Capstone Progress Report 2

**Stakeholder Names and Roles**

| **Stakeholder** | **Role** |
| --- | --- |
| Russell Carlock | Sponsor: rcarlock@k12albemarle.org |
| Heman Shekari | Team mentor: hs9hd@virginia.edu |
| Stephanie Landas | Team member: sfl7ck@virginia.edu |
| Michelle Wu | Team member: mw3ef@virginia.edu |
| Noah Edwards-Thro | Team member: xjb6yb@virginia.edu |
| Sirish Desai | Team member: skd3nz@virginia.edu |

**Project Title:** Does cultural training of teachers improve student outcomes?

**Abstract**

Research supporting culturally responsive teaching has demonstrated high rates of success in accelerating student learning for students of color with economic hardship. This project will analyze students’ math and reading standardized test scores with classroom teachers going through the culturally responsive teaching and those who have yet to do so. We use student data from the 2023 Albemarle County Public School District to analyze whether the cultural training some teachers received had a statistically significant impact on student academic achievement, specifically, on student reading and math test scores. We will use multivariate linear regression and random forest models, which will highlight the prominent factors for growth of student test scores. The results of this project will inform the Albemarle County Public Schools leadership team on whether or not to implement culturally responsive teaching for all classroom teachers for this next upcoming school year of 2024-2025.

**Outline of the Project**

Research supporting culturally responsive teaching has demonstrated high rates of success in accelerating student learning for students of color with economic hardship. This project will analyze students’ math and reading standardized test scores with classroom teachers going through the culturally responsive teaching and those who haven’t yet. The tools that will be used will be multivariate linear regression and random forest models which highlight the prominent factors for the growth of the students in terms of their test scores. The inputs of the model will be demographic information of the students and teachers and the outputs of the model will be evaluated on the growth of students’ test scores. The results of this project will inform the Albemarle County Public Schools leadership team on whether or not to implement culturally responsive teaching for all classroom teachers for this next upcoming school year of 2024-2025. The assumption is that all of the data for each student has been collected correctly, meaning that any analyses abstracted from it are correctly coded.

**Success Criteria**

The goal as stated by our mentor is to have unbiased knowledge of whether the teacher certification was impactful when it comes to growth of student test scores. Our mentor told us that the most important feature for him in general was how the student grew over the year. He also would like for the findings to be statistically significant. He is also curious about whether the certification could work better for each subject.

However, he also has secondary achievements because he is curious to see what other combinations of variables lead to an increase in growth score. For example, he wants to see if the highest increase in scores could be due to some demographic of the teacher. He is also interested in knowing if the vertical scaled score (achievement scaled across grades and subjects) is influenced by any of the other features.

**Data Assumptions and Limitations**

*Data Limitations*One limitation is due to the Family Educational Rights and Privacy Act (FERPA). As we are working with protected information, no analysis with groups smaller than 10 students/data points can be reported on. One way this has affected the data structure is that students of Asian or multi-racial identities had to be combined into an “other” category for the student race variable.

The mathematics dataset has 343 instances of missing data for the “growth” and “expected achievement” variables and 271 instances of missing data for the “vertical scaled score” variable. The reading dataset has 472 instances of missing data for the “value added score” and “expected achievement” variables, and 66 instances of missing data regarding teacher information (license type, years of experience, highest degree, birth year, gender, and race/ethnicity).

*Data Assumptions*

As we have separate datasets for reading test outcomes and math test outcomes, we are assuming that there is no overlap of students in the dataset. This is in compliance with FERPA as well, as linking a student’s demographic information with both their reading and math scores would allow for identification.

Another assumption is that the data has been correctly aligned. Due to FERPA, we cannot gain access to the raw, confidential information, so we are assuming that the scores are correctly aligned with the student’s demographic information.

**Summary of Data Processing, Data Aggregation**

Our data was received in two .csv files (one reading and one math). These files included student demographic information (race, gender, English learner, disabilities, etc.), teacher demographic information (years experience, highest degree, birth year, gender, race, etc.), whether or not a teacher and a principle are certified as culturally responsible educators (and if so, what type of certification), and student results on Virginia SOLs (Standards of Learning) such as growth, expected and actual achievement, and vertical scaled score (achievement scaled across grades and subjects).

Our data processing only consisted of cleaning the data to make sure the variable types were the same and matching variable names. Being that the goal of our project is to investigate if culturally responsive teaching has any impact on student outcomes, we will attempt to account for the impact of all other variables in relation with whether the teacher is culturally responsive certified. Our potential predictors variables therefore include the teacher and student demographic information (specifically investigating teacher experience/qualification) as well as whether or not a teacher is culturally certified and what type of certification there is.

