In [1]:
```
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

In [2]:
```
import pandas as pd
```

In [3]:
```
data = pd.read_csv("C:/Users/DEEPIKA/OneDrive/Desktop/Students.csv")
```

In [4]:
```
# understanding the data
data.head()
```

Out[4]:

	Gender	EthnicGroup	ParentEduc	LunchType	TestPrep	MathScore	ReadingScore	WritingS
0	female	group B	bachelor's degree	standard	none	72	72	
1	female	group C	some college	standard	completed	69	90	
2	female	group B	master's degree	standard	none	90	95	
3	male	group A	associate's degree	free/reduced	none	47	57	
4	male	group C	some college	standard	none	76	78	

In [5]:
```
data.tail()
```

Out[5]:

	Gender	EthnicGroup	ParentEduc	LunchType	TestPrep	MathScore	ReadingScore	Writ
30636	male	group C	some high school	standard	none	56	47	
30637	male	group E	associate's degree	free/reduced	none	74	75	
30638	male	group C	some college	standard	none	36	29	
30639	male	group A	some high school	free/reduced	completed	43	34	
30640	female	group D	associate's degree	standard	none	52	68	

In [6]:
```
data.shape
```

Out[6]:
```
(30641, 8)
```

In [7]:
```
data.describe()
```

Out[7]:

	MathScore	ReadingScore	WritingScore
count	30641.000000	30641.000000	30641.000000
mean	66.749355	69.624980	68.468327
std	15.206049	14.671572	15.307814
min	0.000000	10.000000	5.000000
