Too many requests in 1 hour. Try again Later.

City Year GenAI Hackathon – Final Presentation

Table of Contents.

01

The Gang

A little bit about the dream team...

The Gang

Artem Beer

Artem is a tech wiz who loves data almost as much as he loves his family. Artem believes AI is the future, and the future is NOW!

Cody Little

Despite his last name, Cody is not little at all, in fact he is fully human sized! Cody is a hands-on techy who makes lots of digital jokes.

Annand Mohanta

Annand has tech wisdom way beyond his years! Annand owns an abacus and prefers it to the calculator app on his phone!

Sarah Lakos

Sarah is really tall and loves learning about new technologies! Her favorite reason to use Chat GPT is for crafting e-mail responses!

02

Data Insights

Here’s a look at what we discovered from the data!

Data Insights

Insight 1 – Children with a focus list or received intervention with a focus list had a much higher propencity of meeting Math and ELA growth goals

Insight 2 – Children who received more than 30 total minutes of dosage in either Math or ELA were more likely to meet their goals

Insight 3 – Children who received more than 70 sessions in math or ELA had a much higher chance of meeting their growth goals (2,000 or so children for ELA as an example and 17 for math)

03

Technology Used

Here are all the tools we used throughout our process!

Technology Used

GitHub

Used for storing and sharing code.

Power BI

Used for creating visual dashboards and data represenations.

Python

Python was used for data cleaning and exploratory analysis.

Logistic Regression

Logistic Regression was used to generate probability scores of meeting growth.

Decision Trees

Decision Trees were used for identifying positive impact thresholds for sessions and minutes.

04

Findings

Here is what we found!

Findings

Insight 1 - More Programs Equate to Better Outcomes

Through logistic regression for probability generation, we have found that the strongest determinants for meeting growth goals include being on a focus list and receiving interventions while on a focus list. While overall session counts and total dosage were found to be impactful the relationship to meeting growth goals was not as strong

Insight 2 - Higher Dosage for Math and ELA Increases the Odds of Meeting Growth

Using entropy-based decision trees to generate an initial split on the child level data we can see that students who had at a minimum 30 minutes of Math or ELA dosage faired much better than those who did not. The takeaway from these results indicate that even small interventions can improve outcomes for students

