Assignment 3: Early Childhood Education Assessment

Educational assessments are an efficient method to measure a child's understanding and progress in their learning. In today's multicultural society, we know that socioeconomic factors play a big role in education. Families with more resources often perform well compared to those with less. This points to the need to tailor learning depending on the population as well as providing more support for children who may not have as many resources for learning as their peers. In other words, equality in learning is a way to account for the influence that socioeconomic status plays on an individual's potential and quality of life.
The following analyses aims to examine longitudinal study of kindergarten students using an assessment of reading, math, and general knowledge for Fall 1998 and Spring 1999. Information on the parents' income is also included to consider how resources may affect learning styles or abilities. The dataset used for this analysis is titled 'INF2178_A3_data.csv'.

Data Pre-Processing

The dataset was loaded into a dataframe titled 'kindergarten', containing 11932 rows and 9 columns which were renamed to meaningful/easy to read names. The column Income_Group was changed to type str so that it can be differentiated as a categorical variable. To explore the dataset, the income categories were aggregated into a new table, as well as the minimum and maximum value of household income. This revealed the number of income categories as well as the range of income within each category.

INCOME_GROUP	MIN	MAX
1	1.0	39800.0
2	40000.0	69700.0
3	70000.0	150000.0

Table 1: Income Groups and the range of total household income within each group.

The Household_Income and Income_Rounded are excluded from all subsequent analyses because it was simply used to link the scores to the income group. A summary table of all scores revealed that the students generally improved when comparing their fall and spring scores. One important note is that while the mean and median of all three subjects were relatively close, the maximum value of general knowledge is much lower. Additionally, a correlation matrix was computed to identify some preliminary observations about how certain variables may be linked to one another. From this, our research questions are as follows:

1) How does the end of term **(spring) general knowledge scores** differ across **income groups**, while controlling for **fall general knowledge scores**?
2) How does the **total spring scores** across all three subjects differ across **income groups**, while controlling for **total fall scores**?

Question 1

A boxplot of the general knowledge scores across each level of income group shows that the kindergarten students tend to score increasingly high as we move across each income group, although there are students that score higher or lower than average (i.e., outliers). Comparing the fall and spring scores, the latter scores are higher in that the range of scores increases over time. The amount of improvement (i.e., the range of fall vs. spring scores) also seems to be consistent across each level of income group.

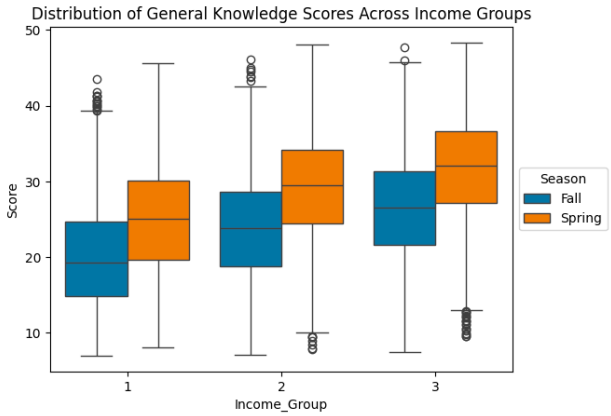

Figure 1: Variation of general knowledge scores across income groups, by season.

The summary statistics for the dependent variable, spring general knowledge scores, and the covariate, fall general knowledge scores, was computed. The results similarly show that there is an improvement in scores over time but also as we move across each income group. It also reveals that the sample size for each income group is consistent for both points of measurements.

	SPRING			FALL(COVARIATE)		
INCOME_GROUP	**n**	**mean**	**std**	**n**	**mean**	**std**
1	4729	25.069	7.248	4729	19.948	6.717
2	3726	29.144	6.965	3726	23.888	6.870
3	3478	31.568	6.928	3478	26.452	7.102

Table 2: Summary statistics of the dependent variable (spring) and covariate (fall) grouped by income group.

A scatterplot of the two scores by income group also shows a moderately positive correlation. This suggests that a one-way ANCOVA may be an appropriate method of statistical analysis to address our research question.

