Name: Qi Li
Student ID: 1005299302
Course: INF2178-Experimental Design for Data Science
Instructor: Shion Guha
Content: Technical Assignment III

Introduction

Math is an important skill for students. In this report, I hope to explore the impact of households' income levels on students' math scores and determine the moderating effect of general knowledge scores.

For simplicity, only the fall data is employed. The variables involved here are 'fallmathscore' (FMS), 'incomegroup' (IG) and 'fallgeneralknowledgescore' (FGKS). The fall general knowledge score is considered a covariate in this study and serves as a baseline. It means the initial level of the dependent variable for each group before treatment (i.e., income effect) is introduced. Additionally, the report will later include 'fallreadingscore' (FRS) as the second covariate since many believe that better reading abilities may contribute to better math performance.

I aim to address three core research questions that serve as guiding principles in curriculum design and can be extended to other subjects of study for enhancing educational outcomes across various academic disciplines:

1. Are there statistically significant differences in math scores among different income groups when adjusting for general knowledge scores, and if so, in which direction?
2. Does the interaction between explanatory variables have a significant impact on students' math scores beyond the main effects of these variables?
3. How does the inclusion of reading scores as an additional covariate in an ANCOVA model affect the explained variance in students' math scores?

Exploratory Data Analysis

First, I used Python to do some basic exploratory data analysis on the dependent variable, FMS, among different income groups.

Income Group	# of Observations	Mean	Standard Deviation
1	4729	23.924504	7.642412
2	3726	27.568468	8.541161
3	3478	31.012720	9.930977

Table 1. Summary statistics of fall math scores by income group

Analysis of Table 1 reveals a clear trend: as income group levels rise, the mean math scores, as well as the standard deviation, increase correspondingly. This suggests that students in higher income groups perform relatively better in math but with greater variation. This trend is visually represented in Fig 1 (left). A similar observation can be seen for the FGKS in Fig 1 (right).

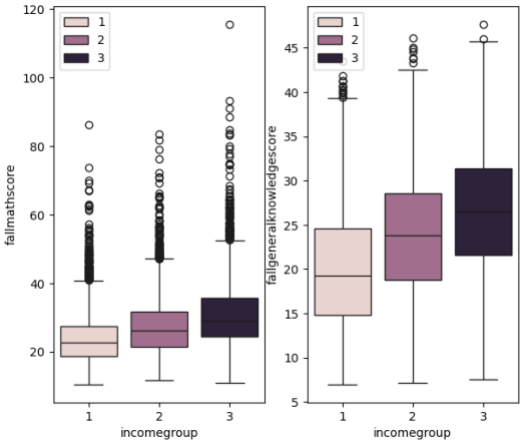

Fig 1. Box plots of FMS and FGKS among different income groups

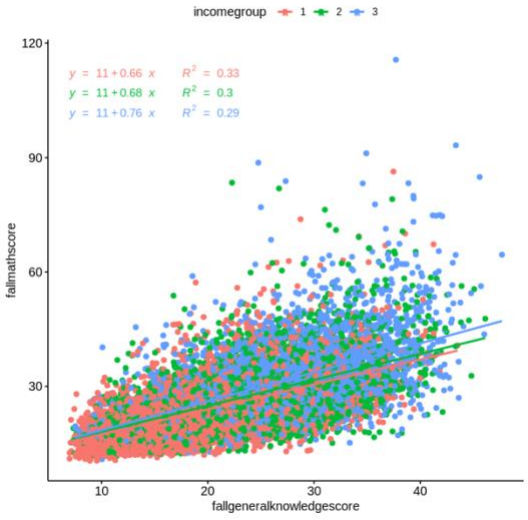

Fig 2. Scatter plot of FGKS against FMS for different income groups

The scatter plot in Fig 3 depicts a positive linear relationship between the covariate (FGKS) and dependent variable (FMS) across all income groups. However, minor differences are observed between income groups as indicated by the varying distances between the regression lines. Our study aims to determine whether such differences are significant.

