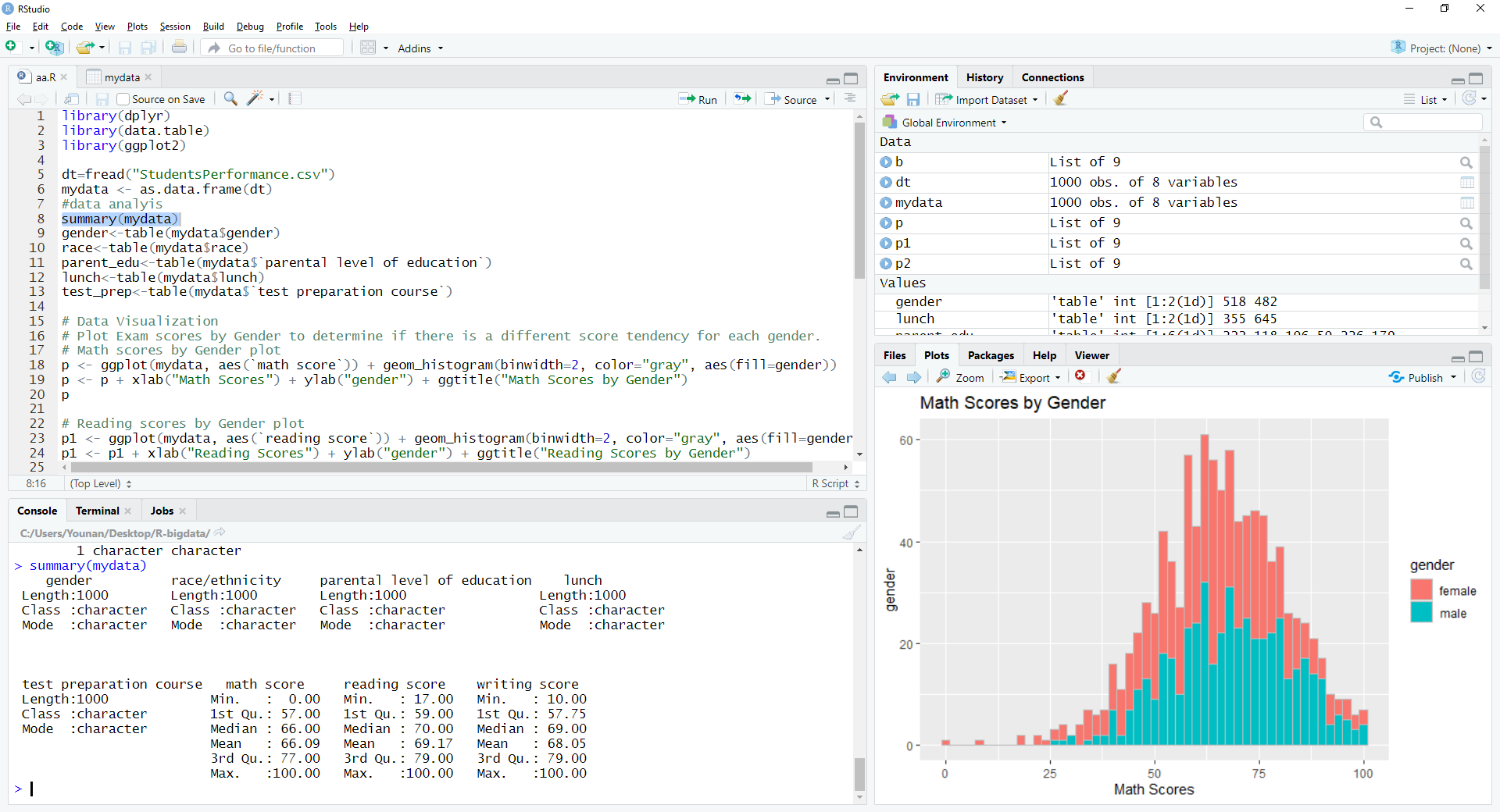
**BigData Final Project**

**Younan Nwesre 307980789 Tofeek Awidat 314731357**

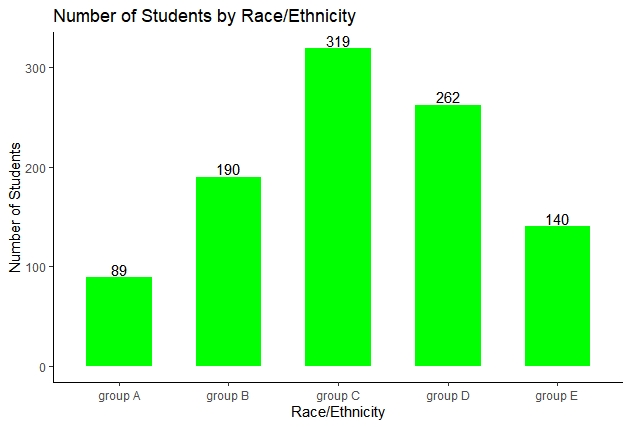
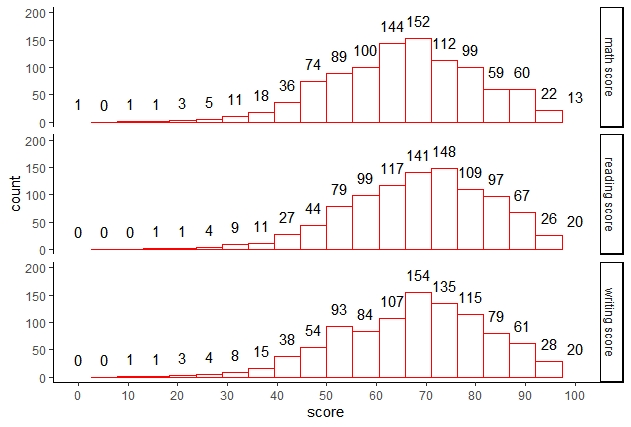
**Understanding the data :**

**The data has eight variables and 1000 observations. The key variables include gender, race/ethnicity, parental level of education ,lunch , test preparation course, math score, reading score and writing score.**



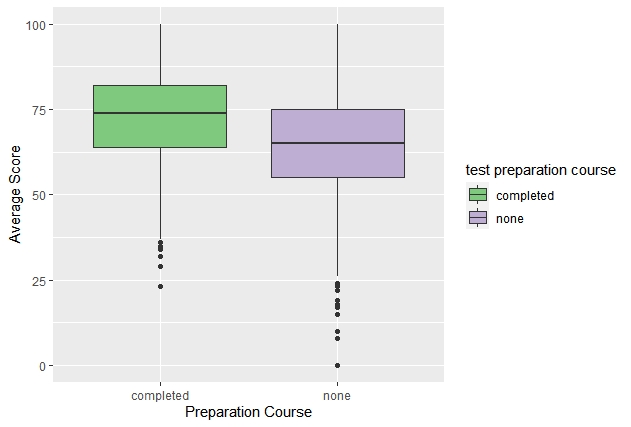
### Exploratory data analysis:

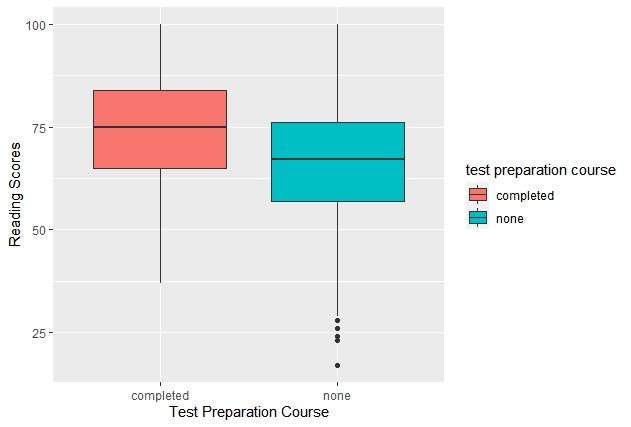
### Students By gender: There are more 518 female students and 482 male students.

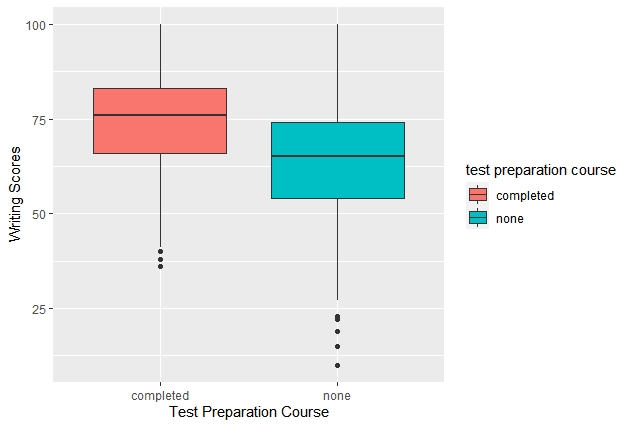
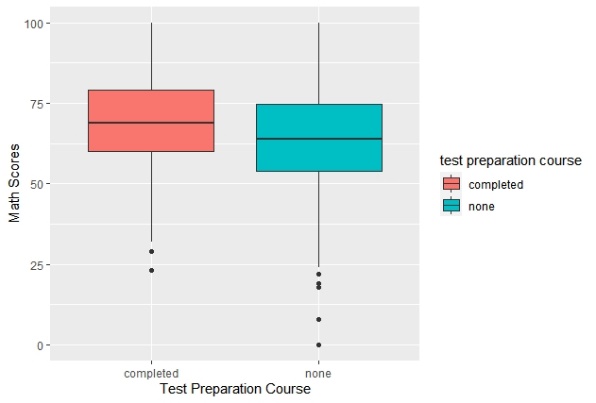
1. **Students By race**: There are 319 students in group C, 262 students in group D while there are only 89 students in group A.
2. **Distribution of scores**: 152 students scored between 65 - 70 marks in math. 154 students scored in the same range in writing. 148 students scored between 70 - 75 marks in reading.

### Preparation courses: we checked the preparation effect on scores 1-Completion: Majority of students do not complete preparation courses. Proportion of students completing preparation course is highest in group E.

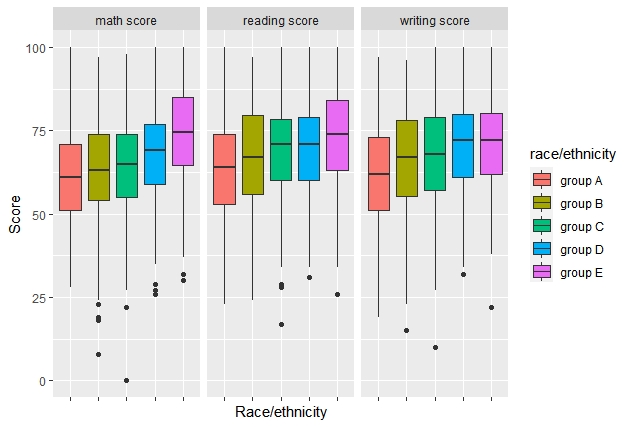
**2- Impact of course**: As we see the that the average of the scores of students who completed preparation course is higher than the average of scores of the students who did not complete the course.



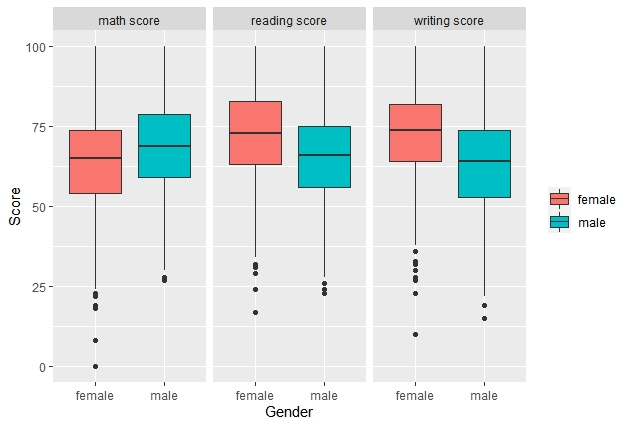
****

****

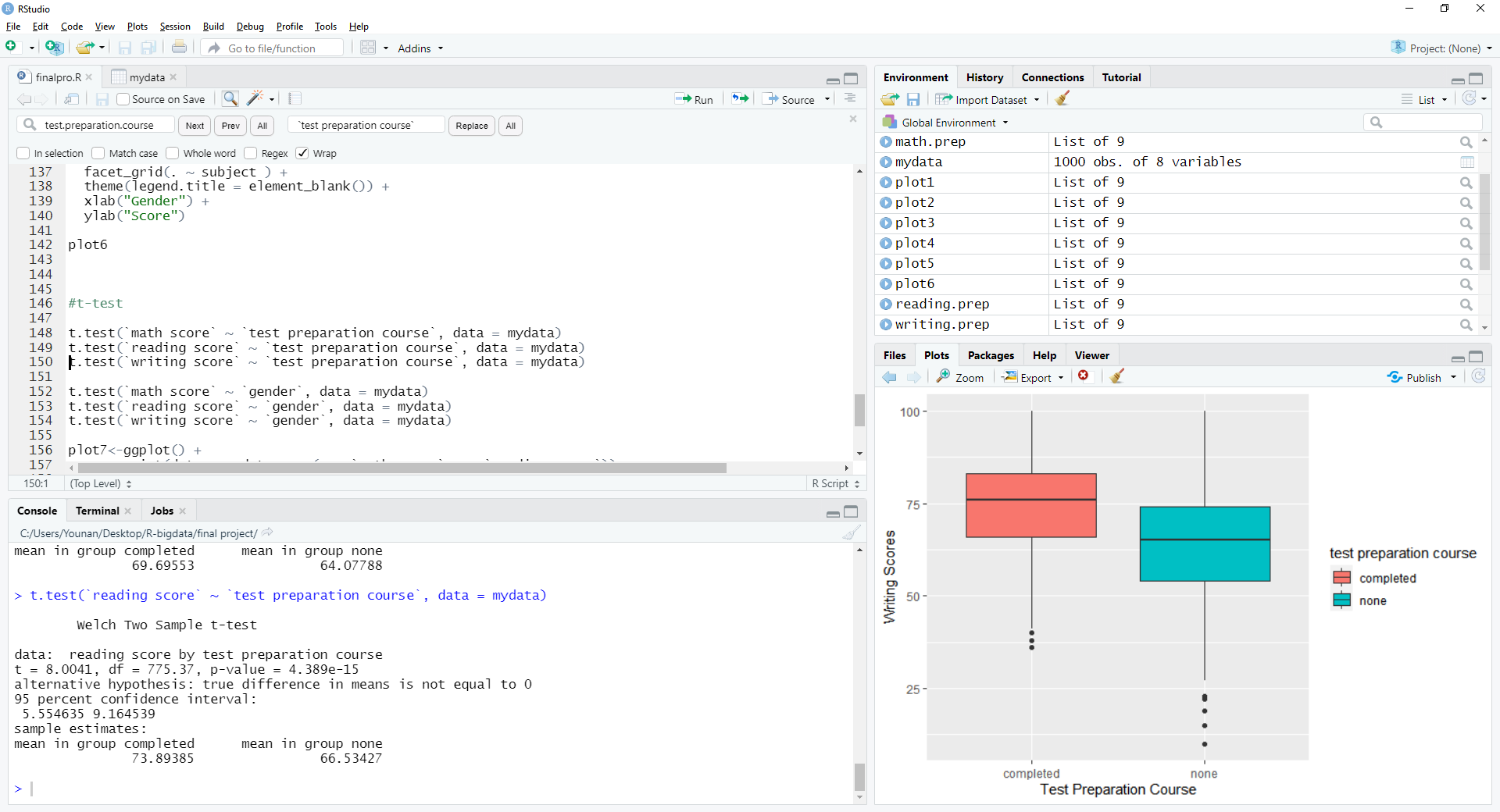
