StudentPerformance HypTesting Project

February 7, 2022

```
First 5 rows of data:
   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
Column 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
Stats on Numerical Columns:
       math score  reading score  writing score
count  1000.00000    1000.000000    1000.000000
mean     66.08900      69.169000      68.054000
std      15.16308      14.600192      15.195657
min       0.00000      17.000000      10.000000
25%      57.00000      59.000000      57.750000
50%      66.00000      70.000000      69.000000
```

```
75%        77.00000        79.000000        79.000000
max       100.00000       100.000000       100.000000

standard          645
free_or_reduced   355
Name: lunch, dtype: int64

some college       226
associate degree   222
high school        196
some high school   179
bachelor degree    118
masters degree      59
Name: parent_ed, dtype: int64

['bachelor degree', 'some college', 'masters degree', 'associate degree', 'high
school', 'some high school']
Categories (6, object): ['some high school' < 'high school' < 'some college' <
'associate degree' < 'bachelor degree' < 'masters degree']
```
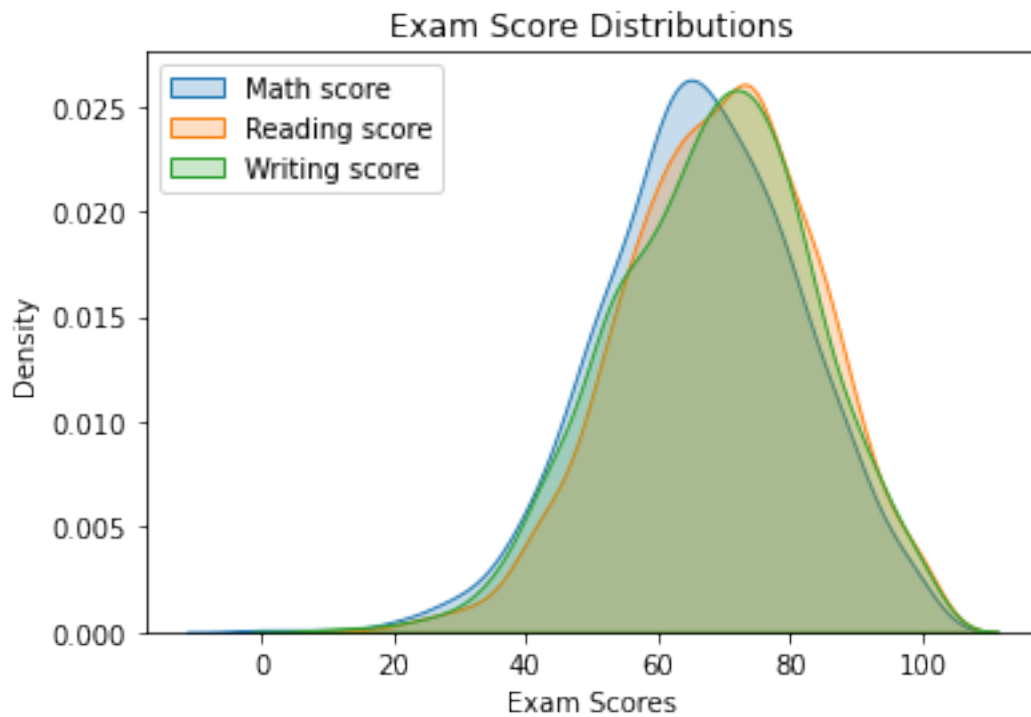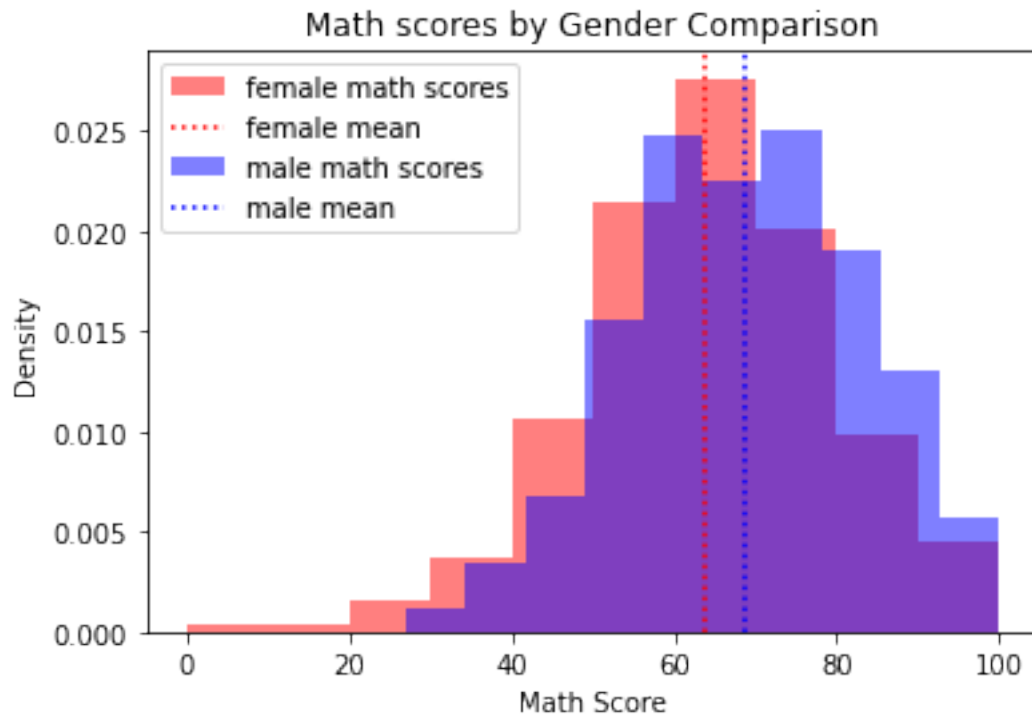


```
<Figure size 432x288 with 0 Axes>

Mean math score of female students: 63.633204633204635
Mean math score of male students: 68.72821576763485
Mean difference: -5.095011134430216
```

Math scores by Gender Comparison

P-Value: 9.120185549328822e-08
There is a statistically significant difference between female and male math
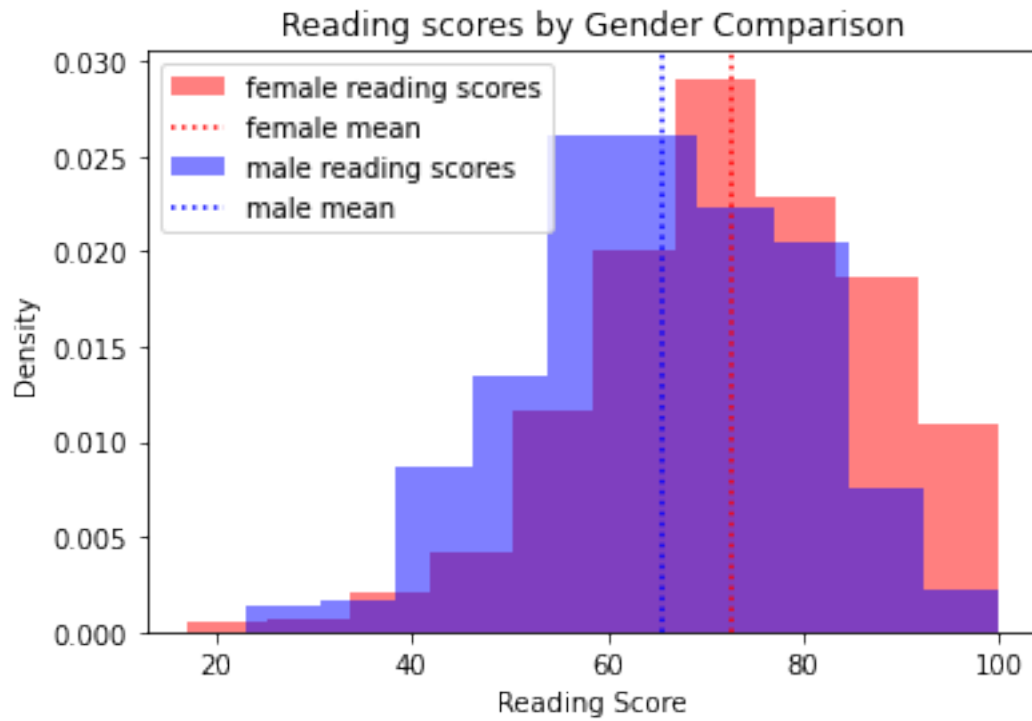scores.
Male students averaged higher with statistical significance.

<Figure size 432x288 with 0 Axes>

Mean reading score of female students: 72.60810810810811
Mean reading score of male students: 65.47302904564316
Mean difference: 7.135079062464953

Reading scores by Gender Comparison

P-Value: 4.680538743933289e-15
There is a statistically significant difference between female and male reading scores.
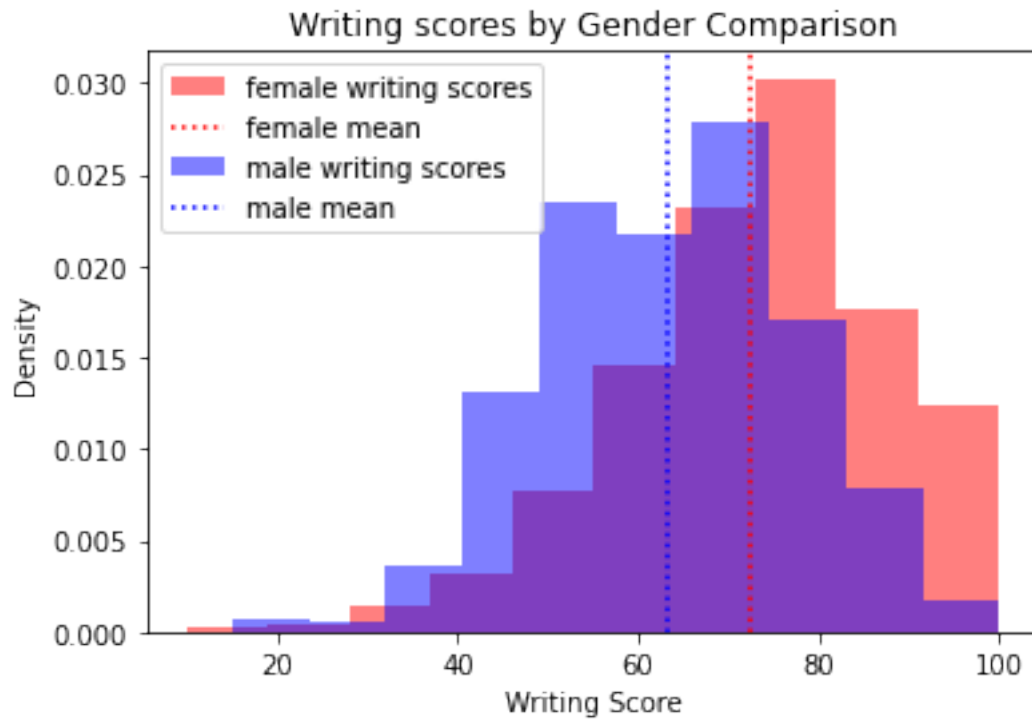Female students averaged higher with statistical significance.

<Figure size 432x288 with 0 Axes>

Mean writing score of female students: 72.46718146718146
Mean writing score of male students: 63.31120331950208
Mean difference: 9.155978147679384

Writing scores by Gender Comparison

P-Value: 2.019877706867934e-22
There is a statistically significant difference between female and male writing scores.
Female students averaged higher with statistical significance.

<Figure size 432x288 with 0 Axes>

Mean math score of standard lunch students: 70.03410852713178
Mean math score of free/reduced lunch students: 58.92112676056338
Mean difference: 11.112981766568396

Math scores by Lunch Status Comparison

P-Value: 2.4131955993137074e-30
There is a statistically significant difference between standard and
free/reduced lunch math scores.
Standard lunch students averaged higher with statistical significance.

<Figure size 432x288 with 0 Axes>

Mean reading score of standard lunch students: 71.65426356589147
Mean reading score of free/reduced lunch students: 64.65352112676057
Mean difference: 7.000742439130903

Reading scores by Lunch Status Comparison

P-Value: 2.0027966545279011e-13
There is a statistically significant difference between standard and
free/reduced lunch reading scores.
Standard lunch students averaged higher with statistical significance.

<Figure size 432x288 with 0 Axes>

Mean writing score of standard lunch students: 70.8232558139535
Mean writing score of free/reduced lunch students: 63.02253521126761
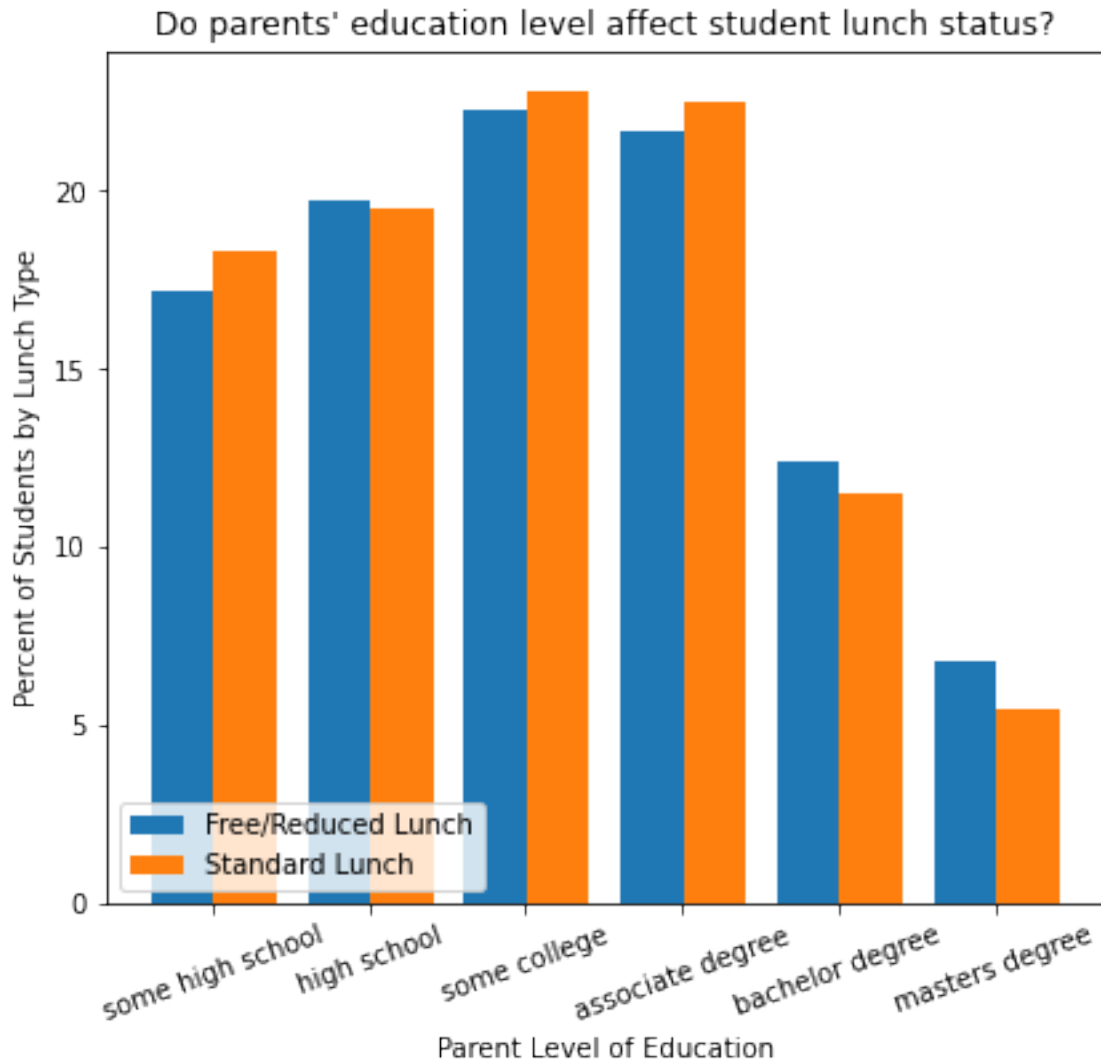Mean difference: 7.800720602685885

P-Value: 3.1861895831664765e-15
There is a statistically significant difference between standard and
free/reduced lunch writing scores.
Standard lunch students averaged higher with statistical significance.

<Figure size 432x288 with 0 Axes>

Do parents' education level affect student lunch status?

P-Value: 0.9531014927218223
There is NOT a statistically significant relationship between parental education level and whether the student receives free/reduced lunch.

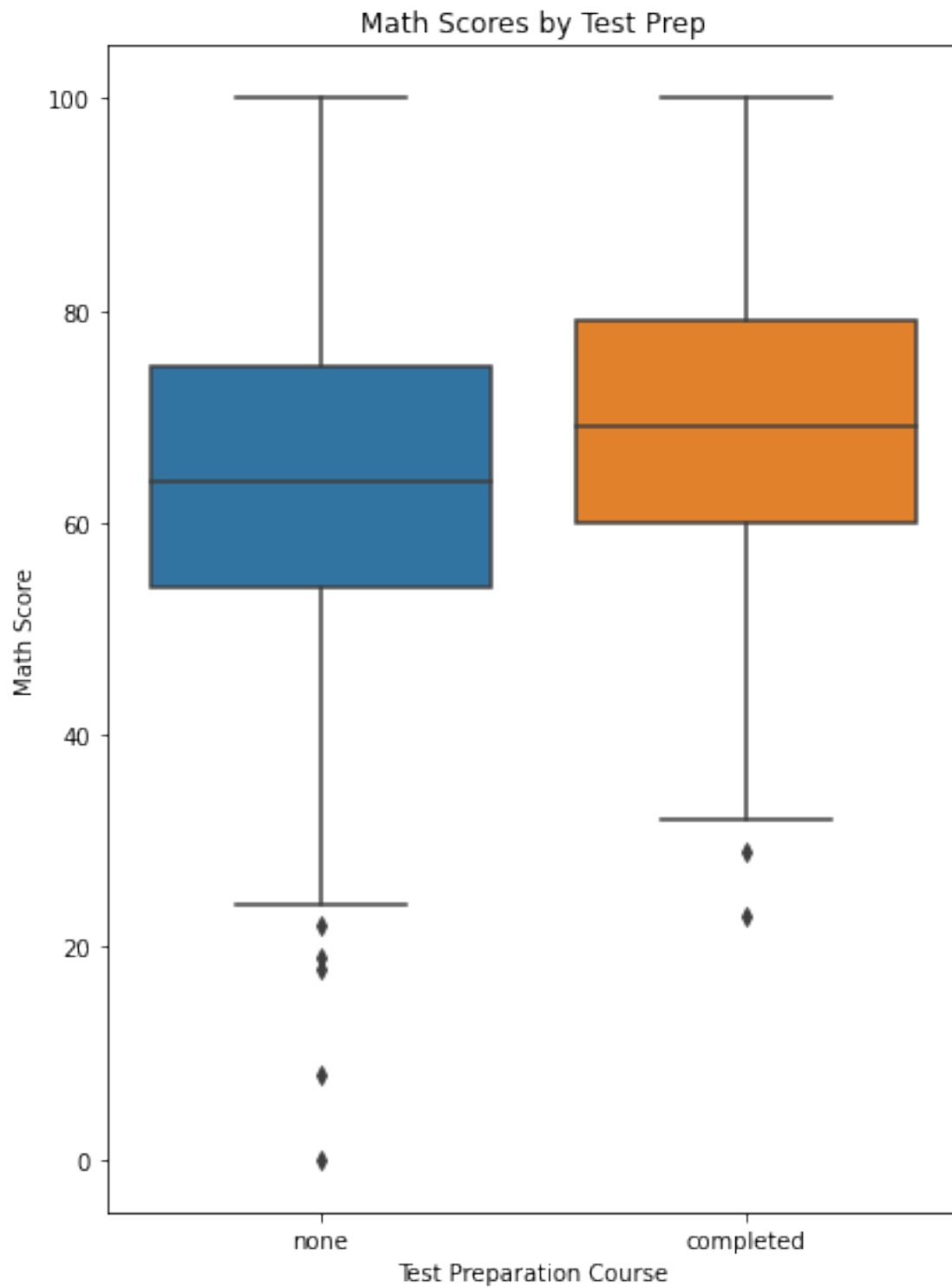<Figure size 432x288 with 0 Axes>

Mean math score of students who completed the Test Prep Course:
69.69553072625699
Mean math score of students who did not complete the Test Prep Course:
64.0778816199377
Mean difference: 5.617649106319291

Math Scores by Test Prep

P-Value: 1.5359134607147415e-08
There is a statistically significant difference between math scores of students
who completed the test prep course vs those who didn't.

Students who completed the test prep course averaged higher with statistical significance.
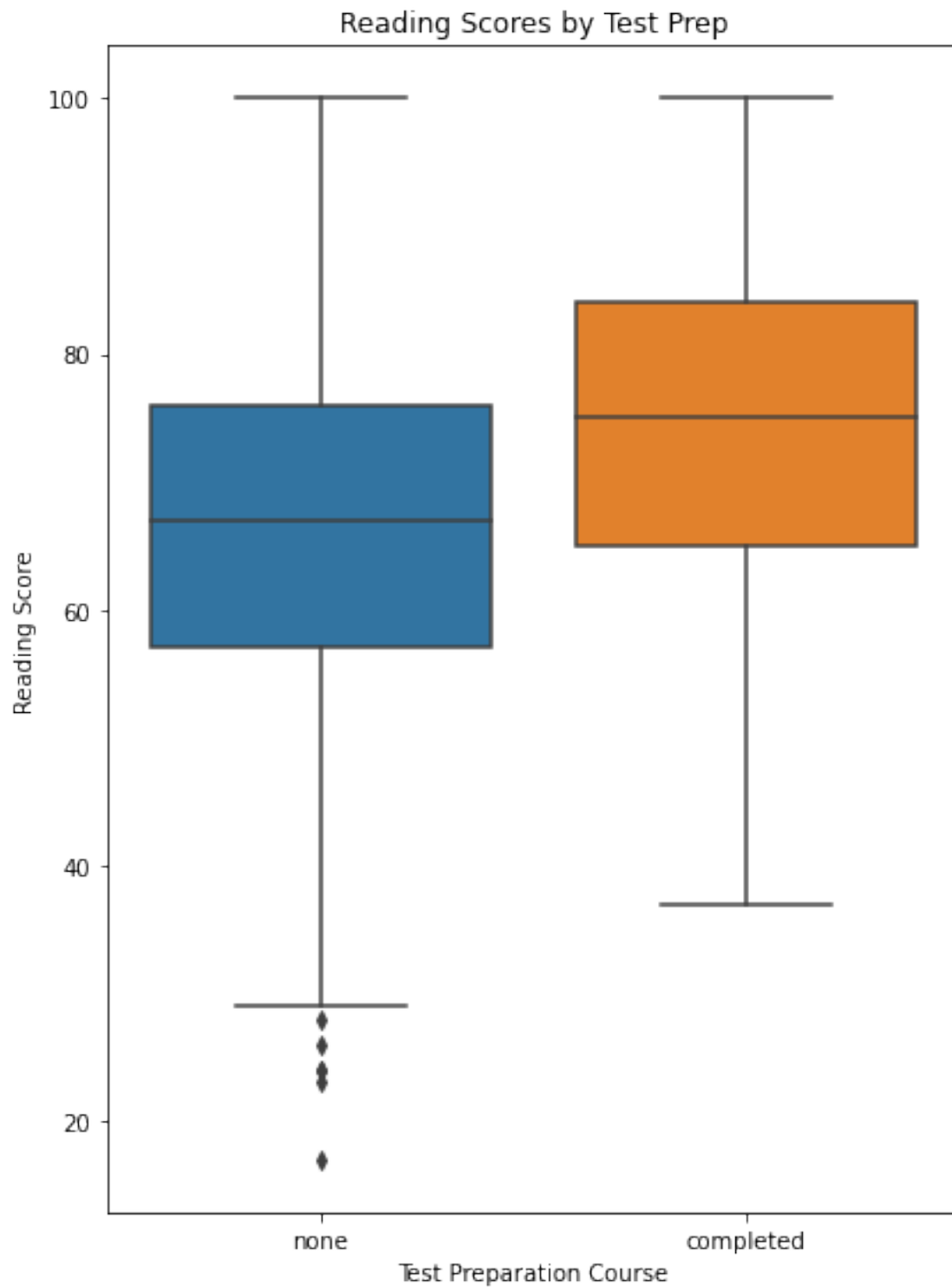

Mean reading score of students who completed the Test Prep Course:
73.89385474860335
Mean reading score of students who did not complete the Test Prep Course:
66.53426791277259
Mean difference: 7.359586835830754

Reading Scores by Test Prep

P-Value: 9.081783336892205e-15
There is a statistically significant difference between reading scores of
students who completed the test prep course vs those who didn't.

Students who completed the test prep course averaged higher with statistical significance.
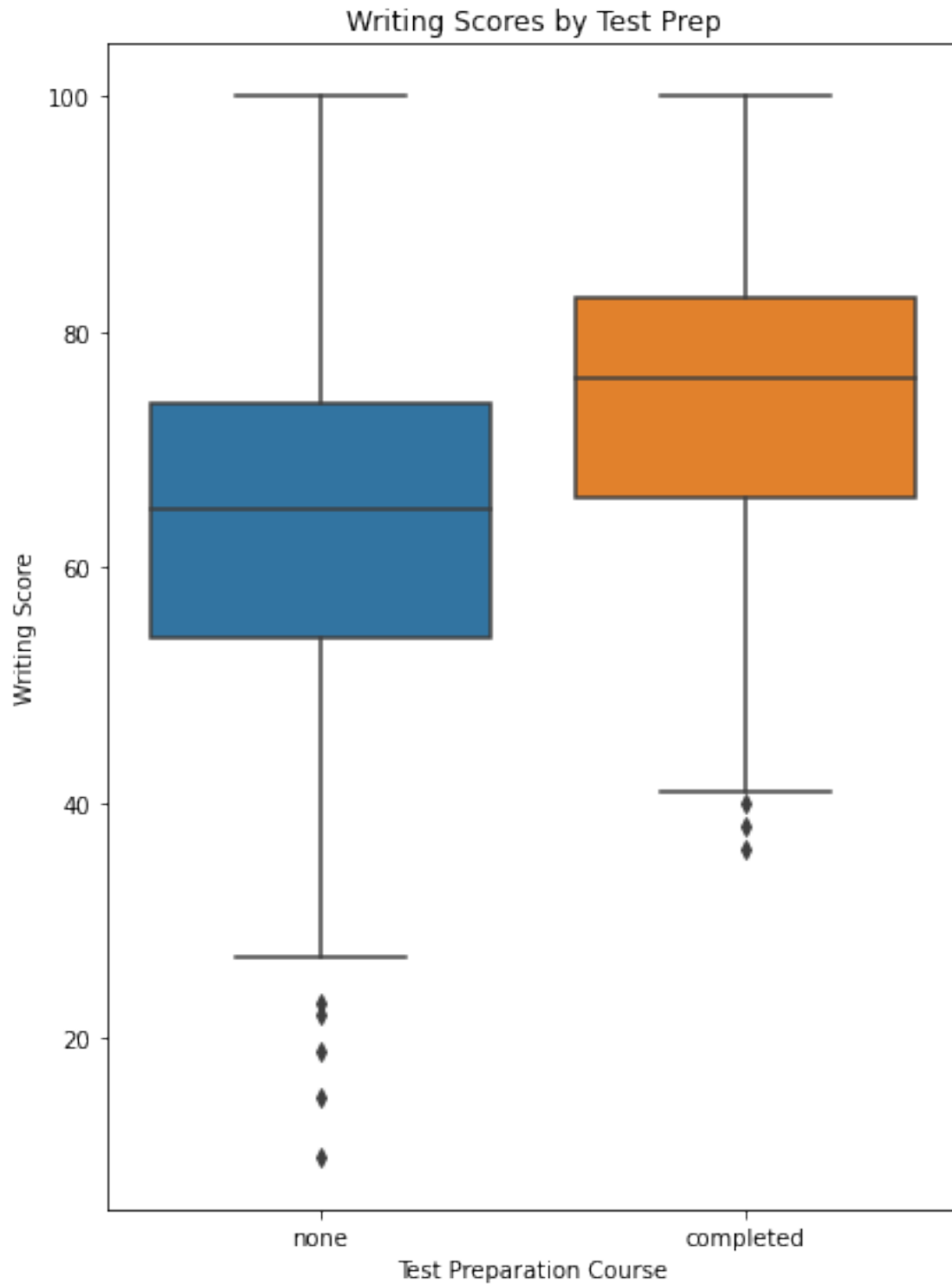

Mean writing score of students who completed the Test Prep Course:
74.41899441340782
Mean writing score of students who did not complete the Test Prep Course:
64.50467289719626
Mean difference: 9.914321516211558

Writing Scores by Test Prep

P-Value: 3.68529173524572e-24
There is a statistically significant difference between writing scores of
students who completed the test prep course vs those who didn't.

Students who completed the test prep course averaged higher with statistical significance.

P-value: 1.3732194030370688e-11
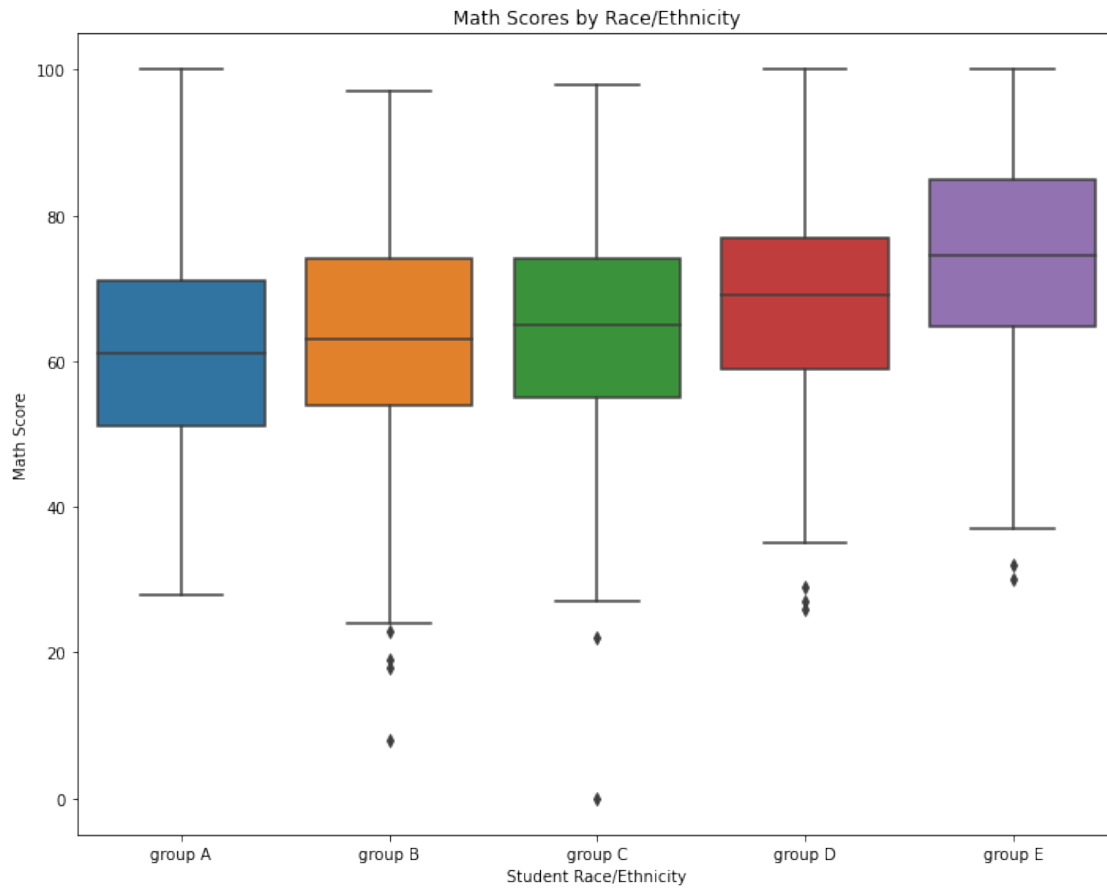There is a statistically significant difference in math scores among the ethnic groups in this data.

```
 Multiple Comparison of Means - Tukey HSD, FWER=0.05
=======================================================
 group1  group2 meandiff p-adj   lower    upper  reject
-------------------------------------------------------
group A group B   1.8234 0.8597   -3.36   7.0068  False
group A group C   2.8347 0.4966 -2.0028   7.6723  False
group A group D   5.7334 0.0138  0.7824  10.6844   True
group A group E  12.1922  0.001  6.7215  17.6629   True
group B group C   1.0113    0.9 -2.6867   4.7094  False
group B group D     3.91 0.0441  0.0647   7.7552   True
group B group E  10.3688  0.001  5.8741  14.8635   True
group C group D   2.8986 0.1287 -0.4659   6.2632  False
group C group E   9.3575  0.001  5.2665  13.4485   True
group D group E   6.4588  0.001  2.2343  10.6834   True
-------------------------------------------------------
```


There are statistically siginificant differences in math scores between:
group A and group D
group A and group E
group B and group D
group B and group E
group C and group E
group D and group E

The largest difference in math scores was 12.1922% between group A and group E

Math Scores by Race/Ethnicity
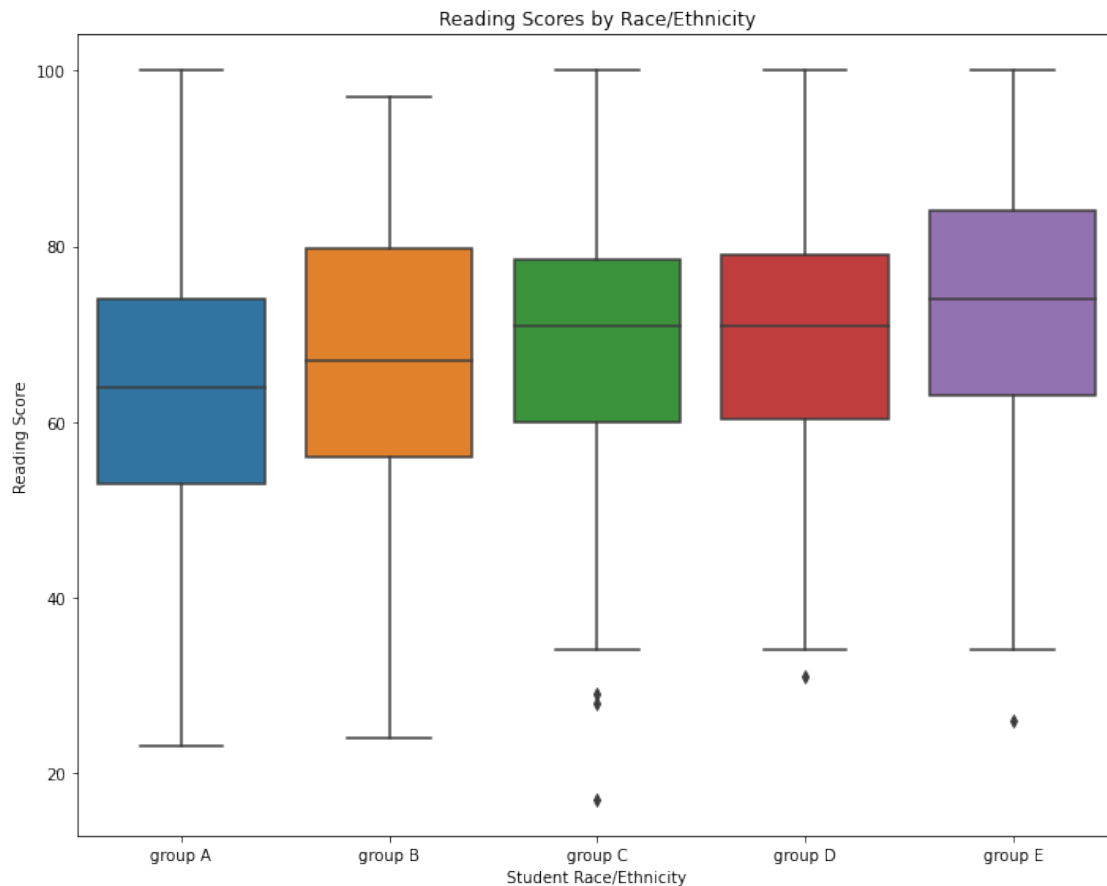
P-value: 0.0001780089103235947
There is a statistically significant difference in reading scores among the
ethnic groups in this data.

```
 Multiple Comparison of Means - Tukey HSD, FWER=0.05
========================================================
 group1  group2 meandiff p-adj   lower   upper  reject
--------------------------------------------------------
group A group B   2.6785 0.5875 -2.3998  7.7567  False
group A group C   4.4293 0.0801 -0.3101  9.1687  False
group A group D   5.3564 0.0219  0.5058 10.2069   True
group A group E   8.3544  0.001  2.9947 13.7141   True
group B group C   1.7508  0.656 -1.8722  5.3738  False
group B group D   2.6779 0.2956 -1.0893  6.4451  False
group B group E   5.6759 0.0041  1.2724 10.0794   True
group C group D   0.9271    0.9 -2.3692  4.2234  False
group C group E   3.9251 0.0583 -0.0829  7.9331  False
group D group E    2.998 0.2768 -1.1408  7.1369  False
--------------------------------------------------------
```

There are statistically siginificant differences in reading scores between:
group A and group D
group A and group E
group B and group E

The largest difference in reading scores was 8.3544% between group A and group E

Reading Scores by Race/Ethnicity

P-value: 1.0979189070067382e-05
There is a statistically significant difference in writing scores among the
ethnic groups in this data.

 Multiple Comparison of Means - Tukey HSD, FWER=0.05
===
 group1 group2 meandiff p-adj   lower   upper   reject
----
group A group B   2.9258 0.5444 -2.3436 8.1953 False
group A group C   5.1534 0.0346 0.2356 10.0712   True
group A group D   7.4709 0.001 2.4377 12.504   True
group A group E   8.733 0.001 3.1715 14.2945   True
group B group C   2.2276 0.4857 -1.5318   5.987 False

```
group B group D    4.545 0.0133    0.636  8.4541    True
group B group E   5.8071 0.0049   1.2379 10.3764    True
group C group D   2.3175 0.3448  -1.1029  5.7378   False
group C group E   3.5796 0.1294  -0.5793  7.7385   False
group D group E   1.2621    0.9  -3.0326  5.5568   False
-------------------------------------------------------
```

There are statistically siginificant differences in writing scores between:
group A and group C
group A and group D
group A and group E
group B and group D
group B and group E

The largest difference in writing scores was 8.733% between group A and group E

Writing Scores by Race/Ethnicity