
```

Out[31]:
```
Index(['gender', 'race', 'parental_level_of_education', 'lunch',
       'test_preparation_course', 'math_score', 'reading_score',
       'writing_score'],
      dtype='object')
```

In [41]:
```
# dropping the unwanted columns
student = student.drop(['parental_level_of_education','race'],axis=1)
```

In [42]:
```
student.head()
```

Out[42]:

	gender	lunch	test_preparation_course	math_score	reading_score	writing_score
0	male	standard	complete	81	85	90
1	female	free	complete	82	86	91
2	male	reduced	none	83	87	92
3	female	standard	none	84	88	93
4	male	free	complete	85	89	94

In [53]:
```
student_new=student[['math_score','reading_score','writing_score']]
```

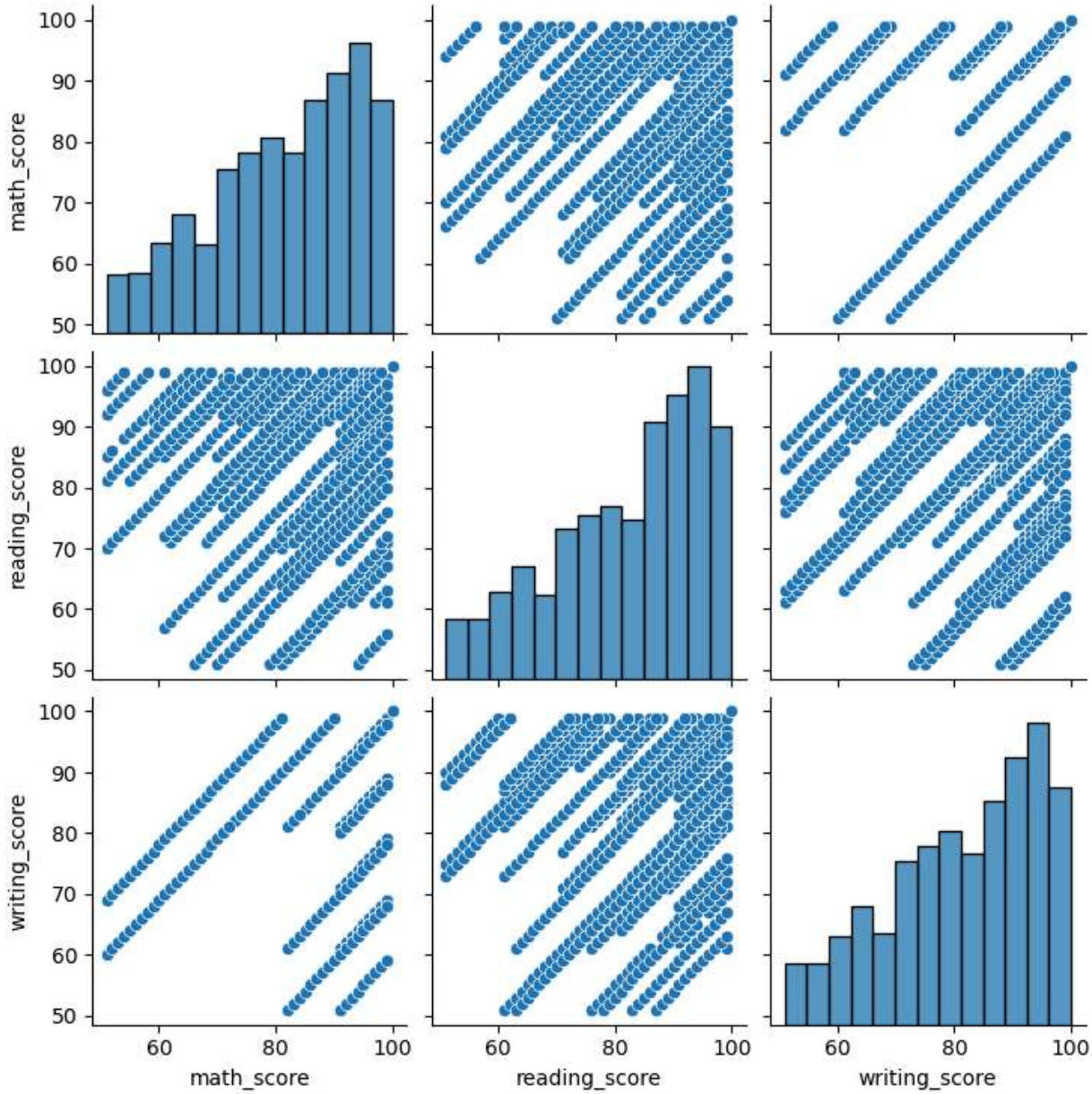
In [56]:
```
# Relationship analysis
corelation = student_new.corr()
```

In [57]:
