ABDULMALIK MUJAHEED

ID: 0758045

DATA ANALYTICS WITH R

TERM PROJECT

TITLE:

SECONDARY SCHOOL STUDENT PERFORMANCES: AN ANALYSIS ON ITS RELATION TO FAMILY INVOLVEMENT

INTRODUCTION

Academic performance is an important aspect of a student's life, it is one of the ways to measure success and it can help children have better opportunities in the future. (Guo et al., 2023). There are many factors that determine or have a significant effect on academic success and failure such as; the involvement of family members and school teachers is a key factor in explaining children's success in academic, socio-emotional development, and socialization (Lawal., 2020). Previous researchers have shared their views on the relationship between these factors and academic performance with varied results.

1.1 SCHOOL & TEACHING METHODS

The school's involvement in student performance in terms of environment cannot be understated. In a study conducted by Araya and Dussaillant (2020), they found the secondary schools attended significantly improved students' performance in language and math. This was achieved using propensity score estimation methods to estimate the effect of attending a selective school on student performance then they conducted a changes-in-changes (**CIC**) analysis to estimate the effect of attending a secondary school under a specific program (Bicentenaro) on student performance over time and then estimates where repeated. This was supported by Ahmetović et al (2020) who suggest that schools can play a significant role in improving student performance through the creation of a conducive learning environment and collaborative atmosphere for optimal motivation. This study also suggests that teachers should focus on the necessities of independence, ability, and relatedness to help improve students' motivation. Idika, & Onuoha (2018) study highlights this importance as they find that students' poor academic performance has been linked to teachers' poor performance in terms of task completion, attitude toward work, and teaching practices, by conducting an independent sample student **t-test** and Analysis of Variance (**ANOVA**) they suggested that teachers with improved qualifications be hired and also improving incentives for existing professionals for increased morale.

1.2 FAMILY INVOLVEMENT

Family plays a crucial role in a child's development which can affect performance in school, a study by Baert & Van der Straeten (2021) mentioned that male students and students with low-educated fathers perform worse than females and students with highly educated fathers. A conditional **logit analysis** was used to estimate the effect of parental divorce on educational performance while controlling for child-fixed effects and time trends. Additionally, the study found that parental divorce has a negative effect on educational performance during secondary education and a high number of siblings negatively influences educational performance. This was supported by Adebusuyi (2018) whose study concluded a higher proportion of students who come from divorce homes have been affected negatively in academic performances using **correlation** as the key inferential analytical method. The study suggested the need for early solutions before the manifestation of the negative effects. These negative effects discussed were discussed in a study by Iqbal et al. (2021) as their findings from Friedman's **Two-way Analysis of Variance by Rank** showed that separated family's children have difficulty recalling the previous day's activities and also show a lack of confidence than the non-separated family's children in the school's activities. Another study by Sæther (2019) suggests that the deterioration of the family leading up to the divorce is more significant in effect on student achievement than the actual divorce itself. The author finds evidence using a regression analysis and concluded that student performance is negatively affected before the family breaks

up, and the effect is visible within a three-year window of the formally recorded divorce. Lastly, Bokhove & Hampden-Thompson (2022) suggest that the effects of family structure on student performance may depend on different things. For example, the type of school a student goes to and the country they live in may affect how much family structure matters. Also, the study found that schools with more single-parent families tended to have lower-performing students, but this was mostly because these schools also had more students from lower-income families.

Previous literature provided various insights as to why family situations are related to student academic performance but gaps in the research show that some of these individual relationships have different effects or no effect at all on student performance and also fail to highlight any moderating effect that could potentially affect student performance.

The overall goal of this study is to test the hypothesis that there is a relationship between Family involvement and the academic performance of secondary school students. The study aims to use an **ANOVA** to compare; the mean academic performance of students from different parental marital status groups and to determine if there is a significant association between parental occupation and academic performance. The study will also explore the potential moderating effects of other factors, such as the education of each parent and its relation to academic performance. This study also aims to improve on previous studies by also exploring the relation of parental occupations to student performances and through their effect sizes. A sample of secondary school student performances from Aman Chauhan (2022) retrieved from a public dataset (Kaggle) is used to achieve this claim.

