Suha Lee
1005155626
23-03-2024

INF2178 Technical Assignment 3 Narrative

Examining Early Childhood Development by Income Class

1. **Introduction**

   Early childhood is a critical period in a child's development, characterized by rapid growth in capacities both physical and mental. In particular, children's cognitive abilities at this age are highly malleable and formative, setting the foundation for future learning and intellectual growth. The skills acquired during these formative years, such as in language, literacy, and numeracy, are crucial building blocks for subsequent educational milestones and effective functioning in daily life.

   These measures are often considered pivotal for their academic success and overall life outcomes, and hence are at the center of numerous interventions aimed at enhancement and equitable access to quality education. Recognizing the importance of these developmental stages, educational policies and practices are increasingly designed to support and nurture early cognitive development, ensuring that all children have the opportunity to reach their full potential.

   We are motivated to investigate the interplay between household income and the cognitive development of children in early childhood education settings.

   Our study seeks to address two research questions:

   Research Question 1 (RQ1): How does household income influence the academic achievements of kindergarten children?

   Research Question 2 (RQ2): Can we identify distinct trajectories in the development of students' skills when categorized by income group?

   By examining the influence of economic factors on learning and development, we aim to uncover insights that could inform targeted interventions. These interventions could help level the playing field, providing children from diverse economic backgrounds with the necessary support to thrive academically and beyond.

2. **Exploratory Data Analysis (EDA)**

   Our study makes use of a subset of data from an early child longitudinal study conducted between 1998-1999. The dataset contains information on students'

performance across three different domains, as well as their household income at the time. The dataset contains a total of 11933 observations.

Students are assigned a categorical class '*incomegroup*' based on their household incomes (Table 1). The students are distributed into three different groups, skewed towards the lower income classes (Figure 1).

incomegroup Class	Total Household Income
1	< 40000
2	≥ 40000 and < 70000
3	≥ 70000

Table 1: Income Group Classifications

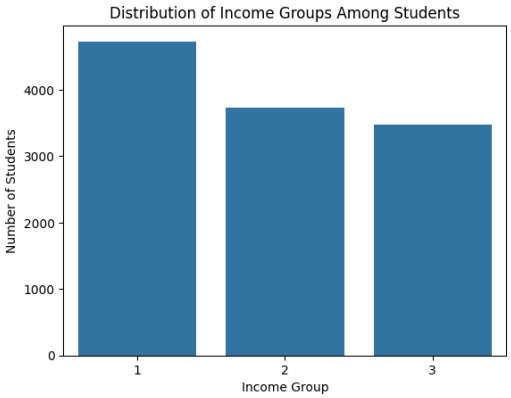

Figure 1: Distribution of Students by Income Group

For each student, the dataset contains six different scores. These scores cover three different areas – math, reading, and general knowledge – across the Fall 1998 semester and the Spring 1999 semester. For this study, we will assess academic achievement as outlined in our research questions using the reported reading and math scores, while using general knowledge as a baseline for aptitude.

To examine the general effect of household income on academic achievement (RQ1) we will focus on the scores attained in the Fall semester, limiting the effect of within-year academic development, and isolating the influence of socio-economic status on initial academic performance. We will then turn our focus to academic development to investigate how changes in academic performance relate to income group distinctions (RQ2) by examining the differences in scores between the Spring and Fall semesters.

The plots in Figure 2 display comparisons of reading and math scores attained in the Fall semester. The boxplots show that mean scores tend to rise with income group, while variance within each group is relatively equal.