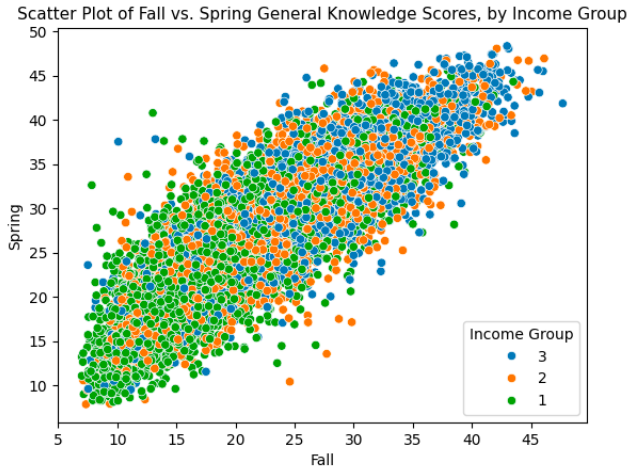

Figure 2: Correlation between fall and spring general knowledge scores by income group.

The one-way ANCOVA was computed to examine whether general knowledge scores in the spring varies across each level of income group, while controlling for general knowledge scores in the fall. Using a significance level of $\alpha = 0.05$, the results showed a p-value of less than 0.001. This suggests

that there is sufficient evidence to reject the null hypothesis that the spring general knowledge scores does not differ by income group.

SOURCE	SS	DF	F	P-UNC	NP2
INCOME_GROUP	1756.904	2	56.908	< 0.001	0.009
FALL	411876.768	1	26682.270	< 0.001	0.691
RESIDUAL	184140.180	11929	NaN	NaN	NaN

Table 3: One-way ANCOVA results for spring general knowledge scores across income group levels, controlling for fall general knowledge scores.

Tukey's HSD was computed to compare the general knowledge scores across each income group. The results revealed a p-value of less than 0.05 for all levels, suggesting that the spring general knowledge scores are significantly different across each income group after controlling for fall general knowledge scores.

Testing ANCOVA Assumptions

Assumption 1: Fall and spring general knowledge scores are linearly related. A linear regression plot of the two variables shows a positive linear relationship across each level of income group, suggesting that this assumption has been met. Note that this assumption was also preliminarily tested through the correlation matrix as well as the scatterplot.

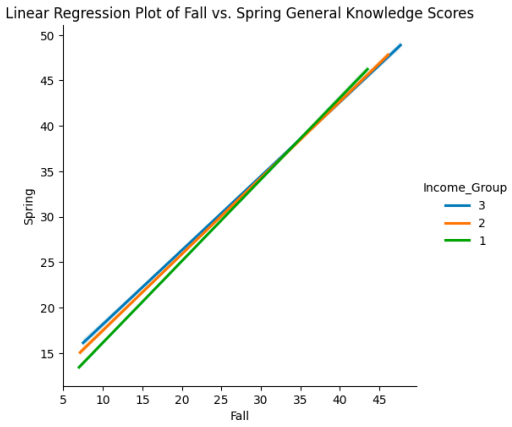

Figure 3: Interaction plot of fall and spring general knowledge scores.

Assumption 2: The residuals follow a normal distribution. The QQ plot show that the residuals have a positive relationship, however it is unclear whether the relationship is linear as it curves slightly at the head and tail of the plot. The histogram of the residuals, however, shows a relatively normal distribution. Using the summary of the Ordinary Least Squares (OLS) regression model, the results of the Omnibus and Jarque-Bera test (both used to test for normal distribution of residuals) yielded a p-value below 0.001. From these results, I am unable to confidently state that the assumption that the residuals follow a normal distribution. This is because a p-value below 0.001 would suggest that a plot of the residuals would very clearly show a non-normal distribution, however this is not the case. Additionally, Central Limit Theorem states that ANCOVA is robust to non-normal distributions given that the sample size is large enough but depending how skewed the distribution is. Thus, we can only state that this assumption has not been met.