METHOD

2.1 PARTICIPANTS

The participants used for analysis in this study were a sample of secondary school students from two Portuguese schools by Aman Chauhan (2022). The sample included students of both male and female gender, different parental jobs such as teacher, health services, at home and other occupations specified as others, different parental marital status groups, such as divorced and still-together, the student family size coerced into a group of two; less than or equal to 3 and greater than or equal to 3.

2.2 DATA COLLECTION & DESCRIPTION

The data used for this analysis were collected from Aman Chauhan (2022) through an open-source website (Kaggle). The original data collected was pre-processed and trimmed to only include relevant variables. It is assumed that permission from relevant educational authorities and schools was taken beforehand by the collector. Two Portuguese schools (Gabriel Pereira and Mousinho da Silveira) were used to collect this data and a diverse range of students of different genders, parental marital status, and parental occupations. The scores of the children across 3 terms were collected and used as the weightage for academic performance.

Table 1: Description of key columns used for analysis

Column Name	**Description**
Gender	Category: M (male) or F (female)
Family size	Category: 'LE3' - less or equal to 3 or 'GT3' - greater than 3)
Parent's Cohabitation Status	parent's cohabitation status (binary: 'T' - living together or 'A' - apart)

Father's Education	numeric: 0 (none), 1 (primary education - 4th grade), 2 (5th to 9th grade), 3 (secondary education) or 4 (higher education)
Mother's Education	numeric: 0 (none), 1 (primary education - 4th grade), 2 (5th to 9th grade), 3 (secondary education) or 4 (higher education)
Father Occupation	Category: 'teacher', 'health' care related, civil 'services' (e.g., administrative or police), 'at home' or 'other')
Mother Occupation	Category: 'teacher', 'health' care related, civil 'services' (e.g., administrative or police), 'at home' or 'other')
Family Educational Support	Category: yes or no
Quality of Family Relationships	Category: from 1 - very bad to 5 - excellent
Final Grade	Numeric: from 0 to 20

2.3 DATA ANALYSIS

To test the hypothesis that there is a relationship between family status and the academic performance of secondary school students, I conducted 4 different ANOVA on different variables and a post hoc analysis using a chi-square test. The data used for this analysis was a subset of the data collected from Aman Chauhan (2022) through Kaggle. The sub-set sample details are as shown in *table 1*. I organized the data into a contingency table that displays the frequencies of each combination of parental educational level and family support. We then conducted a chi-square test to determine if there is a significant association between family support (group: yes or no) and parental educational level (group: none, primary, grade school, secondary, higher). The p-value was calculated using static and post hoc library functions in R studio. Also, variable parental marital status, family support, and parental occupation were all compared individually with final grade scores using a one-way ANOVA. All statistical relationship tested in the hypothesis was done using a *0.05* significance level.

RESULT

Family influence on Student performance; A one-way ANOVA was used to compare the mean academic performance of students in two parental marital statuses (Living together or Living apart) in test 1, test 2 was conducted to compare the mean academic performance of student's parental occupation and test 3 was conducted to determine if there is a significant association between family support and parental educational level using a chi-square test and post hoc analysis.

TEST 1

The test sought to determine whether or not parental marital status had an effect on student final results. 397 students were randomly sampled from two Portuguese schools (355 parents were living together and 42 were living apart). The sample contained no extreme outliers. A **QQ-plot** of the residuals demonstrated normality by group, and a box plot demonstrated heterogeneity of variance. The mean final result for the Living Together group was ***10.30 (SD = 4.64)*** and the mean final result for the Living Apart group was ***11.07 (SD = 4.31)***. since one of the assumptions was violated thus, a non-parametric test will be conducted. A permutated one-way independent ANOVA showed that the mean difference in final result scores across

parental marital status was statistically not significant, ***F(1,395) = 1.07, p > 0.05***. Hence, there is no evidence to show parental marital status had any effect student's final result.