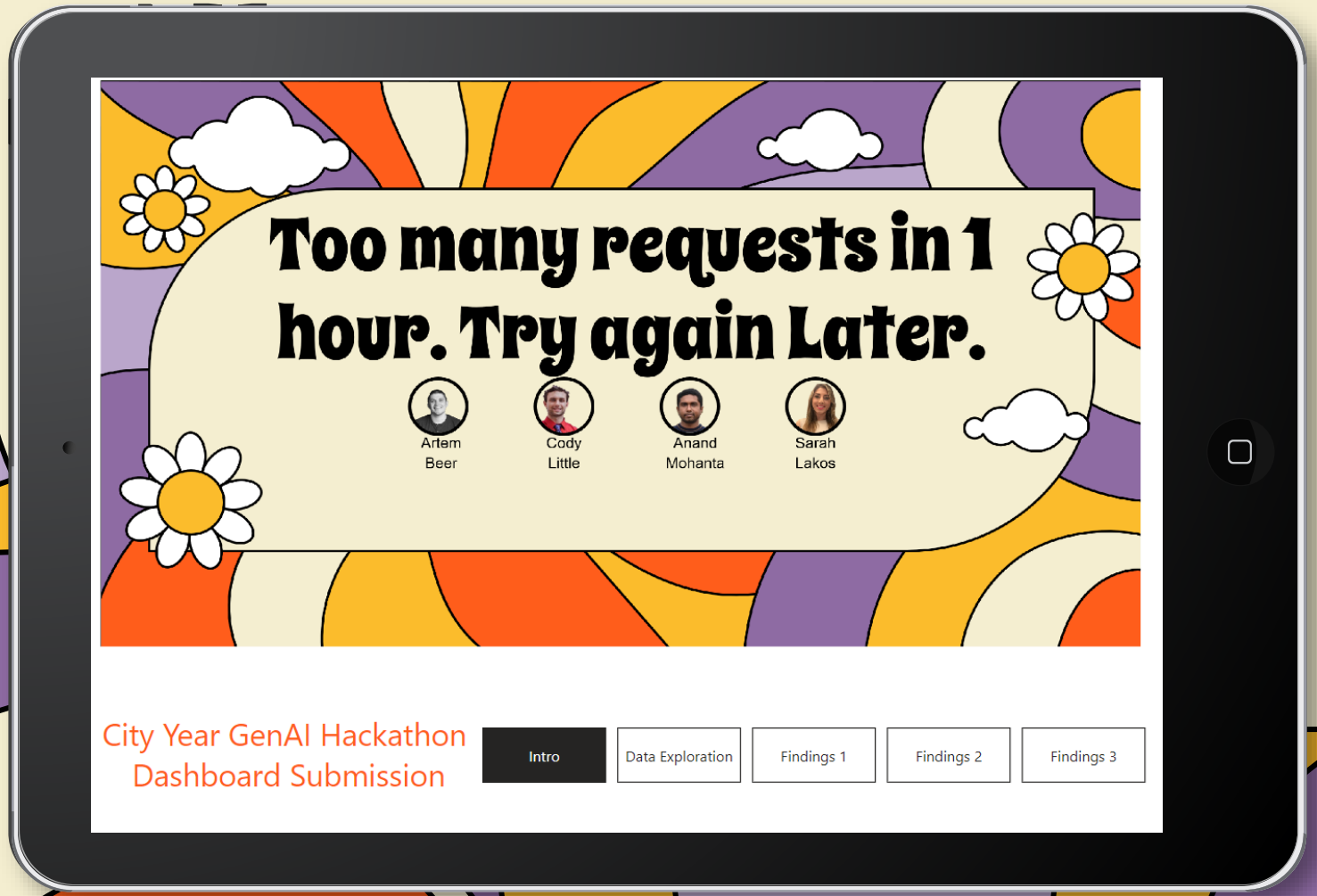
Insight 3 - Higher Session/Lesson Counts Improve Outcomes

Using entropy based decision trees on multiple splits and reviewing the class labels we can see that students who had atleast 17 math sessions or 70 ELA sessions were much more likely to meet their respective growth goals in each area. The major result of this is to show that as session counts increase and approach each threshold the students are more likely to have better educational outcomes

05

Visuals

Look at this!

