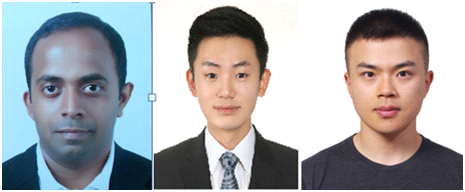
**Factors affecting on student's performance in the school**

**Gayan moraliyage**

**Yeonho Lee**

**Seonbin Lee**



**<ABSTRACT>**

Determinants of students' performance have been the interesting subject matter of ongoing debate among educators, academics, and policy makers. There have been many studies that focused on this issue and their findings revealed that parents’ education, family income and self motivation as factors that have a significant effect on the students performance in exams.

Many practical studies are carried out to investigate factors affecting college students’ performance. Student achievement is not simply a matter of what happens in school. Although schools can and do make a significant difference, research has identified numerous factors which affect student success.

The focus of this research is that student performance in examination is linked with students’ demographic, social and school related features. In this study, we use secondary data and the data had been collected by using school reports and questionnaire. We used both simple linear regression analysis and multiple linear regression analysis to look at the relationship between independent variables and students performance.

The study confirmed that parents education level and going out for fun are strong predictor students performance in Math while study does not provide any evidence that home to school travel time and number of school absence are critical factors for students performance in math. But study concluded that all these factors have impact on students performance in Portuguese language.

**<Introduction>**

There have been lots of studies conducted to figure out what can cause higher grade and what factors make student get lower grade. To analyze these, we get two dataset from two Portuguese high school which represents their math and Portuguese score according to 6 factors; sex, education level of parents, travel time to school, paid for additional education, the number of going out after school and the number of absence. As a result the paper sought to answer one main research questions. (1) What are the factors that affect students performance. To find out what factor can influence on one’s grade we set up 6 hypothesis.

**<Method>**

To conduct this study we first divided 6 factors into 2 types; Categorical, Numerical. Sex and additional paid belongs to categorical category and other 4 factors belong to numerical category. To analyze categorical value we used t-test and we used regression analysis to check other 4 factors; education level, commute time, the number of going out and the number of absence. Also we did multiple regressions after these analysis, since we assume that there would be some interaction between some factors.

Also to conduct this study we scale factors and below table is scale that we use.

∘To measure parents education level, we average two values; mother’s education level and father’s education level.

∘To measure one’s grade we average three values; first, second and final test.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Sex | Paid | Education level | Travel time | Going out | Absence |
| 0 | Women | No extra paid for additional Education | none | N/A | The number of going out after school per week (0~5) | The number of absence student make for one semester  (0~93) |
| 1 | Men | extra paid for additional Education | primary | less than 15 mins |
| 2 | N/A | N/A | 5th to 9th grade | 15 to 30 mins |
| 3 | N/A | N/A | secondary education | 30 mins to one hour |
| 4 | N/A | N/A | higher education | greater than one hour |
| 5 | N/A | N/A | N/A | N/A |

|  |  |  |  |
| --- | --- | --- | --- |
|  | First test | Second test | Final test |
| Grade | 0~20 | 0~20 | 0~20 |

**List of hypothesis**

**H1** : Gender is associated with students exam result.

**H2 :** Whether they paid for additional education is associated with students exam result.

**H3 :** Parents-education level is associated with students exam result.

**H4 :**Commute time to school is associated with students exam result.

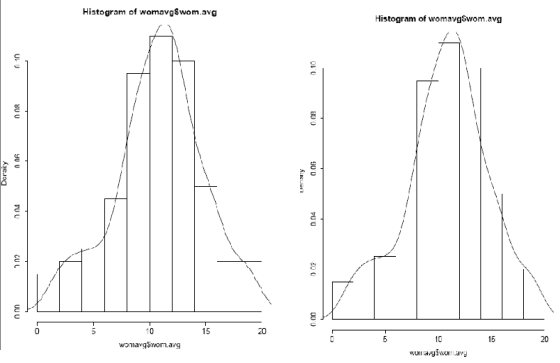
**H5 :** The number of going out is associated with students exam result.

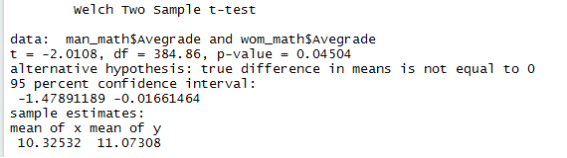
**H6 :** The number of absence in one semester is associated with students exam result.