```
sns.heatmap(corelation,xticklabels=corelation.columns,yticklabels=corelation.columns,a
```

Out[57]:
```
<Axes: >
```

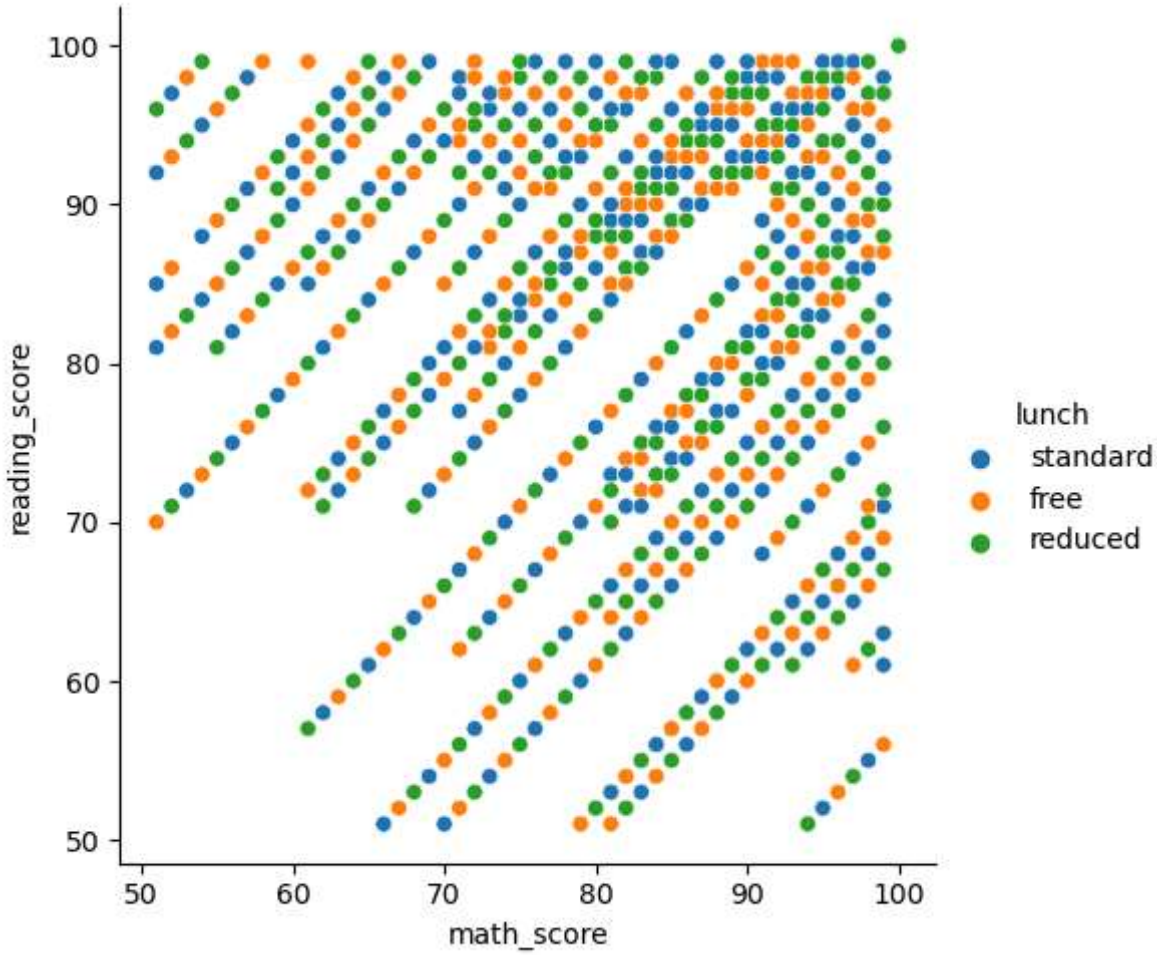
In [60]:
```
sns.pairplot(student)
```

Out[60]:
```
<seaborn.axisgrid.PairGrid at 0x20511fa4490>
```



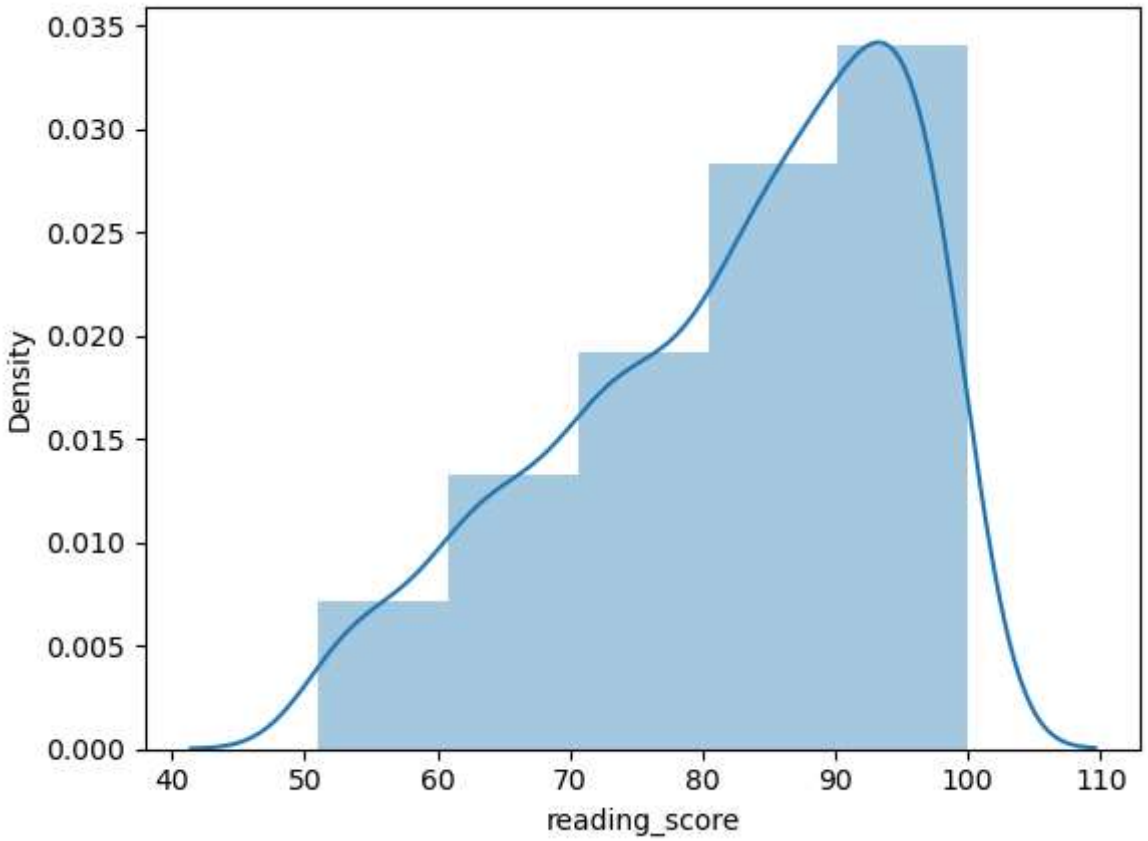
In [62]:
```
sns.relplot(x='math_score',y='reading_score',hue='lunch',data=student)
```