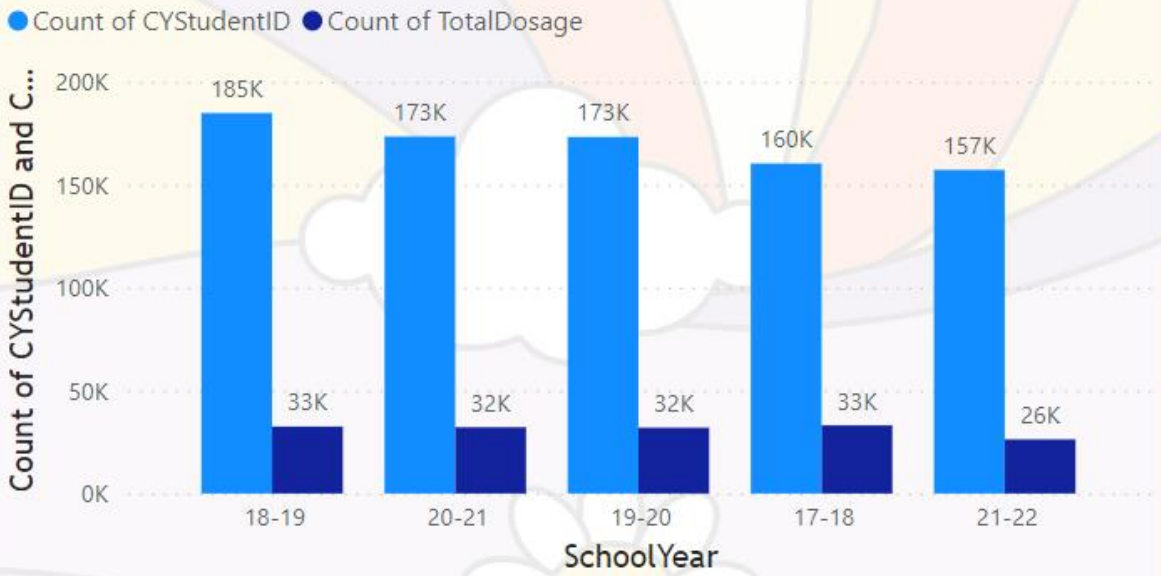


Dashboard

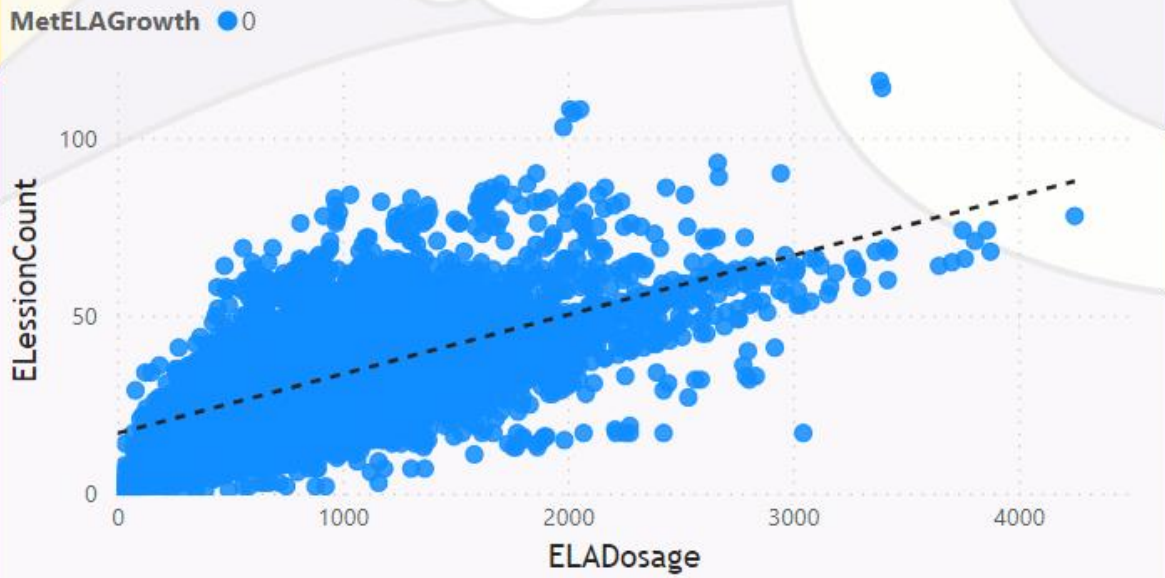
Here's a dashboard to visually depict City Year's data.