25%	56.000000	60.000000	58.000000
50%	67.000000	70.000000	69.000000
75%	78.000000	80.000000	79.000000
max	100.000000	100.000000	100.000000

In [8]:
```
data.columns
```

Out[8]:
```
Index(['Gender', 'EthnicGroup', 'ParentEduc', 'LunchType', 'TestPrep',
       'MathScore', 'ReadingScore', 'WritingScore'],
      dtype='object')
```

In [9]:
```
data.nunique()
```

Out[9]:
```
Gender           2
EthnicGroup      5
ParentEduc       6
LunchType        2
TestPrep         2
MathScore       94
ReadingScore    88
WritingScore    92
dtype: int64
```

In [10]:
```
data['Gender'].unique()
```

Out[10]:
```
array(['female', 'male'], dtype=object)
```

In [11]:
```
data['EthnicGroup'].unique()
```

Out[11]:
```
array(['group B', 'group C', 'group A', 'group D', 'group E'],
      dtype=object)
```

In [12]:
```
data.isnull().sum()
```

Out[12]:
```
Gender          0
EthnicGroup     0
ParentEduc      0
LunchType       0
TestPrep        0
MathScore       0
ReadingScore    0
WritingScore    0
dtype: int64
```

In [13]:
```
student = data.drop(['EthnicGroup','ParentEduc',],axis=1)
```

In [14]:
```
student.head()
```

Out[14]:

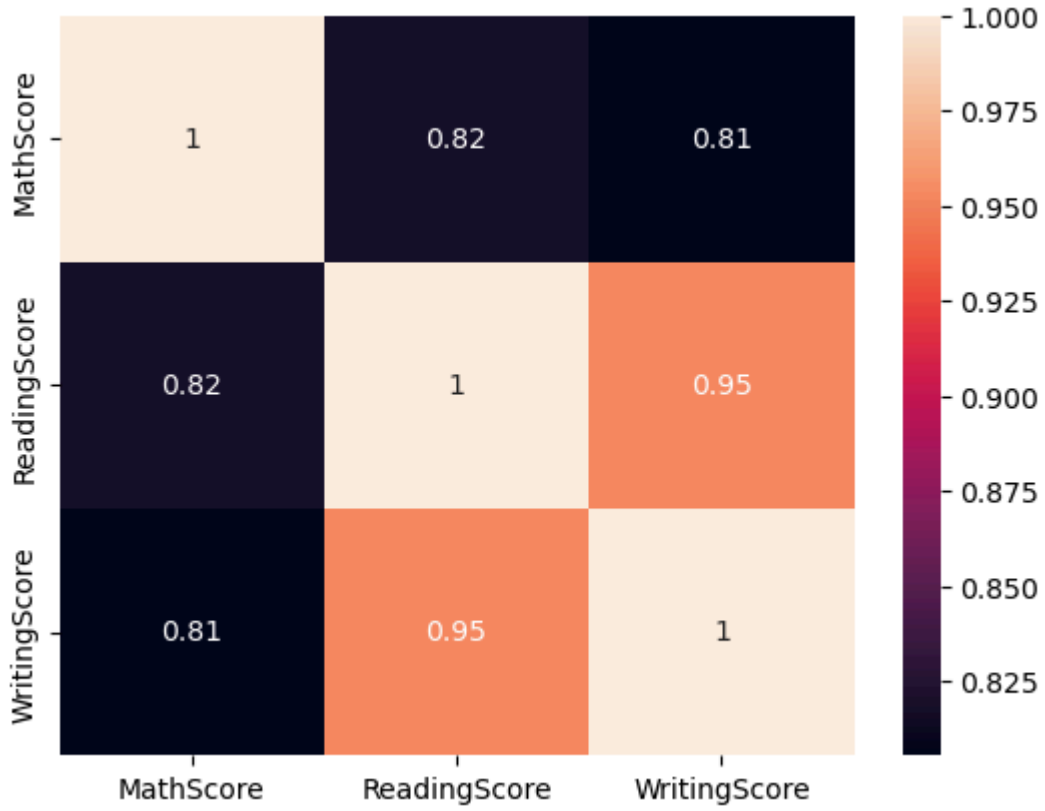
	Gender	LunchType	TestPrep	MathScore	ReadingScore	WritingScore
0	female	standard	none	72	72	74
1	female	standard	completed	69	90	88