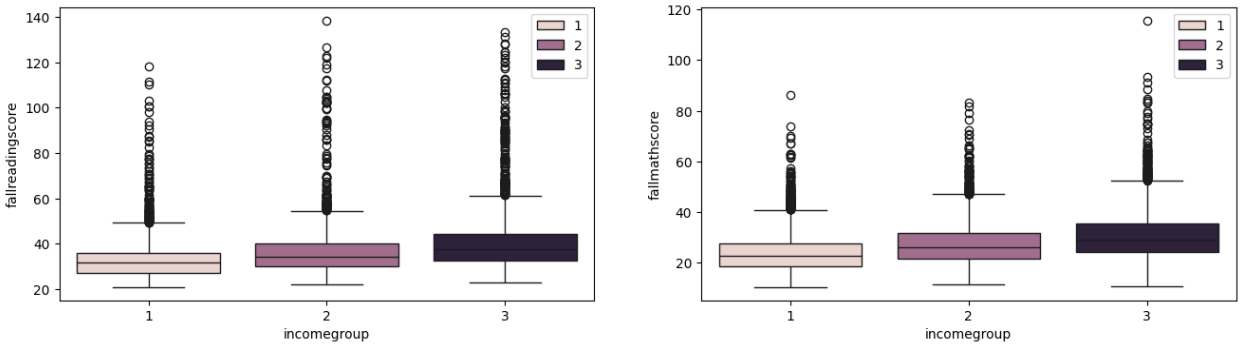

Figure 2: Boxplots of Fall Semester Academic Scores by Income Group

In Figure 3 below, the mean differences in reading and math scores between the Fall and Spring semesters is relatively equal for all three income groups, as are the variances.

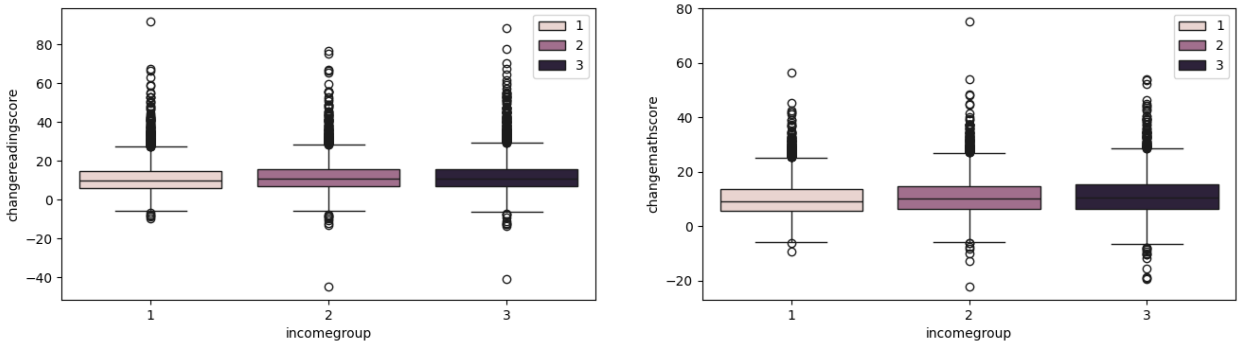

Figure 3: Comparative Analysis of Changes in Academic Scores

Based on these preliminary findings, we have reason to believe that while differences in academic performance can be in part attributed to income class, there does not seem to be discernable differences in how students of different income groups benefit from the existing educational system.

3. One-way ANCOVAs

To further investigate our research questions with more statistical rigour, we will employ one-way ANCOVA. We seek to measure the effect that income has on academic performance, when considering individual baseline aptitude.

Using the Fall scores for reading and math as the dependent variable, we conduct two ANCOVA tests using income group as the independent, categorical variable and the Fall general knowledge scores as the covariate (Table 2). For a confidence level of 0.95, the results of the test are all statistically significant. They indicate that when accounting for general knowledge aptitudes, an increase in income class correlates to an increase in reading and math scores by 1.483 and 1.266 respectively. The covariate is also similarly correlated (albeit at a smaller scale), with coefficients of

0.630 and 0.695 respectively. Interestingly, the intercept for reading is much higher than for math, at 18.603 compared to 8.679. This suggests that children across the board tend to have greater difficulty with numeracy skills, highlighting a potential area for increased focus in the curricula.

Subject	Source	Coefficient	Std. Error	t-statistic	p-value
Reading	Intercept	18.603	0.291	63.875	<0.001
	Income Group	1.483	0.108	13.673	<0.001
	General Knowledge	0.630	0.012	52.238	<0.001
Math	Intercept	8.679	0.231	37.654	<0.001
	Income Group	1.266	0.081	14.721	<0.001
	General Knowledge	0.695	0.010	71.619	<0.001

Table 2: Summary of ANCOVA Results for Fall Scores

To test the appropriateness of ANCOVA as our method, we address each of the required assumptions. In particular, we are interested in normality of distributions, homogeneity of variances, and homogeneity of regression slopes. For normality we conduct the Kolmogorov-Smirnov test and obtain a p-value <0.001. For homogeneity of variances, we conduct a Levene test and again obtain a p-value <0.001. These results indicate that for a confidence level of 0.95, we have statistically significant evidence to reject the null hypotheses of normality and equal variances respectively, failing the two assumptions. In addition, we check for homogeneity of regression slopes by comparing the effects of the covariate on the dependent by income group (Figure 4).

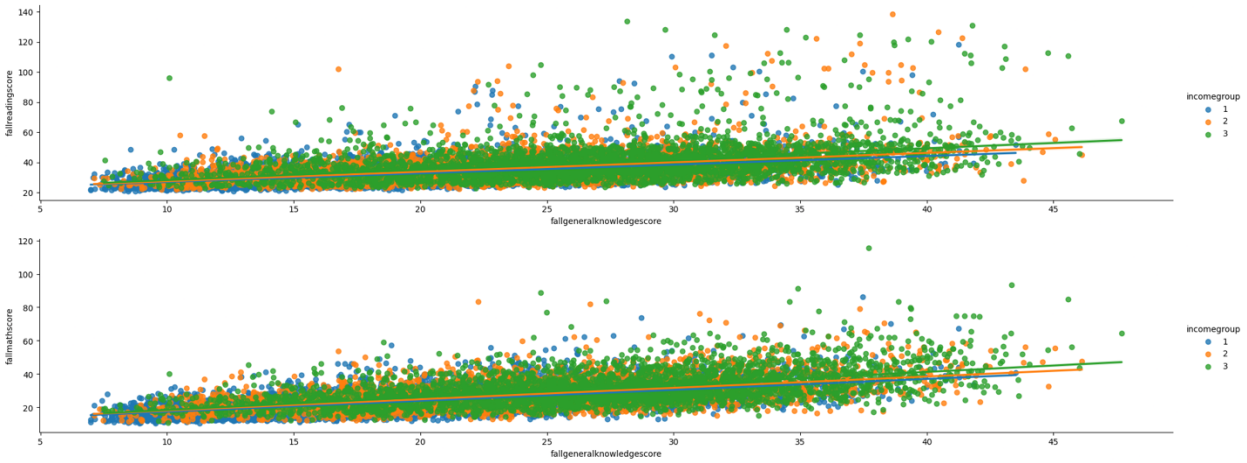

Figure 4: Interaction Plots for General Knowledge on Academic Scores by Income Group

As the slopes are generally parallel, we can assume that there is no significant interaction between the covariate and the dependent, as required. However, given the lack of clear distinction between the groups, our findings may not be reliable.

Overall, the findings from our assumption tests give us low confidence in the results of our ANCOVA test.

To investigate RQ2, we want a measure of the change in students' performance after being exposed to the educational system. We will use the net difference in scores between the second and first semesters of the academic year, compared by income group while accounting for the net difference in general knowledge. We conduct two such ANCOVAs, one for reading scores and one for math scores (Table 3).

Subject	Source	Coefficient	Std. Error	t-statistic	p-value
(Net) Reading	Intercept	8.965	0.206	43.455	<0.001
	Income Group	0.720	0.089	8.084	<0.001
	General Knowledge	0.238	0.018	13.155	<0.001
(Net) Math	Intercept	8.008	0.174	46.008	<0.001
	Income Group	0.730	0.075	9.711	<0.001
	General Knowledge	0.248	0.015	16.255	<0.001

Table 3: Summary of ANCOVA Results for Net Difference in Scores

The findings from our two ANCOVAs are statistically significant for every source. Compared with the results in Table 2, the results for reading and math are much more similar. The intercepts show that regardless of income group and general knowledge, students tend to improve by about 8-9 points between the Fall and Spring semesters in both subjects. When a student is placed in a higher income group, their level of improvement is expected to increase by about 8%. Their coefficient corresponding to general knowledge is much lower, indicating that for every point of growth in general knowledge they can expect an increase of only 0.24 in their net improvement. This may suggest that major development in these subjects requires more specialized instruction.