Dashboard Link

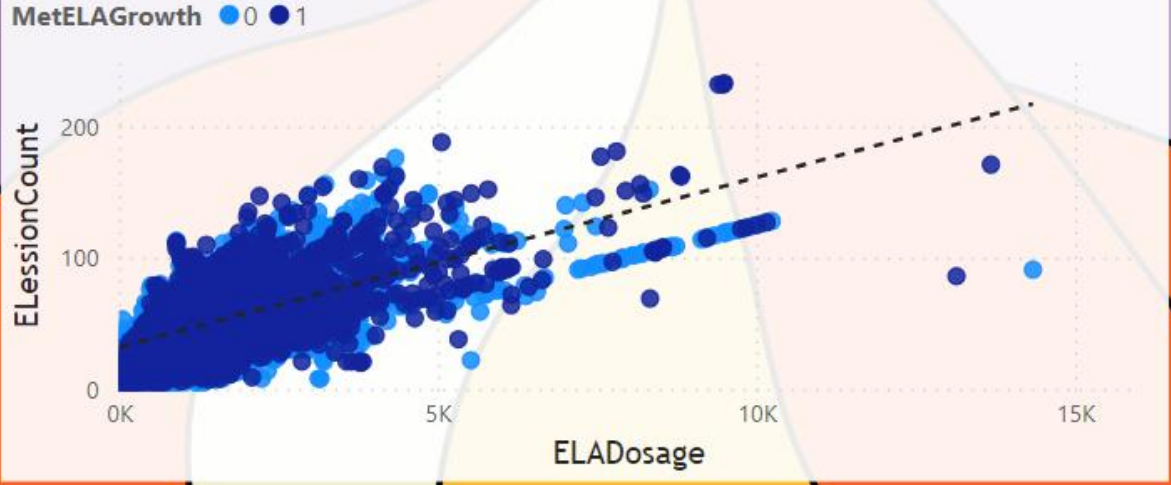
How many students for each year actually had data?



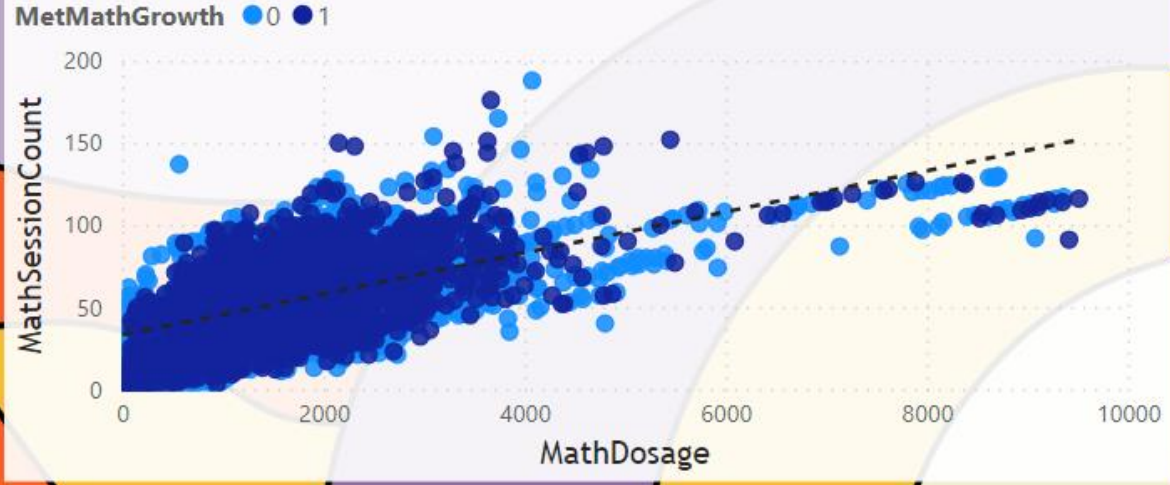
19-20 year was removed as data on goals is missing



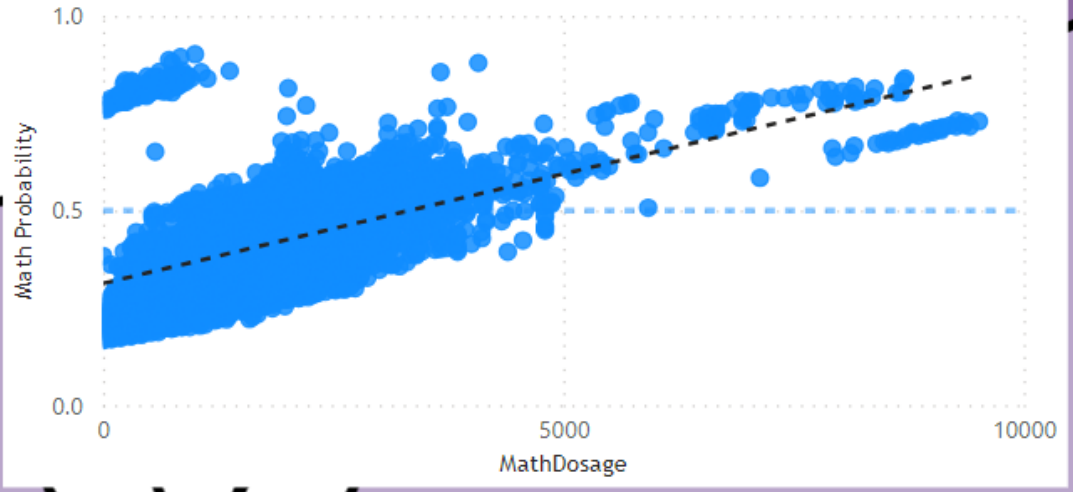
Did session count and length had an impact on whether students met their goals for English?