2	female	standard	none	90	95	93
3	male	free/reduced	none	47	57	44
4	male	standard	none	76	78	75

In [15]:
```
# Relationship Analysis
```

In [16]:
```
corelation = student.corr()
```

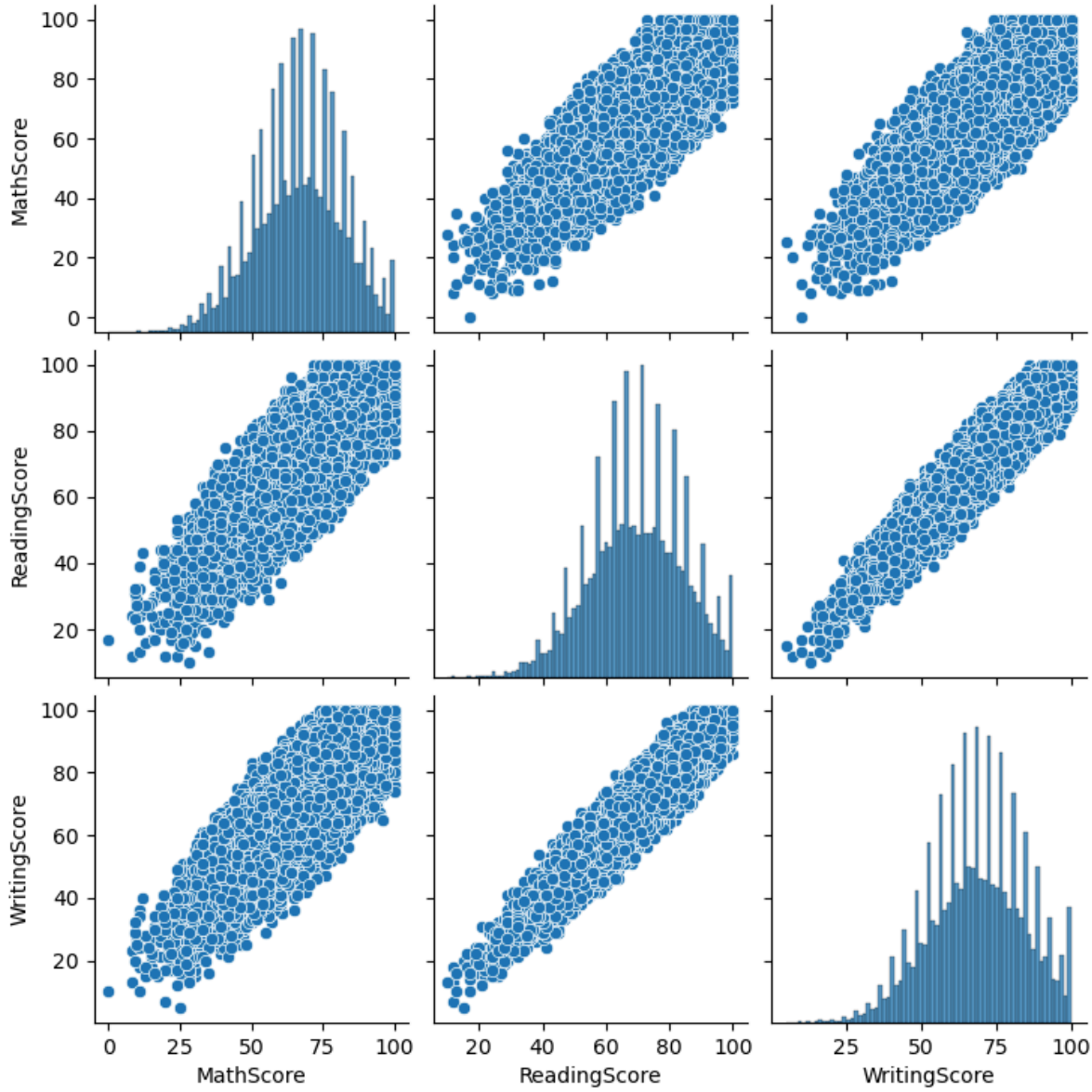
In [17]:
```
sns.heatmap(corelation,xticklabels=corelation.columns,yticklabels=corelation.columr
```

Out[17]:
```
<AxesSubplot: >
```



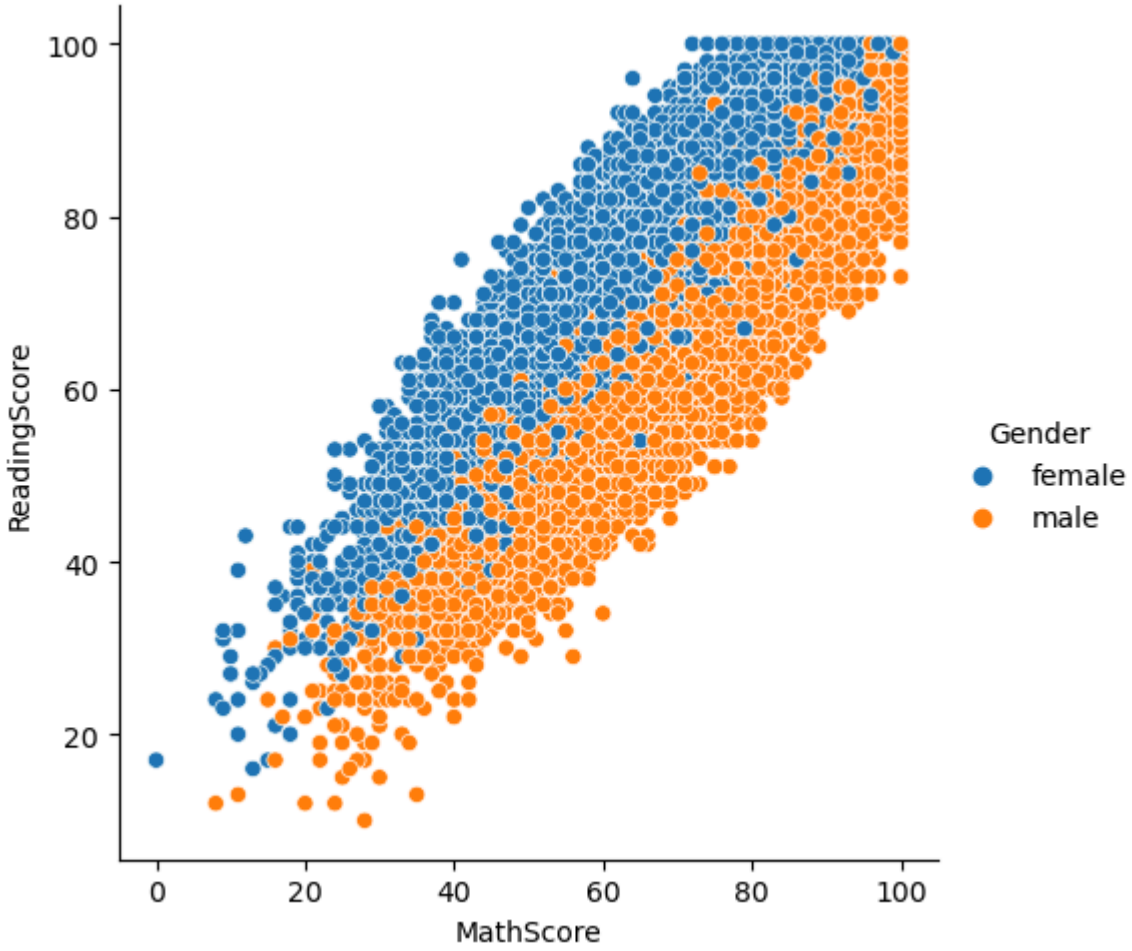
In [18]:
```
sns.pairplot(student)
```

Out[18]:
```
<seaborn.axisgrid.PairGrid at 0x21abebe0790>
```



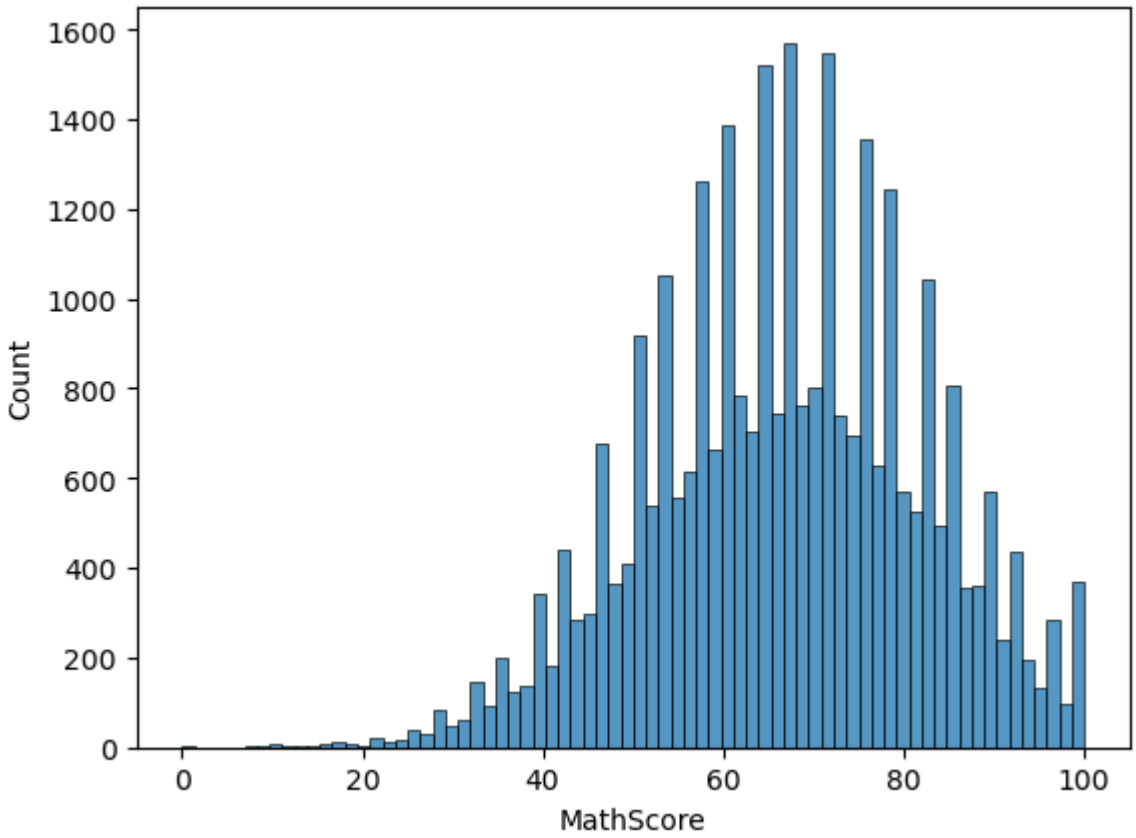
In [19]:
```
#Scatterplot
sns.relplot(x='MathScore', y='ReadingScore', hue='Gender', data=student)
```

Out[19]:
```
<seaborn.axisgrid.FacetGrid at 0x21ac0d6c7c0>
```



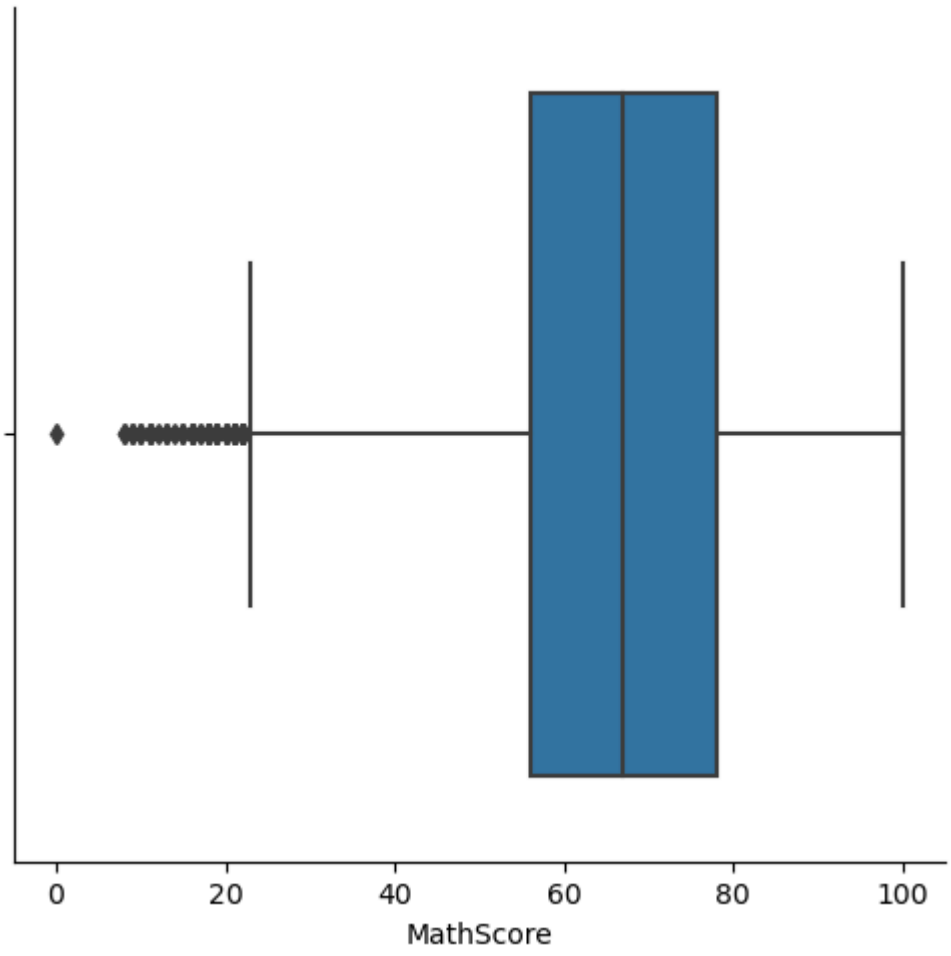
In [20]:
```
sns.histplot(data=student, x='MathScore')
```

Out[20]:
```
<AxesSubplot: xlabel='MathScore', ylabel='Count'>
```



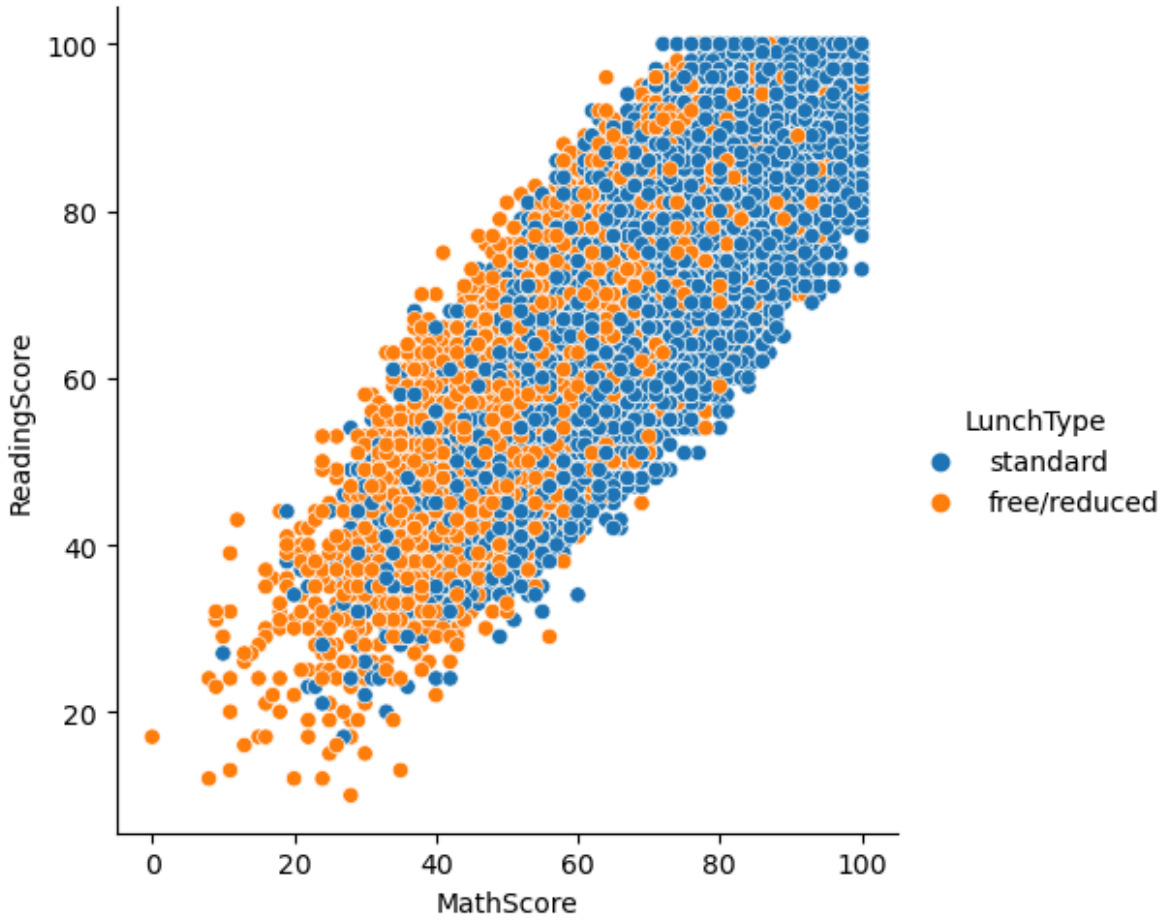
In [21]:
```
sns.catplot(x='MathScore',kind='box',data=student)
```

Out[21]: <seaborn.axisgrid.FacetGrid at 0x21ac10c3c10>



In [22]:
```
sns.relplot(x='MathScore', y='ReadingScore', hue='LunchType', data=student)
```

Out[22]: <seaborn.axisgrid.FacetGrid at 0x21ac1150550>


100
80
60
40
20
ReadingScore
0
20
40
60
80
100
MathScore
LunchType
standard
free/reduced

In [ ]:

In [ ]: