Stat 107 Individual Project Report

In this project, I choose the Students Performance in Exams dataset from Kaggle to analyze. This data set consists of the marks secured by the students in various subjects, which contains areas from the lunch standard of a student, test preparation for a course to the education level of a student’s parent, along with students’ math, reading and writing scores. Some people argue that if the educational level of a parent is higher, their children could have a better performance in tests because the high-educated parents would take tests more seriously than those low-educated parents. But others argue that the reason a student could performance better in a test may not be the educational level of the student’s parent, but the student has prepared well for the test. Therefore, in the analysis, I would like to find out that does parents’ educational level really affect the test performance of their children?

Table

Description automatically generated

The dataset has several columns: gender, race/ethnicity, parental level of education, lunch, test preparation course, math score, reading score, writing score. Except for the test scores, all other columns have string values. Since many columns in the dataset are not directly related to test performance such as gender, race and lunch and my analysis is to discover whether parental educational level can really make an impact on students’ performance in tests, I will focus on analyzing two columns of data, one is parental level of education and the other is test preparation course. I choose students that have completed their test preparation course, separating them into different groups based on parental education level. From these data, I then choose three groups to make comparison: some college group, bachelor group and high school group. Since the size of each group is greater than 30, I decide to use z-test for the next step of analysis.

Graphical user interface, application, Word

Description automatically generated

I make three comparisons between the groups I have chosen, high school group to bachelor group, high school group to some college group and some college group to bachelor group. And for each comparison, I conduct three z tests for math score, reading score and writing score respectively. For all the tests I have conducted, the Ho hypothesis is that there is no difference in the mean score of both groups and the Ha hypothesis is that there is difference in the mean score of both groups. The first value of the result is test statistic and the second value of the result is p-value. In addition, the α level in the z tests is 0.05. The first test is between high school group and bachelor group. The result is shown in the picture below:Graphical user interface, text, application, email

Description automatically generated

Since all the p-values in the tests are less than α level of 0.05, it is convincible that students whose parents have bachelor’s degree would perform better in the math, reading and writing tests than other students whose parents are high school education level.

Graphical user interface, text, application

Description automatically generated

The second comparison is between high school group and some college group. Also, since all the p-values in the tests are less than α level of 0.05, it is convincible that students whose parents have some college education level would perform better in the math, reading and writing tests than other students whose parents are high school education level.

Graphical user interface, text, application, email

Description automatically generated

The last comparison is between some college group and bachelor group. Based on the result from the other two comparisons, it is suggested that the bachelor group could perform better than some college group. But to my surprise, all the p-values in the tests are larger than α level of 0.05, and even larger than α level of 0.1. This result indicates there is no evidence that students whose parents have bachelor’s degree would perform better in the math, reading and writing tests than other students whose parents are some college education levels.

Based on the analysis above, the argument that if the educational level of a parent is higher, their children could have a better performance in tests does apply to the conditions between parents with high school level of education and parents with college/bachelor level of education. However, when both parents have some college education level or even higher, the parents’ educational level will not impact the test performance of their children in school, which indicates that the preparation of children to the tests could be a more important factor in determining the test performance than parent’s educational level at this time.