TEST 2

1. **MOTHER**; The test sought to determine whether or not parental occupation status had an effect on student final results. 397 students were randomly sampled from two Portuguese schools with 5 occupational splits for their mothers (60 at home, 34 for health, 142 for others, 103 for services, and 58 for teachers). The sample contained no extreme outliers. A QQ-plot of the residuals demonstrated normality by group, and a box plot demonstrated homogeneity of variance. The mean final result for at home group was ***9.10 (SD = 4.71),*** health group was ***12.15 (SD = 4.23),*** others group was ***9.75 (SD = 4.42),*** services group was ***11.02 (SD = 4.76)*** and the mean final result for the teacher group was ***11.05 (SD = 4.40)***. A one-way independent ANOVA showed that the mean difference in final result scores across parental marital status was statistically significant, ***F(4,392) = 3.99, p = 0.003, η 2 = 0.04***. A TukeyHSD pairwise comparison showed that the health group had significantly higher children's final scores than both the at-home and other groups of mothers' children and there were no significant proportional differences amongst other groups.
2. **FATHER**; The test sought to determine whether or not parental occupation status had an effect on student final results. 397 students were randomly sampled from two Portuguese schools with 5 occupational splits for their fathers (20 at home, 18 for health, 214 for others, 112 for services, and 30 for teachers). The sample contained no extreme outliers. A QQ-plot of the residuals demonstrated normality by group, and a box plot demonstrated heterogeneity of variance. The mean final result for at home group was ***10.15 (SD = 5.32),*** health group was ***11.61 (SD = 3.24),*** others group was ***10.19 (SD = 4.51),*** services group was ***10.20 (SD = 4.56)*** and the mean final result for the teacher group was ***11.77 (SD = 5.50)***. A permutated one-way independent ANOVA showed that the mean difference in final result scores across parental marital status was statistically significant, ***F(4,392) = 1.15, p > 0.05***.

Hence, the analysis shows that mothers have more impact or effect on their children's academic performance than fathers. Also, children with mothers who work in health have significantly higher scores than children with mothers who stay at home.

TEST 3

1. **MOTHER**; This test research sought to determine whether parental support with school work is associated with the level of education their mother has. 396 mothers (2 from the none group, 60 from the primary school group, 103 from the grade-school group, 99 from the secondary school group, and 132 from the higher group) in a randomized sample reported if they support their children with school work ***(N = 242)*** or if they do not ***(N = 154),*** the sample failed the assumption of greater than or equal to 5 expected values. A simulated Chi-square Test of Independence parental support with school work and the level of education mothers were significantly related, ***χ 2 (4, N = 396) = NA, p = 0.008, V= 0.15***. According to Cohen's (1988) conventions, this effect was small. A post hoc test for proportions revealed no proportional differences between groups.
2. **FATHER**; This test research sought to determine whether parental support with school work are associated with the level of education their father has. 397 fathers (2 from the

none group, 83 from the primary school group, 115 from the grade-school group, 100 from the secondary school group and 97 from the higher group) in a randomized sample reported if they support their children with school work ***(N = 242)*** or if they do not ***(N = 155),*** the sample failed the assumption of greater than or equal to 5 expected values. A simulated Chi-square Test of Independence parental support with school work and the level of education fathers were significantly related, ***χ 2 (4, N = 397) = NA, p = 0.006, V= 0.16***. According to Cohen's (1988) conventions, this effect was small. A post hoc test for proportions revealed that fathers with primary education were significantly less likely to support their children with schoolwork, while fathers with higher education were significantly more likely to support their children with schoolwork, this is displayed graphically in *figure 1* as the proportion of expected (yes or no) and observed (yes or no) values are displayed per group.