**Scores Average :**

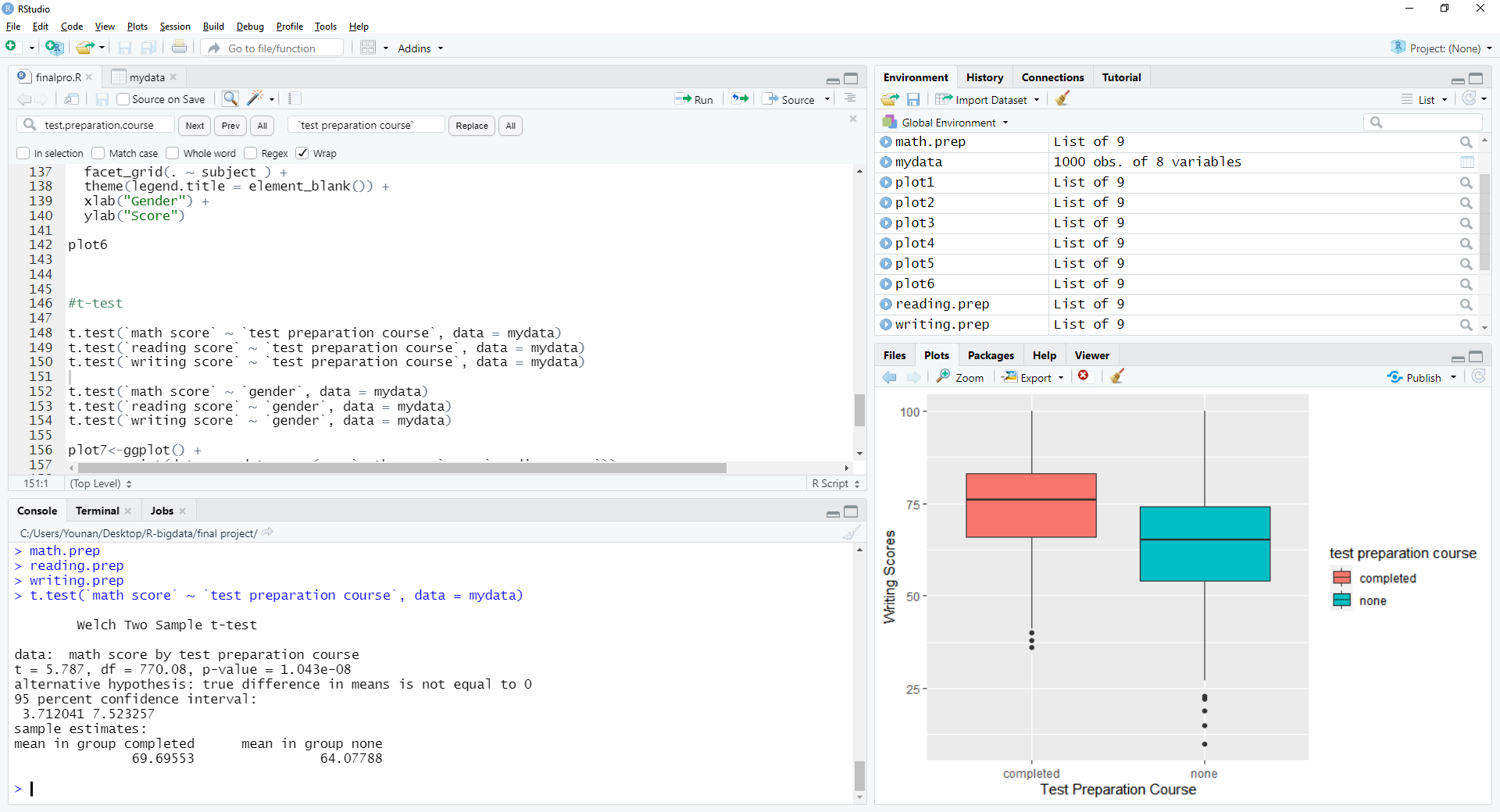