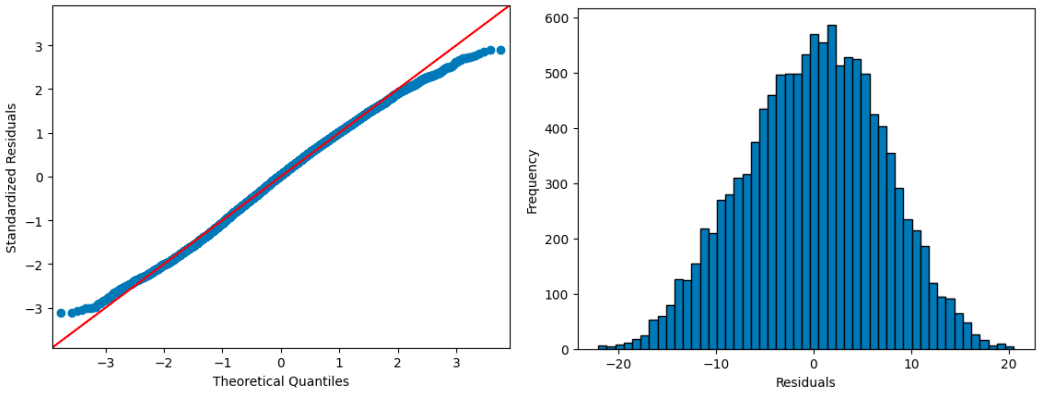

Figure 4: Distribution of residuals as a QQ plot (left) and histogram (right) for one-way ANCOVA of spring general knowledge scores and income group levels, controlling for fall general knowledge scores.

Assumption 3: Each level of income groups has an equal variance. Levene's test was used to test this assumption due to the normal distribution of the residuals. The results yielded a p-value below our significance level, suggesting that we reject the null hypothesis that the income groups do not have equal variance.

We are unable to use ANCOVA to interpret whether spring general knowledge scores are affected by income group, while controlling for fall general knowledge scores. It may be possible to transform the data (i.e., removing outliers or applying log transformations) and repeat testing to see if it improves the distribution of the data. Alternatively, a non-parametric test can be conducted that does not require a normal distribution of the data.

Question 2

A single, total score was calculated for all three subjects for fall and spring. The purpose is to examine how the spring total scores may be affected by income groups levels, while controlling for fall total scores. A boxplot of the scores across each level of income group shows that the kindergarten students tend to score increasingly high as we move across each income group. The range of scores remains quite low, with each level and term of testing having many students scoring exceptionally high (i.e., many outliers above the range of scores). Comparing the fall and spring scores, the latter scores are higher in that the range of scores increases over time. The amount of improvement (i.e., the range of fall vs. spring scores) also seems to be consistent across each level of income group.

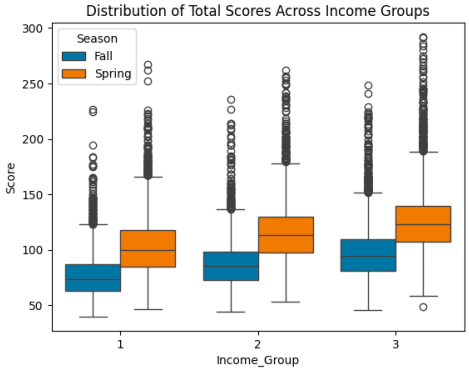

Figure 5: Variation of total scores across income groups, by season.

The summary statistics for the dependent variable, spring total scores, and the covariate, fall total scores, was computed. The results show that the students scored higher during the spring testing compared to the fall, and that students from households of a higher income group tend to score better than those of a lower income group.

	SPRING			FALL(COVARIATE)		
INCOME_GROUP	**n**	**mean**	**std**	**n**	**mean**	**std**
1	4729	102.618	25.765	4729	76.659	19.250
2	3726	115.618	26.889	3726	87.749	21.342
3	3478	126.186	30.241	3478	97.363	24.676

Table 4: Summary statistics of the dependent variable (spring) and covariate (fall) grouped by income group.

A scatterplot of the two variables by income group shows a moderately positive correlation. This suggests that a one-way ANCOVA may be an appropriate method of statistical analysis to address our research question.

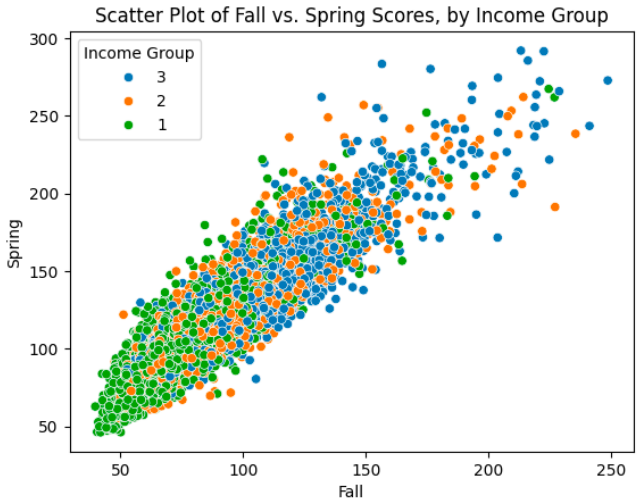

Figure 6: Correlation between fall and spring total scores by income group.