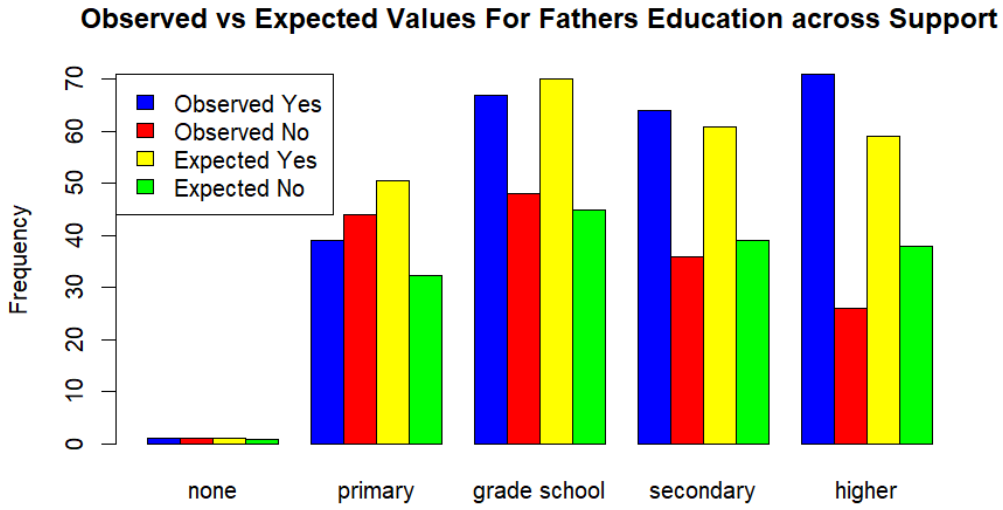

Figure 1: Bar Plot of The Expected and Observed Values from The Result of Test 2

Hence, this analysis shows that the level of support for helping these students at home depends on the educational level of both mothers and fathers.

Note: All write-up formats were retrieved from Henning (2023).

TEST 4

The test sought to determine whether or not family size had an effect on student final results. 397 students were randomly sampled from two Portuguese schools with 2 family size splits (283 families with more than 3 family size (**GT3**) and 114 families with less than or equal to 3 family size (**LE3**)). The sample contained no extreme outliers. A QQ-plot of the residuals demonstrated normality by group, and a box plot demonstrated homogeneity of variance. The mean final result for at GT3 group was ***10.13 (SD = 4.73)*** and the mean final result for the LE3 group was ***11.00 (SD = 4.22)***. A one-way independent ANOVA showed that the mean difference in final result scores across parental marital status was not statistically significant, ***F(1,395) = 2.93, p = 0.08***.

DISCUSSION

Academic performance is a crucial aspect of a student's life, and family involvement and school teachers are key factors in promoting academic achievement and socio-emotional development. However, previous research has produced varied results on the relationship between these factors and academic performance, underscoring the need for further investigation.

The analysis conducted in this study aimed to investigate the relationship between family involvement and secondary school student performance. The data used for this analysis were collected from two Portuguese schools, Gabriel Pereira and Mousinho da Silveira, and included a diverse range of students of different genders, parental marital status, and parental occupations. The scores of the children across three terms were collected and used as the weightage for academic performance.

The results showed that not all family involvement has a significant impact on a student's academic success. A one-way ANOVA showed that there isn’t any influence of parental marital status and family size on student academic performance. On the other hand, the same statistical inference showed that the occupation of their mothers has an effect on student performance and fathers do not, further tests for effect sizes (TukeyHSD) showed that children with mothers who work under health had significantly better results than stay-at-home moms. A chi-square test of independence showed that the level of support children get with school work from family at home is dependent on the educational level of both mothers and fathers. The analysis also showed that fathers who have only primary school education are significantly less likely to help and those with higher education are significantly more likely to help with their offspring’s school work.