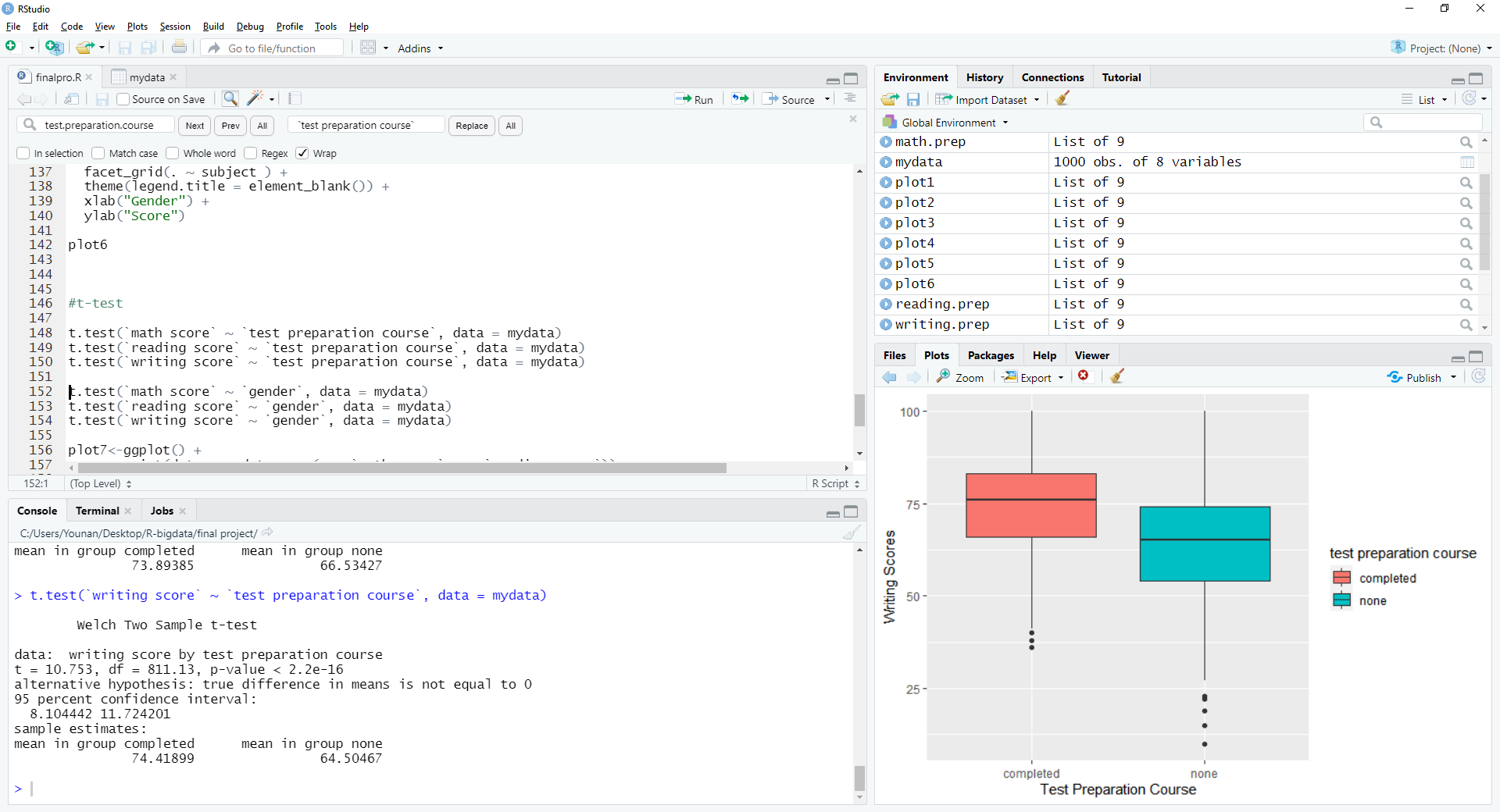
**1-Scores by Race:**