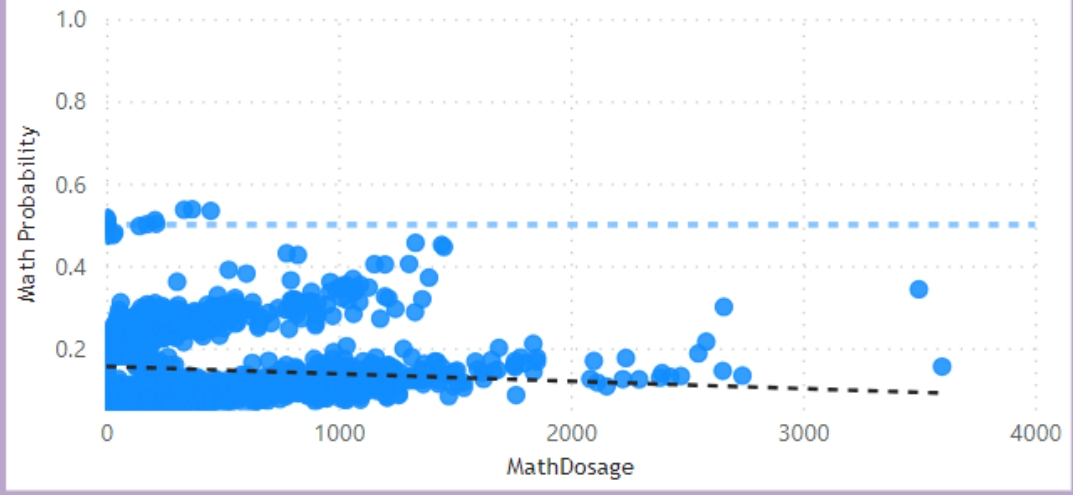
Did session count and length had an impact on whether students met their goals for Math?