ANOVA

Before proceeding with ANCOVA, I perform an ANOVA with ‘fallmathscore ~ incomegroup*fallgeneralknowledgescore’ in hopes of getting an initial understanding of the relationship between variables, which can inform the subsequent ANCOVA. The results suggest the effect of the general knowledge score on math score is statistically significant with a p-value < 0.001. Similarly, both the effect of income group and the interaction effect are statistically significant with p-values below the conventional significance level of 0.05. However, the

practical impact might be limited for the latter two effects suggested by the small generalized eta squared (ges) values.

Effect	DFn	DFd	F-statistic	P-value	P<.05	ges
IG	2	11927	112.106	5.82e-49	*	0.018
FGKS	1	11927	5291.155	0.00e+00	*	0.307
IG*FGKS	2	11927	10.487	2.82e-05	*	0.002

Table 2. ANOVA table

ANCOVA

The first ANCOVA analysis was performed to evaluate the impact of IG on FMS while controlling for FGKS as a covariate.

Source	Sum of Squares	Degrees of Freedom	F-statistic	P-value Uncorrected	Partial Eta Squared
IG	11585.864980	2	111.928370	6.927877e-49	0.018420
FGKS	273412.600276	1	5282.752173	0.000000e+00	0.306927
Residual	617393.889005	11929	NaN	NaN	NaN

Table 3. ANCOVA table with 'fallmathscore ~ incomegroup + fallgeneralknowledgescore'

The resulting p-values, far below 0.05, indicate that both IG and FGKS have significant impacts on FMS at the 0.05 significance level. Since ANCOVA is a combination of linear regression and ANOVA, I also used Ordinary Least Squares (OLS) regression to perform the same ANCOVA. In the more comprehensive table obtained from OLS regression, the adjusted R-squared equals 0.377, indicating a relatively small proportion of variation explained by the model. To improve the model's explanatory power, I am going to incorporate the second covariate, FRS.

Term	Coefficient	Standard Error	T-statistic	P > |t|
IG	0.0831	0.070	8.235	0.000
FRS	0.4630	0.006	78.546	0.000
FGKS	0.4030	0.009	46.801	0.000

Table 4. (Partial) ANCOVA table with 'fallmathscore ~ incomegroup + fallreadingscore + fallgeneralknowledgescore'

The inclusion of FRS as an additional covariate substantially improves the model's explanatory power, as evidenced by an increase in the adjusted R-squared from 0.377 to 0.590. Moreover, the coefficient for FRS is significant with a p-value < 0.001. It means that reading scores are a statistically significant predictor of math scores at the significance level of 0.05. This finding supports the hypothesis that reading ability may contribute to math performance, which reflects the importance of reading in learning and understanding math-related concepts.

Term	Coefficient	Standard Error	T-statistic	P > |t|

IG	0.0461	0.223	0.207	0.836
FGKS	0.3593	0.019	18.556	0.000
IG*FGKS	0.0229	0.009	2.519	0.012
FRS	0.4624	0.006	78.371	0.000

Table 5. (Partial) ANCOVA table with 'fallmathscore ~ incomegroup* fallgeneralknowledgescore + fallreadingscore '

Given the significant main effects of both IG and FGKS on FMS, in addition to the findings from ANOVA, there is a theoretical justification suggesting that the effect of IG on the dependent variable could depend on the level of FGKS. I will further extend the ANCOVA model by including interactivity.

In Table 5, it's noteworthy that the main effect of IG becomes insignificant at the 0.05 significance level, despite being significant in the additive model. While the interaction term itself is significant.

Explanation

This indicates that the effect of IG on FMS depends on the level of FGKS. Income group alone doesn't tell the whole story of its effect on the math scores. For example, in some income groups, higher general knowledge is associated with better math performance, while the impact is less pronounced in other groups. In general, the effect of being in certain income groups on math scores varies at different levels of general knowledge.

This finding aligns with the result from ANOVA, where the generalized eta squared (ges) value for the IG term is very small, indicating the variation in math scores explained by the income group alone is minimal. It highlights the need to look at the implications derived from combining related variables rather than testing predictors in isolation. What we can learn from this is that any curricular policies aimed to improve students' math scores should consider both income group and general knowledge level interactively.

