

The logo is an oval shape. The top half is white, and the bottom half is black. A series of horizontal wavy lines, alternating between white and black, separate the two halves. A dark gray rectangular box is positioned on the left side of the oval, overlapping both the white and black background areas. Inside this box, the text "LAFAYETTE ADULT RESOURCE ACADEMY" is written in a white, serif, all-caps font, centered vertically. Below the name, the tagline "The place where dreams come true." is written in a smaller, white, serif font.

# LAFAYETTE ADULT RESOURCE ACADEMY

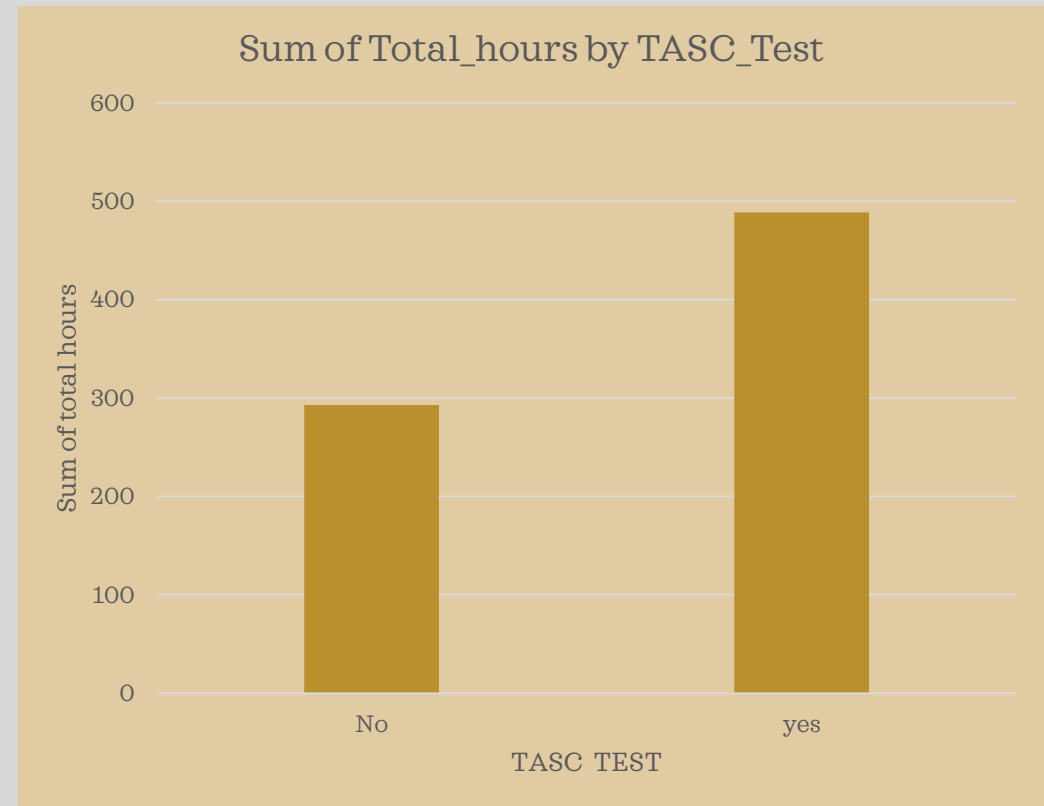
The place where dreams come true.

# Introduction

- **Project Headline:** Here is the analysis of the project that Amy asked me to undertake: trending regular vs not regular students and those are taking vs not taking a post test. I collected data for August and September 2021 . The data set was not large; however, I was able to demonstrate some correlations. Testing information collected by the teachers and students class hours was accumulated by the data director. It was hard for me to gather all data from one resource but with the help of staff I was able to get some of the information needed for my analysis.
- **About the Data:** During fetching up data from different resources and I found many errors and the data was not clean. Some students' names were spelled incorrectly, last names were different, students were assigned in wrong classes, entries were wrong, and some had the teacher's name instead of student's name. Data cleaning was challenging, but fun.