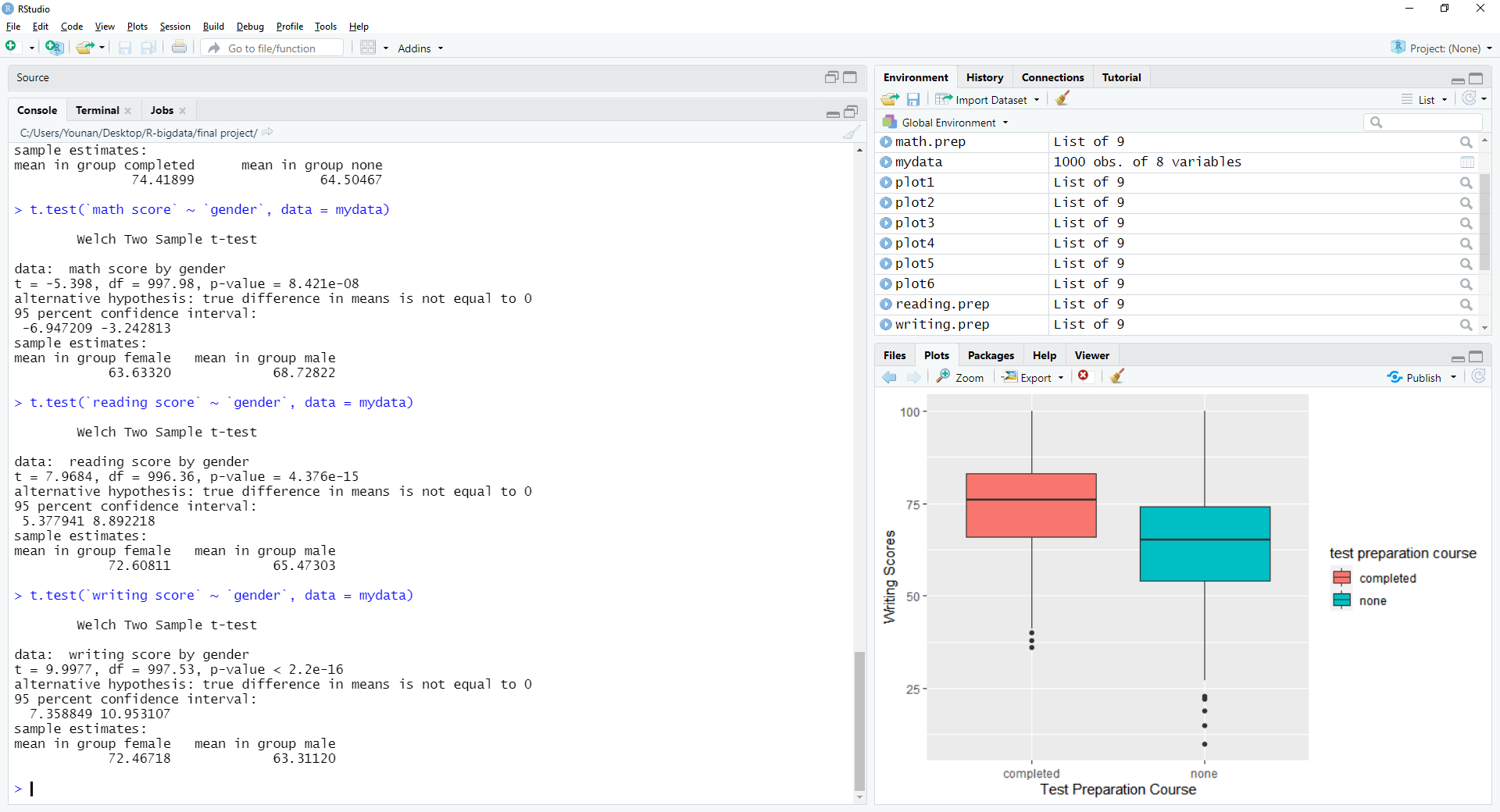
**2-Scores by Gender :**

****

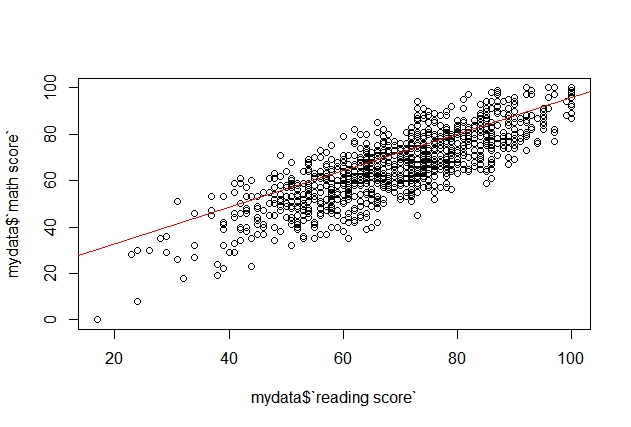
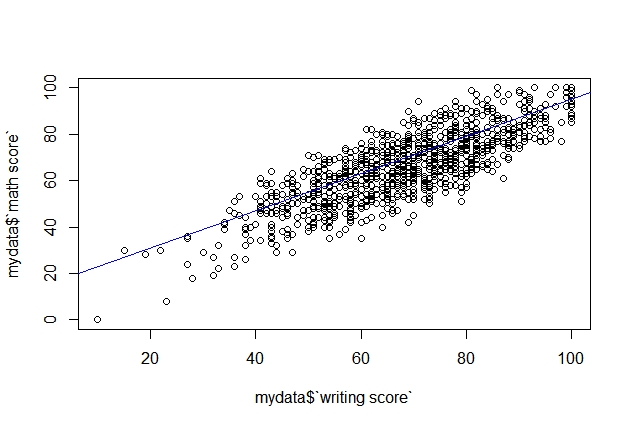
**T-test :**

We did a t-test between all the courses scores to courses preparation:



Also we did t-test between all the courses scores to students gender:

**Liner Model :**

****

**Conclusion :**

**As we see from our work on this dataset that students with preparation have a better chance to get higher scores than non-preparation students .**

**Also we can see that female get higher scores in reading and writing from males that gets a higher score in math .**

**If we check scores by race we can spot that race E have the higher scores in almost every subject.**