Post-hoc Test

Term	Y	Group1	Group2	df	Statistic	P	p.adj	p.adj.signif
FGKS*IG	FMS	1	2	11929	-5.564928	2.679181e-08	8.037542e-08	****
FGKS*IG	FMS	1	3	11929	-14.862992	1.588335e-49	4.765006e-49	****
FGKS*IG	FMS	2	3	11929	-9.684706	4.223559e-22	1.267068e-21	****

Table 6. Pairwise comparison of estimated marginal means (EMM)

To read Table 6, the 'statistic' column shows the standardized mean difference between groups classified by IG. Since values in this column are all negative, it means that the former group in each row (Group1) has a lower standardized mean of FMS than the latter group (Group2). The column 'p.adj.signif' demonstrates the significance level of pairwise differences using asterisks based on adjusted p-values. The presence of four asterisks for all comparisons suggests that there are significant differences in FMS among different levels of IG with the interaction of FGKS. The adjusted p-values are below the significance level of 0.05, indicating strong evidence against the null hypothesis of no difference.

FGKS	IG	Emmean	se	df	Conf.low	Conf.high	Method
23.07369	1	26.10007	0.1088130	11929	25.88677	26.31336	Emmeans test
23.07369	2	27.00183	0.1181152	11929	26.77030	27.23335	Emmeans test
23.07369	3	28.66168	0.1262030	11929	28.41430	28.90905	Emmeans test

Table 7. Table of estimated marginal means by IG

As can be seen from Table 7, the estimated marginal means for FMS increase with the IG when the covariate FGKS has been controlled for, suggesting higher income groups have higher adjusted mean math scores. This proves that the minor differences observed among income groups, as indicated by the varying distances between the regression lines in Fig 2, are indeed significant. The means of the IG are statistically significantly different from each other given the outcome FMS when holding the continuous variable FGKS constant.

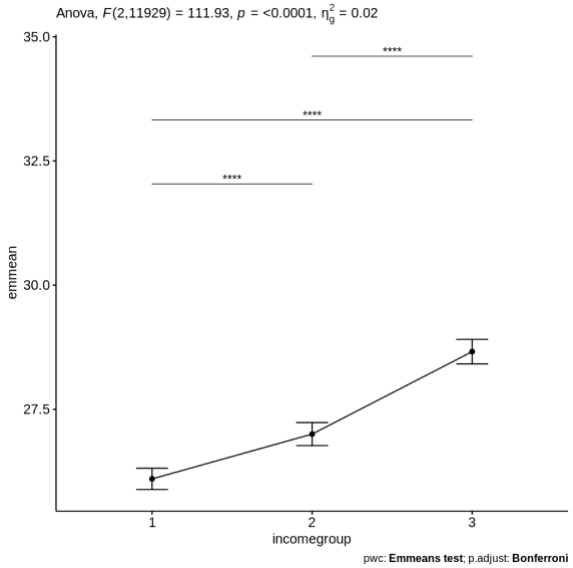

Fig 3. Line plot of Emmeans

Fig 3 contains visual representations of information obtained in Table 6 and Table 7.

Checking ANCOVA Assumptions

I constructed a linear regression model, and the Shapiro-Wilk test was employed to assess the normality assumption. The p-value is far below the significance level of 0.05 providing strong evidence against the null hypothesis that states the sample (i.e., residual) came from a normally distributed population.

The p-value obtained from Levene's test is less than the significance level of 0.05, indicating a rejection of the null hypothesis, thereby suggesting the assumption of consistent variances across groups is not met.

However, ANCOVA is often robust to mild violations of normality and homogeneity of variances assumptions, particularly with larger sample sizes. It's more important to verify the fixed-x assumption and independent errors assumptions.

Fig 2 shows a linear relationship between the covariate and dependent variable. Moreover, there is little difference in the regression coefficients between each income group. We can conclude that the effect of covariate on the dependent variable is approximately constant across all levels of income groups. Hence, the fixed-x assumption is met.

The independent errors assumption is also met because separate subjects are included in the datasets, and the subjects are not allowed to collaborate.