The one-way ANCOVA was computed to examine whether the total scores in the spring varies across each level of income group, while controlling for the total scores in the fall. Using a significance level of $\alpha = 0.05$, the results showed a p-value above 0.05. This suggests that there isn't sufficient evidence to reject the null hypothesis and the spring total scores may not be affected by income groups.

SOURCE	**SS**	**DF**	**F**	**P-UNC**	**NP2**
INCOME_GROUP	701.257	2	2.055	0.128	< 0.001
FALL	6976112.522	1	40883.557	< 0.001	0.774
RESIDUAL	2035489.362	11929	NaN	NaN	NaN

Table 5: One-way ANCOVA results for spring total scores across income group levels, controlling for fall total scores.

Tukey's HSD was computed to compare the general knowledge scores across each income group. The results revealed a p-value of less than 0.05 for all three levels. Testing of ANCOVA assumptions may reveal why the results of these two tests are contradictory. It may also be that there is no significant effect of overall income group, but at each pairwise level.

Testing ANCOVA Assumptions

Assumption 1: Fall and spring total scores are linearly related. This assumption has been met by looking at a linear regression plot of the two variables, which shows a positive linear relationship.

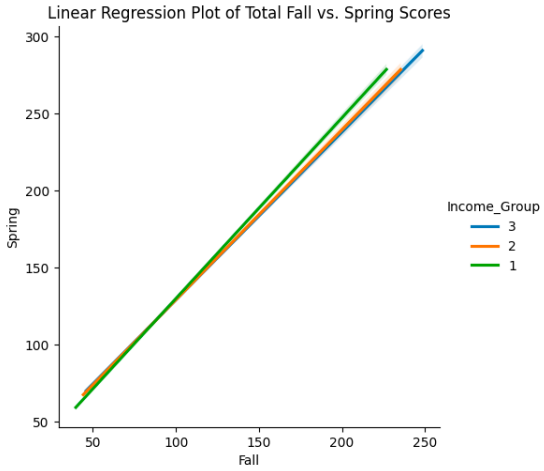

Figure 7: Interaction plot of fall and spring total scores.

Assumption 2: The residuals follow a normal distribution. The QQ plot show that the residuals curves upwards rather than in a linear fashion, suggesting that the model may not be well-fitted. The histogram also shows a single peak that is positively skewed, suggesting that the residuals do not follow a normal distribution. Our observations are further confirmed by the Omnibus and Jarque-Bera test of the OLS summary, which both yielded a p-value below 0.001. We reject the null hypothesis that the residuals follow a normal distribution.

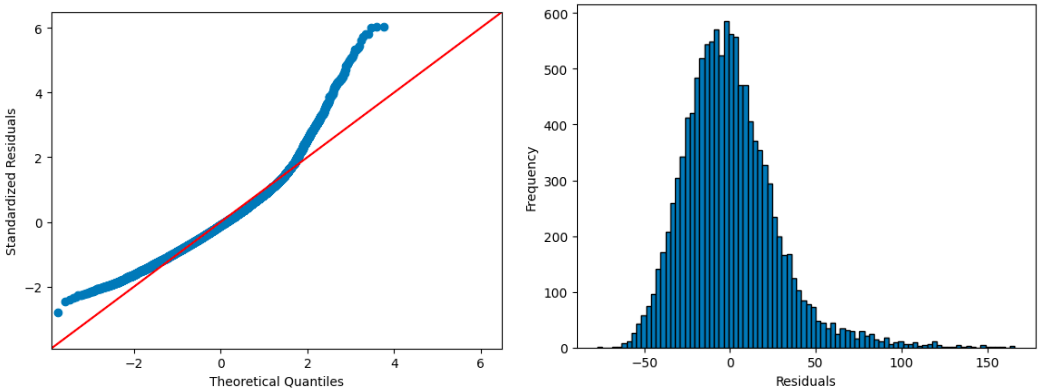

Figure 8: Distribution of residuals as a QQ plot (left) and histogram (right) for one-way ANCOVA of spring total scores and income group levels, controlling for fall total scores.

Assumption 3: Each level of income groups has an equal variance. Levene's test yielded a p-value below our significance level, suggesting that we reject the null hypothesis that the income groups do not have equal variance.

In conclusion, ANCOVA may not be an appropriate statistical test to examine whether spring total scores are affected by income group, while controlling for fall total scores. It is recommended to use a non-parametric test that does not require a normal distribution of the data.