Out[62]:
```
<seaborn.axisgrid.FacetGrid at 0x20513658610>
```



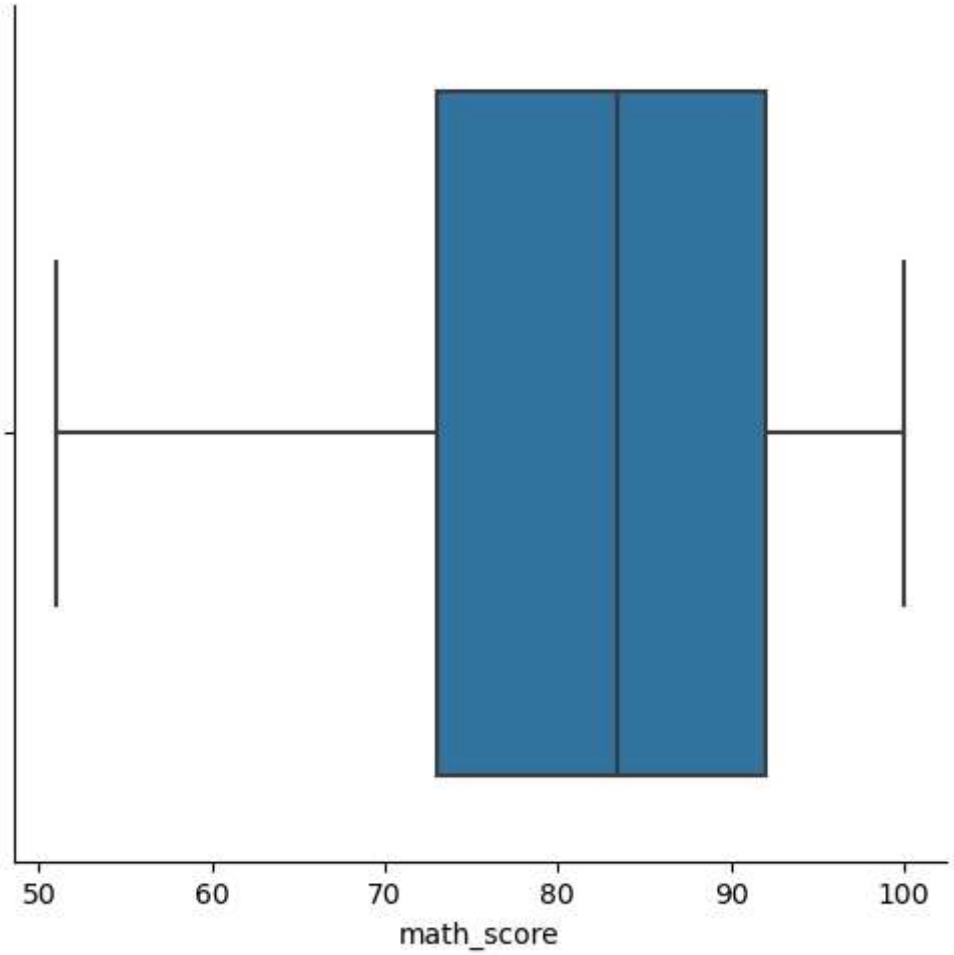
In [64]:
```
sns.distplot(student['reading_score'],bins=5)
```

Out[64]:
```
<Axes: xlabel='reading_score', ylabel='Density'>
```



In [66]:
```
sns.catplot(x='math_score',kind='box',data=student)
```

Out[66]:
```
<seaborn.axisgrid.FacetGrid at 0x20515ad8c10>
```


50
60
70
80
90
100
math_score