**Results & Discussion**

**First**, we analyzed categorical variable ; sex and additional pay for study which are not possible to be analyzed by its numerical scale.

**1. Histogram of math score in male and female.**



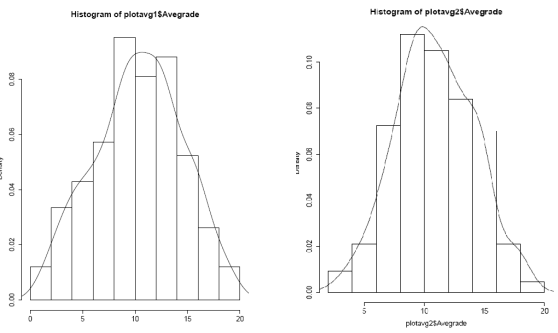


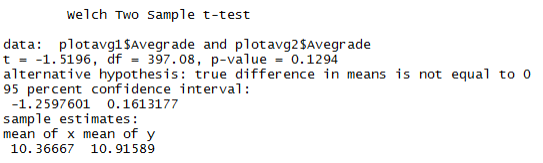
It seems data of math score is normally distributed, so we can compare the difference of math score by t-test.

From the t-test we can infer that there exists the difference in math grade between male and female.

-> Null hypothesis is rejected and H1can be accepted.

**2. Histogram of math score by additional pay for extra education.**

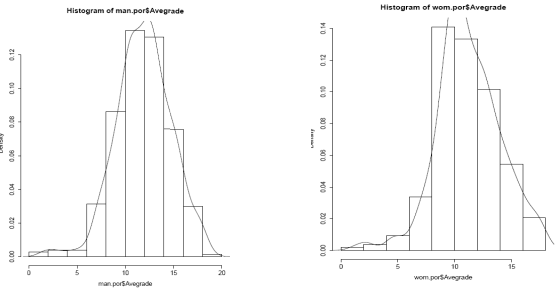


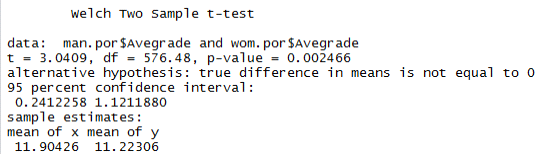
They seems normally distributed, so we can compare how different math the students have between expenditure for extra education by t-test.

Since the p-value is greater than .05, from the t-test we can infer that there is no difference between math score by additional pay for extra education. So, we can conclude that there is no relationship between additional pay and math score.

-> Null hypothesis is cannot be rejected.

**3. Histogram of Portuguese score in male and female.**





It seems that data of Portuguese score of male and female are normally distributed, so we can compare the difference of Portuguese score by t-test.

Since the p-value is less than .05, from the t-test we can infer that there is a difference between

Portuguese score and student gender.

-> Null hypothesis is rejected and H1can be accepted.

Then we analyze numerical variables.

**1. P-value, Coefficient, Correlation between factors and Math score**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***Math Score*** | **AveEdu** | **travel time** | **go out** | **absences** |
| **P-value** | 8.33E-0.6 | 0.01076 | 0.002073 | 0.9068 |
| **Coefficient** | 0.835 | -0.6794 | -0.5131 | -0.002729 |
| **Correlations** | 0.222128 | -0.1281972 | -0.1545113 | -0.0059088 |

**1) Parents education level and math grade**

Since P-value is less than 0.05 we can say that education level of parents is significant on student’s math grade. Also we can see correlation between two factors is 0.22. Since coefficient for the factor is 0.835 we can say that the more parents are educated, the math grade their children get is likely to increase.

-> Null hypothesis is rejected and H3can be accepted.

**2) Travel time to school and math grade**

Since P-value is less than 0.05 we can say that travel time to school is significant on student’s math grade. Also we can see correlation between two factors is –0.12. Since coefficient of for the factor is –0.6794 we can say that the more students use their time to go to school, the math grade they get is likely to decrease.

-> Null hypothesis is rejected and H4can be accepted.

**3) The number of going out and math grade**

Since P-value is less than 0.05 we can say that the number of going out is significant on student’s math grade. Also we can see correlation between two factors is –0.15. Since coefficient of for the factor is –0.5131 we can say that the more students are going out, the math grade they get is likely to decrease.

-> Null hypothesis is rejected and H5 can be accepted.

**4) The number of absences and math grade**

Since P-value is greater than 0.05, we can conclude that there is no relationship between the number of absences and math grade.