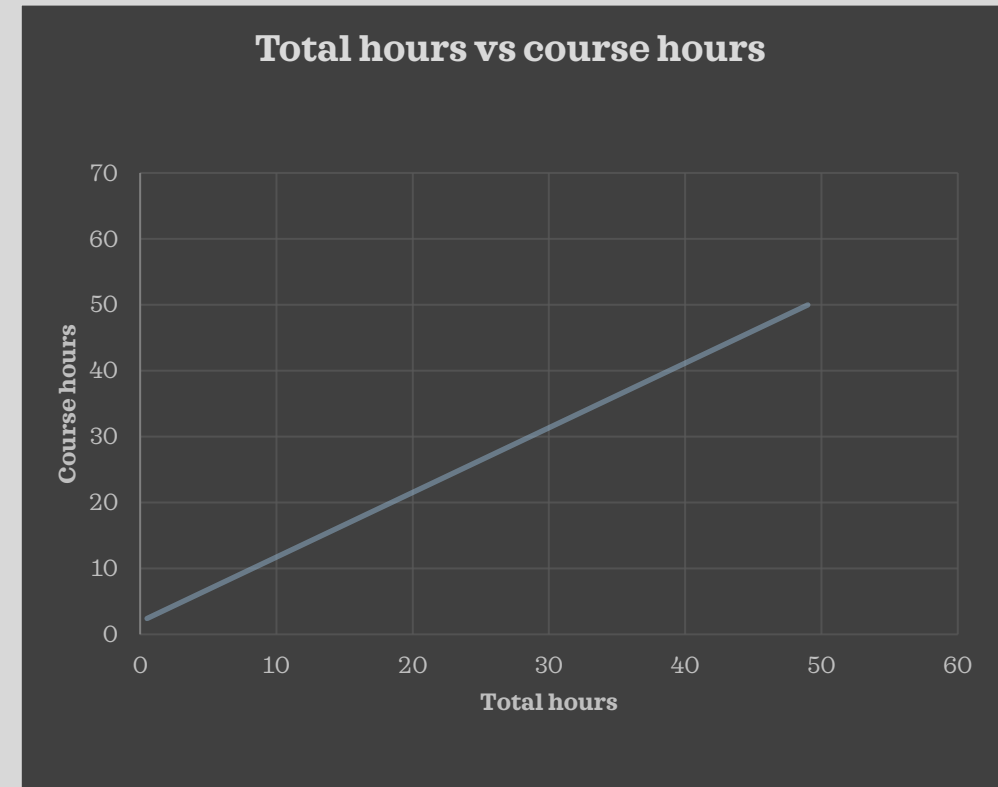
## Sum of total hours vs TASC TEST(Yes/No)

- We can see that those who took the TASC test were more likely on average to have put in more hours than those that didn't.



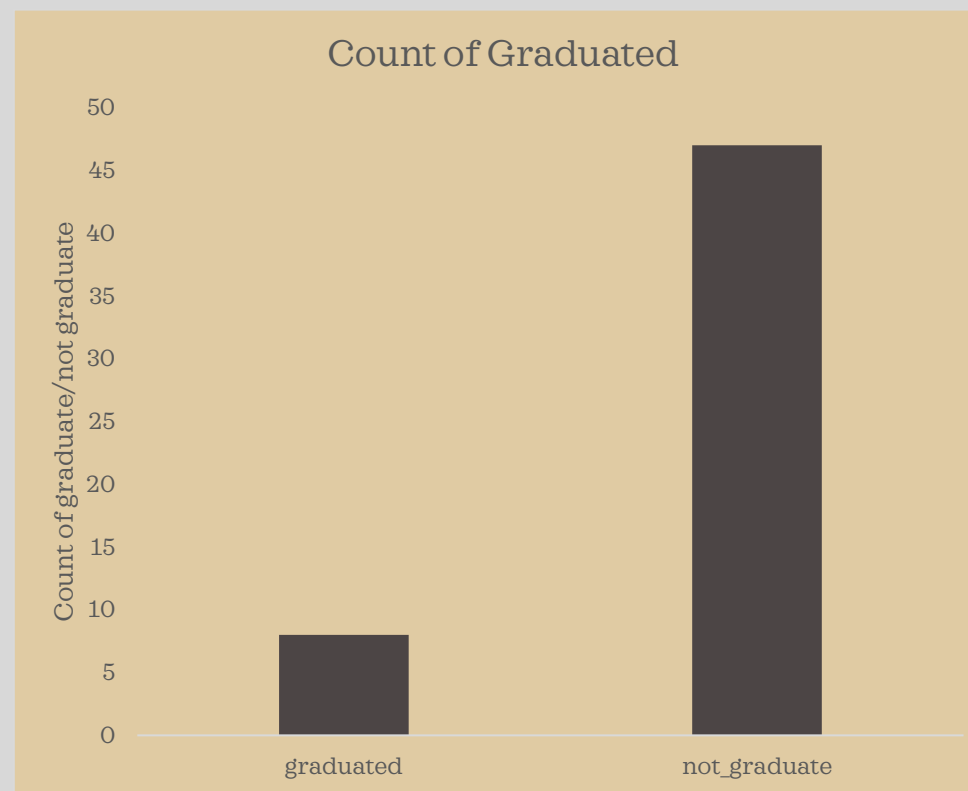
## Scatter chart of Total hours vs course hours HSC

- There is a positive correlation between total hours and course hours, some points are not in the trend, these are called outlier, outliers happens when entries are wrong, we can easily visualize that we have 2 outliers.