Students that participated in the FLMath

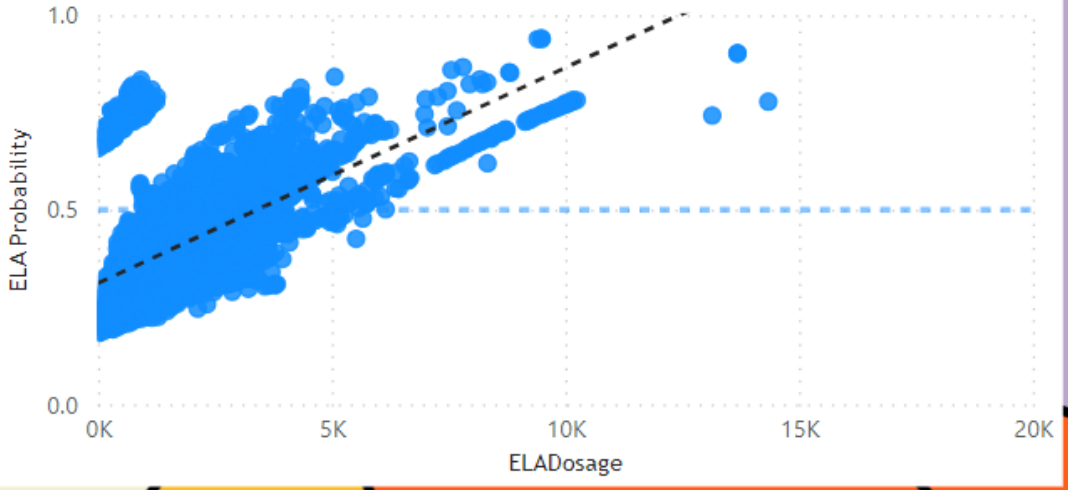


Students that did not participate in the FLMath

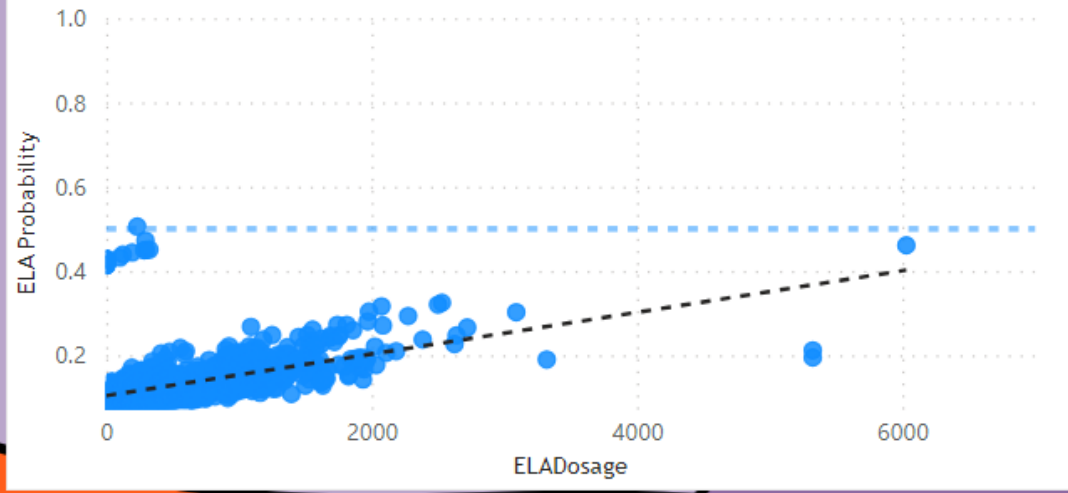


Participation in FLMath and FLELA programs has significant impact on meeting Math and English goals

Students that participated in the FLELA



Students that did not participate in the FLELA



06

Accuracy

Here’s how confident we are in our model!

GitHub:
https://github.com/codlittle/AIGuild-CityYear-Hackathon2023-

Confidence in our Model

Due to the data concerns we took a non-traditional approach to machine learning for City Year

- Most Machine Learning seeks to “predict”
- We sought to find what the best predicters “are”, what we can learn from the mathematical fit of these models, and relate those findings to actionable outcomes

We are confident in our three main insights using these more academic oriented methods which are focused on interpretability. Overall, the insights aggregate to one comprehensive theme….

More involvement or focus for children **improves the likelihood of meeting their growth goals** in ELA and Math

Assumptions/ Suggestions

Assumption

- If City Year was able to capture additional data elements for each student and session, we would better be able to assess the program and each session's success as well as the root cause.
- As we took an academic approach by utilizing machine learning to understand the best predictors and provide findings, we do not have traditional metrics such as accuracy, Cohen's Kapp Score, or AUC, F1 etc.

Data Suggestions and Caveats

To understand the target # of sessions, it would be helpful to have data on students who participated in a large range of sessions and compare their baseline results with their results at different time periods throughout their school year

To understand the ideal length of session, our model would be helpful to assess results for individuals with various session durations and compare the change in their probability to meet goals. Having a lower-level granularity of each session and its length would help inform additional insights

The results presented in these findings utilized dosage, number of sessions, age, involvements in programs such as flexible support and focus lists for both Math and ELA. Data from the COVID year of 2019-2020 was dropped and individuals with 0 total dosage minutes were note included

Appendix

This is where you section ends. Duplicate this set of slides as many times you need to go over all your sections.