Once again following the approach above to test the assumptions for our ANCOVAs, we obtain p-values less than 0.001 for both the Kolmogorov-Smirnov for normality and the Levene test for homogeneity of variances. We thus reject the null hypotheses stating that the distribution is normal and the variances are equal, failing two assumptions required for an ANCOVA. For our third assumption test we again plot the interactions between the covariate and the dependent variables (Figure 5). As with our previous ANCOVAs, while the slopes are visually parallel, the dispersion of the data points makes it difficult to assign a line of best fit with confidence.

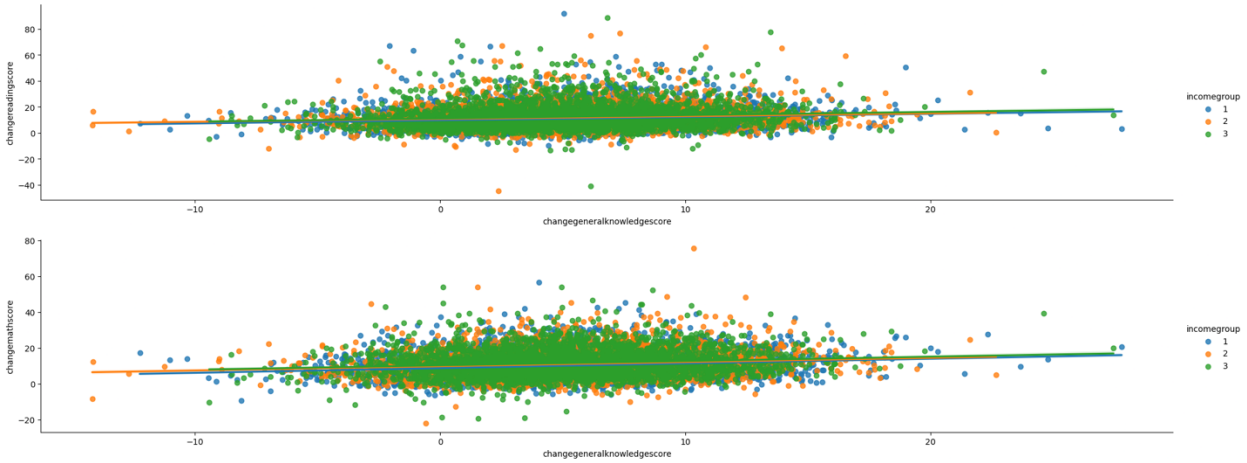

Figure 5: Interaction Plots for General Knowledge on Net Improvement by Income Group

4. Conclusion

The analysis of kindergarten academic achievements in relation to household income (RQ1) reveals that income level is a significant predictor of academic performance in both reading and math. Controlling for baseline aptitude as measured by fall general knowledge scores, our ANCOVA results show that children from higher income groups tend to have higher scores in reading and math. This suggests a clear socioeconomic gradient in educational outcomes, even at the kindergarten level. Additionally, the notably higher intercept for reading scores implies that, overall, students may find reading less challenging than math, which could guide educators in focusing more resources towards improving early numeracy skills.

With respect to RQ2, our analysis indicates that income groups do indeed follow distinct trajectories in the development of reading and math skills over the academic year. With all other factors being constant, children from higher income groups show a greater increase in their performance from fall to spring, as evidenced by the positive coefficients for income group in both reading and math scores. This incremental growth, while modest, suggests that socioeconomic factors play a role in the rate of academic development. The relatively smaller coefficient for general knowledge underscores that, although baseline aptitude contributes to skill development, the progress attributable to general knowledge alone is less pronounced. This implies that while general cognitive ability lays the foundation, the magnitude of skill improvement is influenced more heavily by income-related factors, potentially access to resources or learning opportunities that are not captured solely by general knowledge.