-> Null hypothesis cannot be rejected

**2. P-value, Coefficient, Correlation between factors and Portuguese score**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***P score*** | **AveEdu** | **traveltime** | **go out** | **absences** |
| **P-value** | 1.68E-12 | 0.0001119 | 0.03143 | 0.001417 |
| **Coefficient** | 0.761 | -0.5717 | -0.2035 | -0.07633 |
| **Correlations** | 0.272366 | -0.1510663 | -0.08446692 | -0.1250152 |

**1) Parents education level and Portuguese grade**

Since P-value is less than 0.05 we can say that education level of parents is significant on student’s Portuguese grade. Also we can see correlation between two factors is 0.27. Since coefficient for the factor is 0.761 we can say that the more parents are educated, the Portuguese grade their children get is likely to increase.

-> Null hypothesis is rejected and H3can be accepted.

**2) Travel time to school and Portuguese grade**

Since P-value is less than 0.05 we can say that travel time to school is significant on student’s Portuguese grade. Also we can see correlation between two factors is –0.15. Since coefficient for the factor is –0.5717 we can say that the more students use their time to go to school, the Portuguese grade they get is likely to decrease.

-> Null hypothesis is rejected and H4can be accepted.

**3) The number of going out and Portuguese grade**

Since P-value is less than 0.05 we can say that the number of going out is significant on student’s Portuguese grade. Also we can see correlation between two factors is –0.08. Since coefficient for the factor is –0.2 we can say that the more students are going out, the Portuguese grade they get is likely to decrease.

-> Null hypothesis is rejected and H5can be accepted.

**4) The number of absences and Portuguese grade**

Since P-value is less than 0.05 we can say that the number of absence is significant on student’s Portuguese grade. Also we can see correlation between two factors is –0.12. Since coefficient for the factor is –0.07 we can say that the more students are absent, the Portuguese grade they get is likely to decrease.

-> Null hypothesis is rejected and H6can be accepted.

Since we find the result that parent’s education level, travel time to school and the number of going out have relationship between both math grade and Portuguese grade, we can conclude there is relationship between those factors and student’s performance. However the results indicates that the number of absence only has relationship with Portuguese grade not math grade, we can’t conclude that there is relationship between student performance and the number of absence.

**Last** step is analyzing by using multiple regression.

Multiple Regression Analyze method was used to test the proposed hypotheses of this study. Multiple regression analysis is a statistical technique used to estimate the relationships among variables.

Results of the multiple regression analysis method as follows,

*Output 6 : Regression Analysis Results- Students Math score*

from the results of the multiple regression analysis, we can infer as following:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Hypothesis** | | **β** | **P-value** | **Results** |
| H3 | Parents edu. Level Students Math score | 0.82 | 1.42e-05 \*\*\* | Supported |
| H4 | Commute to school Students Math score | -0.44 | 0.089864 | Not supported |
| H5 | Going out for fun Students Math score | -0.54 | 0.000792 \*\*\* | Supported |
| H6 | Absences to school Students Math score | -0.006 | 0.760869 | Not supported |

The dependent variable for this test was student Math score and the independent variables were Parents education level, commute to school, going out for fun, Absences to school. Parents education level, Going out for fun were found to be significant enabler in students math score while Commute to school, Absences to school were not statistically significant under the 5% significant level. Parents education level has strong positive relationship (β = 0.82) while going out for fun has strong negative relationship (β = -0.54) on students math score

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Hypothesis** | | **β** | **P-value** | **Results** |
| H3 | Parents edu. Level Students Math score | 0.71 | 1.07e-10 \*\*\* | Supported |
| H4 | Commute to school Students Math score | -0.31 | 0.03600 \* | Supported |
| H5 | Going out for fun Students Math score | -0.18 | 0.04757 \* | Supported |
| H6 | Absences to school Students Math score | -0.07 | 0.00116 \*\* | Supported |

All the 4 hypothesis were found to be statistically significant on students Portuguese score. Parents education level has strong positive relationship to the Portuguese score while other 3 hypothesis have negative relationship.

According to the both simple regression and multiple regression analysis results, all the factors we considered showed same results except commute time and math performance.

Future directions

Until now, we analyzed how various factors can affect the exam score in Portugal. We only analyzed few factors in one country. In order to further generalization of our study, we need more datasets from various schools in various regions. Also if we can get data from different countries, it would be more helpful for generalizing of our findings. In this direction, we can study how various factor affect the accomplishment of students. Here we consider only few factors and probably there are more factors that may affect for student’s performance. Therefore we need to consider those factors in future studies. If we can use these findings for policy implication in education sector, that will be helpful for students to improve their ability.