In conclusion, the analysis conducted in this study provides valuable insights into the family involvement effects on secondary school student performances through parental occupation and parental educational level. The results suggest that the mother’s occupation is an important factor that can positively impact a student's academic success and the level of education obtained by both parents will determine if they will help their children with school work. Therefore, it is recommended that schools and families work together to promote family involvement in students' education. This can be done by possibly collecting parental information and paying more attention to students who fall under the subgroups of parents like those without any level of education in case there is a need for extra help academically.

This research had limitations, as this study was conducted using the student's results in math the main quantifier for academic performance thus, not giving a full description of their total academic grade point average as they are other subjects. Also, the data used in this analysis was collected from only two Portuguese schools, which may limit the generalizability of the findings to other schools and contexts. Additionally, the data was collected using self-reported measures, which may be subject to bias and may not accurately reflect the true level of family involvement or academic performance.

For future research it is needed to address these limitations, to provide a more complete understanding of the factors that contribute to student success, and also explore other factors

within family involvement like such as socioeconomic status, cultural background, and individual differences in learning styles and abilities.

REFERENCES

1. Guo, Z., Qi, C., Yang, J., Xu, Y., & Li, S. (2023, March 31). How family structure influences middle-school students' involvement in physical exercise and their academic achievement in China. Humanities and Social Sciences Communications, 10(1). https://doi.org/10.1057/s41599-023-01636-8
2. Aman Chauhan (2022, October). *Student Performance* [Data set]. Retrieved [2023, July 17] from https://www.kaggle.com/datasets/whenamancodes/student-performance
3. Lawal. (2020). Age, home and school involvements, and home-school communication in alcohol use among secondary school students. Vulnerable Children and Youth Studies, 15(3), 236–245. https://doi.org/10.1080/17450128.2019.1695993
4. Araya, & Dussaillant, F. (2020). Does attending a selective secondary school improve student performance? Evidence from the Bicentenario schools in Chile. *School Effectiveness and School Improvement*, *31*(3), 426–444. https://doi.org/10.1080/09243453.2019.1697299
5. Ahmetović, Bećirović, S., & Dubravac, V. (2020). Motivation, anxiety and students' performance. European Journal of Contemporary Education, 9(2), 271–289. https://doi.org/10.13187/ejced.2020.2.271
6. Idika, & Onuoha, J. C. (2018). Influence of economics teachers' personality on secondary school students' classroom performance in public secondary schools in Nsukka local government area of Enugu state. Journal of Social Science Education, 17(3), 100–106. https://doi.org/10.4119/UNIBI/jsse-v17-i3-1667
7. Baert, & Van der Straeten, G. (2021). Secondary School Success in Times of Parental Divorce. Family Relations, 70(2), 575–586. https://doi.org/10.1111/fare.12476
8. Adebusuyi. (2018). Effects of divorce and separation on academic performance of vulnerable children in Nigeria : social work as a panacea. Gender & Behaviour, 16(3), 11894–11907.
9. Iqbal, Khalid, M. S., Rehman, N., & Yanping, L. (2021). Parental Divorce: Impact on Socio-Psychological Behavior and Academic Performance of Students in Teacher's Perception. Journal of Divorce & Remarriage, 62(6), 475–492. https://doi.org/10.1080/10502556.2021.1925854
10. Sæther. (2019). Childhood Family Dissolution and School Outcomes. The Timing of Dissolution Effects. Marriage & Family Review, 55(7), 686–700. https://doi.org/10.1080/01494929.2019.1592057
11. Bokhove, & Hampden-Thompson, G. (2022). Country and school family composition's effects on mathematics achievement. School Effectiveness and School Improvement, 33(2), 280–302. https://doi.org/10.1080/09243453.2021.2012207
12. Henning, C. (2023). *AMOD-5210 Workshop 1: Introduction to Basic and Inferential Statistics in R.* Trent Blackboard. Retrieved July 10, 2023, from https://trentu.blackboard.com/webapps/blackboard/execute/content/file?cmd=view&content_id=_1551876_1&course_id=_50501_1
13. Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Hillsdale, NJ: Lawrence Erlbaum.