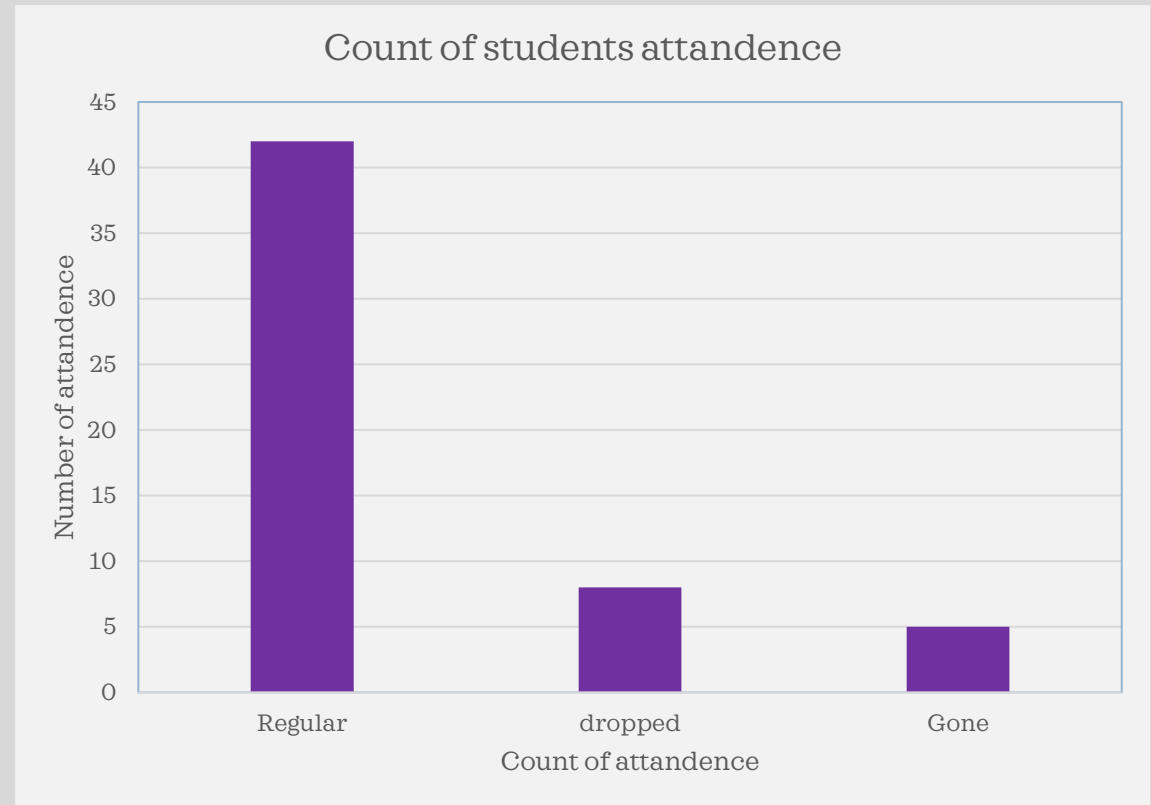
## Count of Graduate and not Graduate students in HSE/GED

- From the data set I was given; we can see most of the students have not graduated so far.



