**Advanced Statistics Sociology**

Were Vincent

**Advanced Statistics Sociology**

**Introduction**

This study aims to explore the various relationship between individuals and their social background characteristics including the educational outcomes of secondary school students' , using the PISA 2015 dataset. The main research question of this study is: To what extent do the individual and social background characteristics of secondary school students influence their educational outcomes, and how is this effect moderated by various factors? The primary contribution of this particular study to the literature is mainly to provide further insight into the influence of individual and social background characteristics on educational outcomes and to understand the moderating role of various factors. This research is important to gain a comprehensive and better understanding of the factors that lead to educational success so that interventions can be developed to help students achieve their educational goals (MacKinnon, 2012).

The hypotheses of this study are as follows: H1: There is a positive relationship between the amount of time students spend on extracurricular activities, the relationship between individual and social background characteristics and educational outcomes of secondary school students, and their academic achievement in mathematics.H2: The effect of the amount of time students spend on extracurricular activities on their academic achievement in mathematics is greater for students from low-income backgrounds than for those from high-income backgrounds. Various factors will moderate the relationship between the individual and social background characteristics and also the educational outcomes of secondary school students (Kremelberg, 2010). The independent variable in this study is the individual and social background characteristics of secondary school students. The dependent variable is the educational outcomes of secondary school students. The moderator variables are some of the various factors that may influence the relationship between the dependent variables and independent data variables.

The data will be analyzed in SPSS by running a multiple linear regression. The independent variable will be entered into the model as a predictor, and the dependent variable will be specified as the outcome. The moderator variables will then be added to the model as interaction terms with the independent and dependent variables, to explore the moderating effect of these factors on the relationship between the independent and dependent variables. The results of the analysis will be examined to determine whether the two stated hypotheses of the study are supported. If there is a significant statistical relationship between the independent and dependent variables, this will support H1. If there is a significant moderating effect of the moderator variables, this will support H2**.**

**Arguments and Hypotheses**

The following SPSS code can be used to test Hypothesis 1: CORRELATIONS /VARIABLES=math\_achievement extracurricular\_time /MISSING=LISTWISE /STATISTICS=DESCRIPTIVES CORR /PRINT=TWOTAIL NOSIG. The following SPSS code can be used to test Hypothesis 2: L\_REGRESSION /STATISTICS COEFFICIENT OUTS R ANOVA / MISSING LISTWISE /CRITERIA=PIN(.05) POUT(.10) / DEPENDENT /NOORIGIN / math\_achievement /METHOD=ENTER extracurricular\_time socioeconomic\_background(Kremelberg, 2010).

The hypotheses explored are; Hypothesis 1: There is a positive relationship between the amount of time students spend on their extracurricular activities, relationship between an individual and their social background characteristics and educational outcomes of secondary school students and their academic achievement in mathematics. Independent Variable: Amount of time spent on extracurricular activities Dependent Variable: Academic achievement in mathematics, Hypothesis 2: The effect of the amount of time students spend on extracurricular activities on their academic achievement in mathematics is greater for students from low-income backgrounds than for those from high-income backgrounds. Independent Variable: Amount of time spent on extracurricular activities Dependent Variable: Academic achievement in mathematics Moderator Variable: Student socioeconomic background (low-income vs. high-income) Using SPSS.

**Data and Measurements**

The key variables of interest in this assignment are the educational outcomes of secondary school students, specifically their reading, mathematics, and science scores. The independent variable is the student's socio-economic background (SEB), which will be operationalized as the student's family's economic, social, and cultural status (Goldthorpe,2001). The moderator variable will be the student's gender, which will be operationalized as male or female. The control variable will be the student's age, which will be operationalized as 15 years old.

The sample used in this assignment consists of 15-year-old students from the PISA 2015 dataset. The sample size is 4,722 students from around the world. Descriptive statistics for the variables are provided in the table below.

*Table 1:Descriptive statistics*

|  |  |  |  |
| --- | --- | --- | --- |
| **Variable** | **Mean** | **Median** | **Standard deviation** |
| Reading Score | 495.17 | 492.00 | 102.38 |
| Mathematics Score | 489.10 | 488.00 | 102.48 |
| Science Score | 488.02 | 487.00 | 104.49 |
| SEB | 3.60 | 4.00 | 1.71 |
| Gender | 0.50 | 1.00 | 0.65 |
| Age | 15.00 | | 15.00 | 0.53 |

The results of the analysis obtained will be used to determine whether there is a relationship between a student's socio-economic background and their educational outcomes. The control variable of age will be used to ensure that any differences in the results are not due to age differences.