**Data Visualizations**

Figure 1: Kruskal-Wallis Test (Teacher Certification on Growth) for Math Dataset

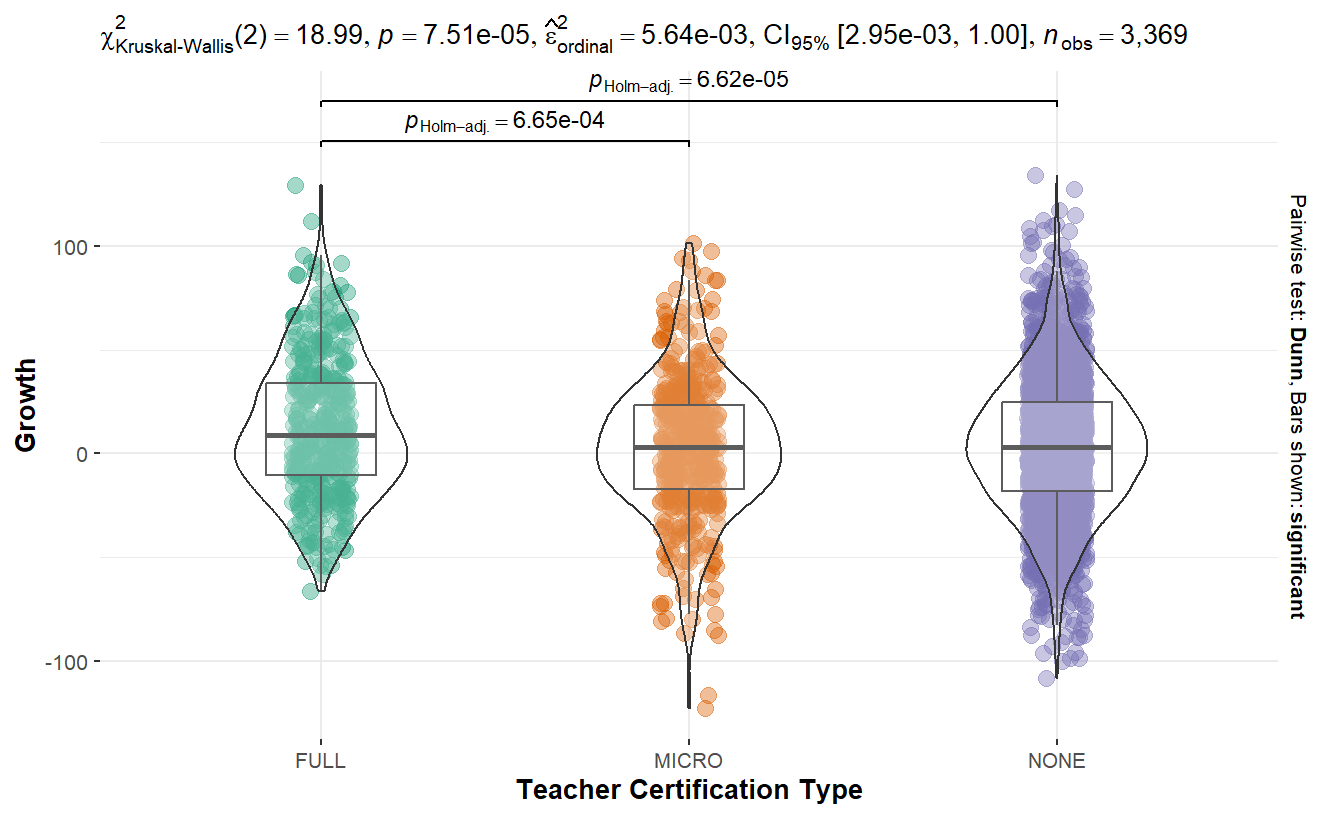


Figure 2: Kruskal-Wallis Test (Teacher Certification on Growth) for Read Dataset

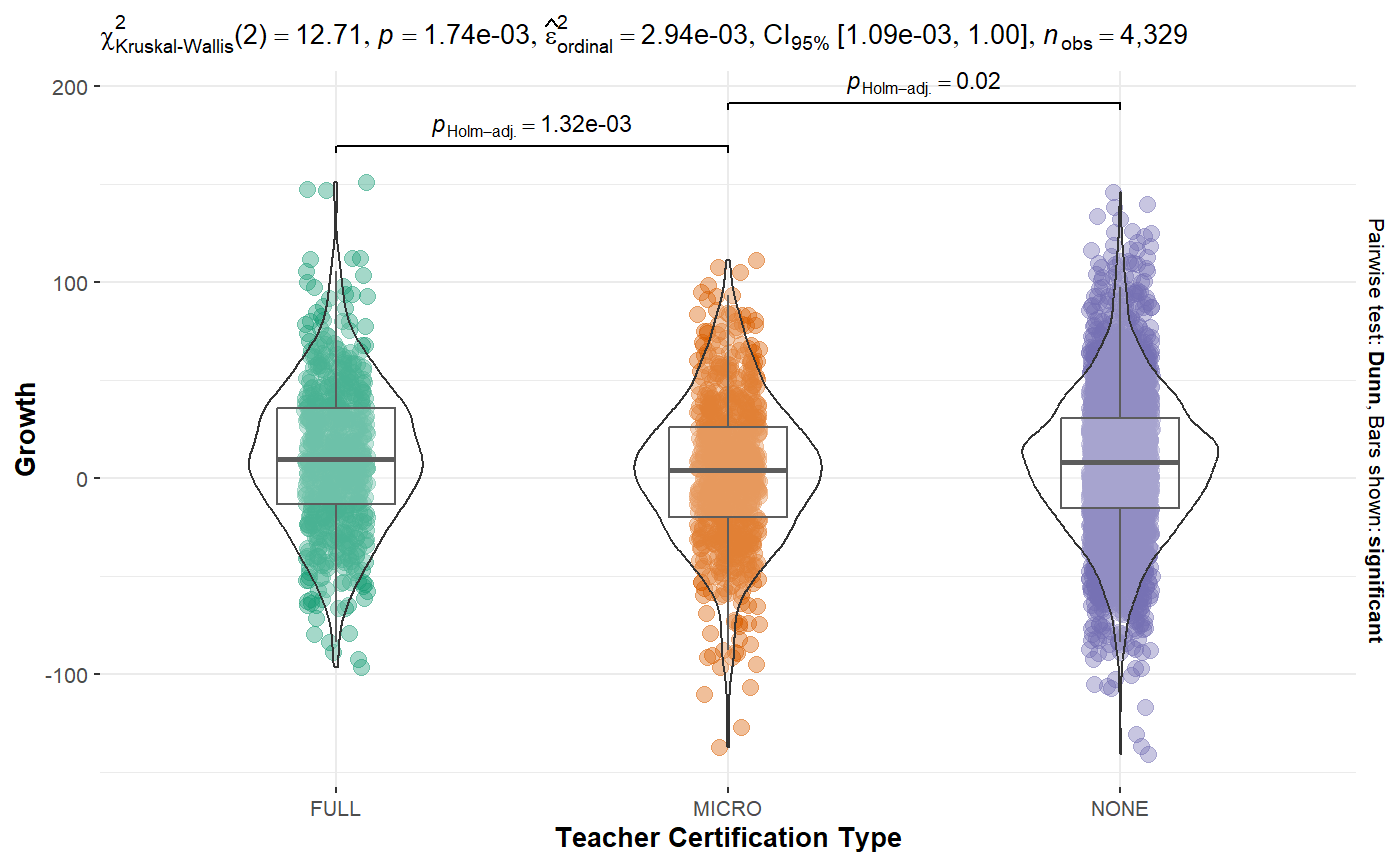


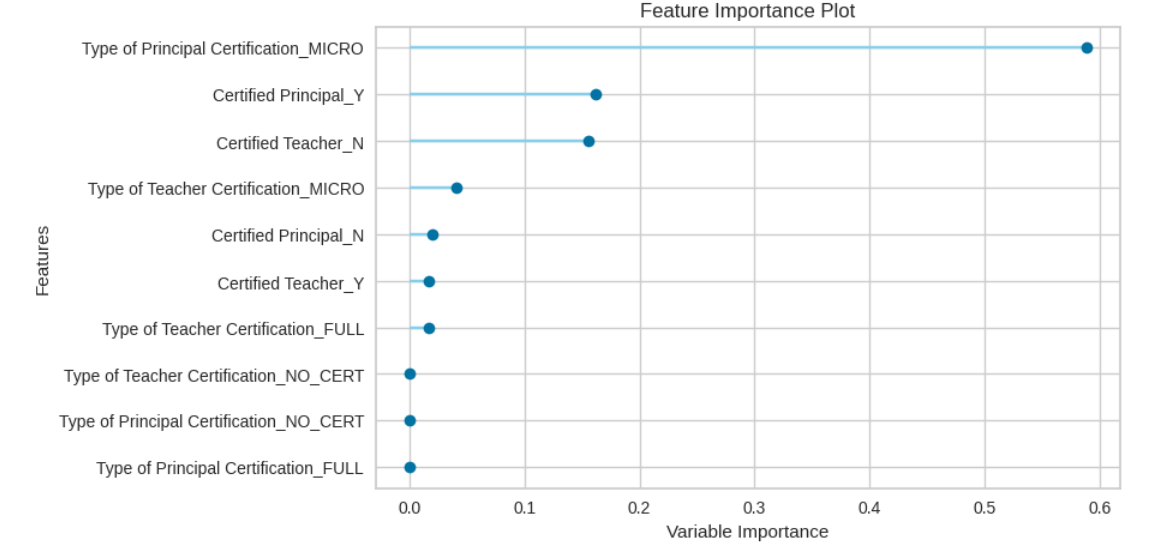
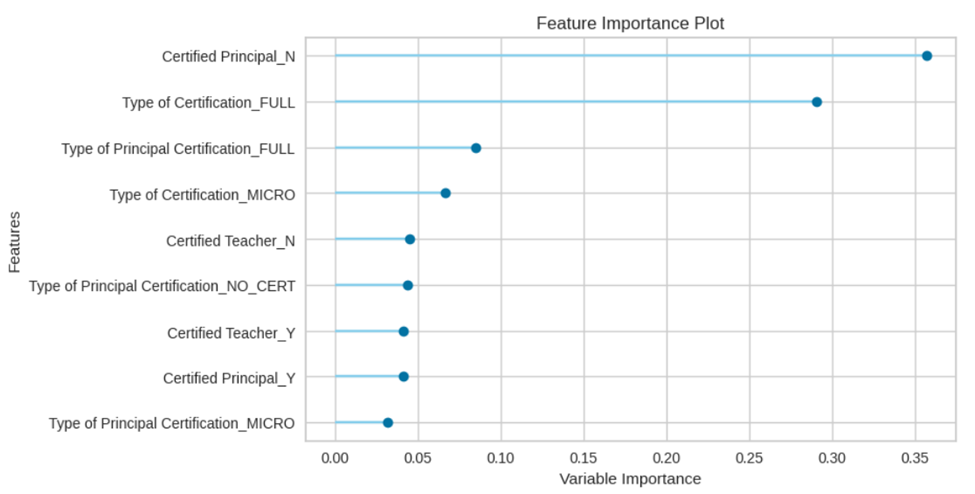
Figure 3: Decision Tree on the Reading Dataset for Teacher + Principal Certification (y = Growth)

Figure 3: Decision Tree on the Math Dataset for Teacher + Principal Certification (y = Growth)

**Summary of Modeling and Analysis**

With the goal of determining if culturally responsive certifications impact student outcomes, we plan to investigate a variety of models. First, we plan to look at multivariate linear regression, specifically looking at which variables are statistically significant. Additionally, we plan to investigate random forests and decision trees, seeking to understand which variables have a high “variable importance” score. Finally, we plan to conduct statistical significance tests on a variety of variable groupings.

**Future Work Plan**

| **Plan** | **Due** |
| --- | --- |