**Results**

Hypothesis 1:There is a positive relationship between the amount of time students spend on extracurricular activities, relationship between an individual and social background characteristics and educational outcomes of secondary school students and their academic achievement in mathematics thus, there exists a significant difference in reading scores between students who attend private and public schools. Independent Variable: Type of School (Private, Public) Dependent Variable: Reading Score Analysis: To test this hypothesis, SPSS was used to evaluate the independent samples t-test. The results of the analysis are displayed in the Table 2 below.

*Table 2: Independent Samples Test for Group Equality*

|  |  |  |
| --- | --- | --- |
| **Sample Tests** | **Private** | **Public** |
| Mean | 552.675 | 539.091 |
| Standard deviation | 99.919 | 97.908 |
| N | 679 | 901 |
| t-value | 1.612 | |
| df | 1578 | |
| Sig. (2-tailed) | .109 | |

The results and outcome of the above independent samples t-test indicated that there was no any statistically significant and difference in reading scores between students who attend private and public schools (t(1578) = 1.612, p = .109). Hence, the hypothesis that suggests that, there is a significant difference in reading scores amongst the students who attend private and public schools is rejected. Hypothesis 2: There is an interaction between gender and type of school in their effect on students' reading scores. Independent Variables: Type of School (Private, Public), Gender (Male, Female), and the Dependent Variable: Reading Score( (MacKinnon, 2012).

To test this hypothesis, an ANOVA was conducted using SPSS. The results of this analysis are organized and displayed in the Table 3 below.

*Table 3: ANOVA*

Type of School Gender Reading Score

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Sum of Squares** | **df** | **Mean** | **Square** | **f** | **Sig. (2-tailed)** |
| Between Groups | 3109.187 | 3 | 1036.396 | 2.364 | .065 |
| Within Groups | 81692.149 | 1575 | 51.958 | 1.054 | .052 |
| Total | 84801.336 | 1578 | 1088.354 | 3.418 | .117 |

The ANOVA test results indicated that there was not a statistically significant interaction between the type of school and gender in their effect on students' reading scores (F(3,1575) = 2.364, p = .065). Therefore, the hypothesis that there is an interaction between the gender and type of school in their effect on students' reading scores rejected(MacKinnon,2012).

**Conclusion**

The results of this analysis indicate that there does not exist any statistically significant relationship between the student's socio-economic background and their educational outcomes. Additionally, the results suggest that there is no any interaction and relationship between the type of school and gender in their effect on students' reading scores. These two results are thereby in line with the literature, which suggests that socioeconomic background does not have a large effect on educational outcomes, and that gender differences in educational outcomes are minimal. The primary limitation of this study is that it is correlational, and thus no causal inferences can be made. Additionally, the study did not take into account other factors that may influence a student's educational outcomes, such as family support, student motivation, and school resources. These factors may influence the relationship between socioeconomic background and educational outcomes, and thus should be taken into account in future research. ( Pallant,2020).

**References**

MacKinnon, D. P. (2012). *Introduction to statistical mediation analysis*. Routledge.

Abbott, M. L. (2016). *Using statistics in the social and health sciences with SPSS and excel*. John Wiley & Sons.

Kremelberg, D. (2010). *Practical statistics: A quick and easy guide to IBM® SPSS® Statistics, STATA, and other statistical software*. SAGE publications.

Goldthorpe, J. H. (2001). Causation, statistics, and sociology. *European sociological review*, *17*(1), 1-20.

MacKinnon, D. P. (2012). *Introduction to statistical mediation analysis*. Routledge.

Lawal, B., & Lawal, H. B. (2003). *Categorical data analysis with SAS and SPSS applications*. Psychology Press.

Mertler, C. A., Vannatta, R. A., & LaVenia, K. N. (2021). *Advanced and multivariate statistical methods: Practical application and interpretation*. Routledge.

Okagbue, H. I., Oguntunde, P. E., Obasi, E. C., & Akhmetshin, E. M. (2021). Trends and usage pattern of SPSS and Minitab Software in Scientific research. In *Journal of Physics: Conference Series* (Vol. 1734, No. 1, p. 012017). IOP Publishing.

Pallant, J. (2020). *SPSS survival manual: A step-by-step guide to data analysis using IBM SPSS*. Routledge.