| Continue to run regressions and try to identify and justify which variables are of high importance. Try to come to a consensus as to what is believed to be directionally influential to an increase in growth.  Looking into some features of the data as to why growth has increased in the way it has. | April 8 |
| Continue to build and evaluate models and test for statistical significance | April 15 |
| Write report and finalize presentation | April 28-May 3 |

**Potential Concerns [C] and Blockers [B]**

| **Identifier** | **Description** |
| --- | --- |
| C | Upon further analysis with various regression models, it seems like the whole dataset has high levels of multicollinearity, which has made it difficult to fit a good model on the data. We are going to be looking into interaction terms and other data engineering models to try to fit the data. |
| C | There is limited and unbalanced data, which decreases our confidence in some of the analysis we are making. |

**Questions for Reflection**

*Reflection Question 1: What was the biggest challenge that you faced with this project?*

Our biggest challenge came from our dataset. For starters, it was a small dataset that at first only encapsulated one of the tests and had not many unique variables. Due to this we had a big issue with multicollinearity also. Since the dataset is also small we have issues when it comes to having very significant data.

*Reflection Question 2: Did this project stretch you to grow?* If so, how?

Yes, this project stretched us to grow. This dataset is extremely messy and limited, which matches a lot of real-world datasets that we may face in our careers. By adapting to these limitations, we can apply the skills learned in an academic vacuum to a more realistic dataset.

*Reflection Question 3: Do you believe the capstone experience will be helpful for your career? If so, how?*

This capstone experience has thus far been helpful for our careers just by giving us another experience to work with data science in a real world scenario that is different from our current fields. While the project itself is not very intensive and has many limitations, the experience of working with another organization and communicating with that organization throughout the process has been beneficial.

*Reflection Question 4: Anything else that you would like to share?*

We’ve enjoyed working with our sponsor Albemarle County Schools and Rusty Carlock.