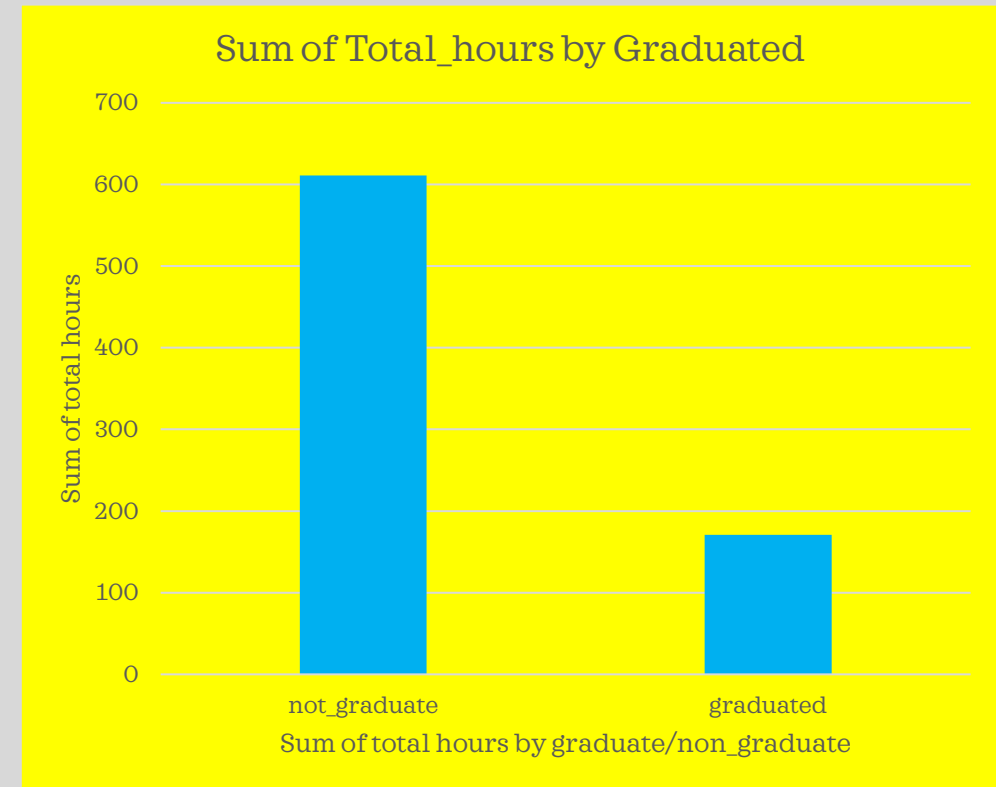
## Clustered Column chart of Count of student's attendance HSC

- 76% of students are regular, 14% students dropped and 9% of students are graduate and they are gone.



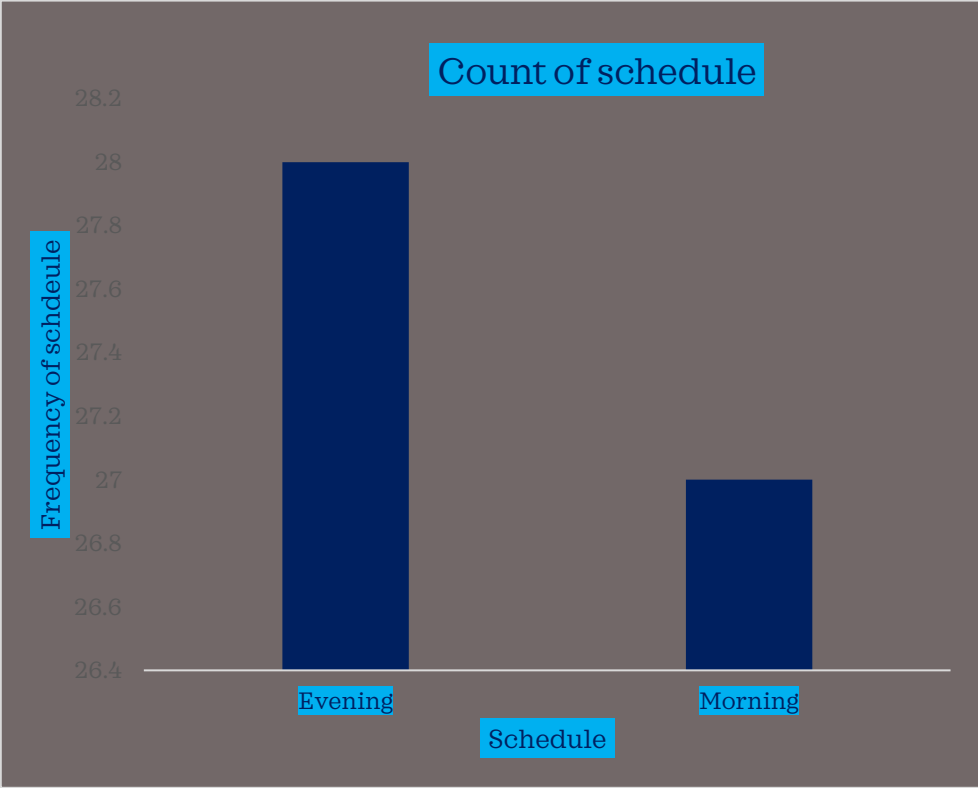
## Sum of total hours vs Graduated/not\_graduated HSC

- When looking at the sum total of all hours, we can see graduate students have less total hours than not graduated students.



# Clustered column chart of Count of schedule HSC

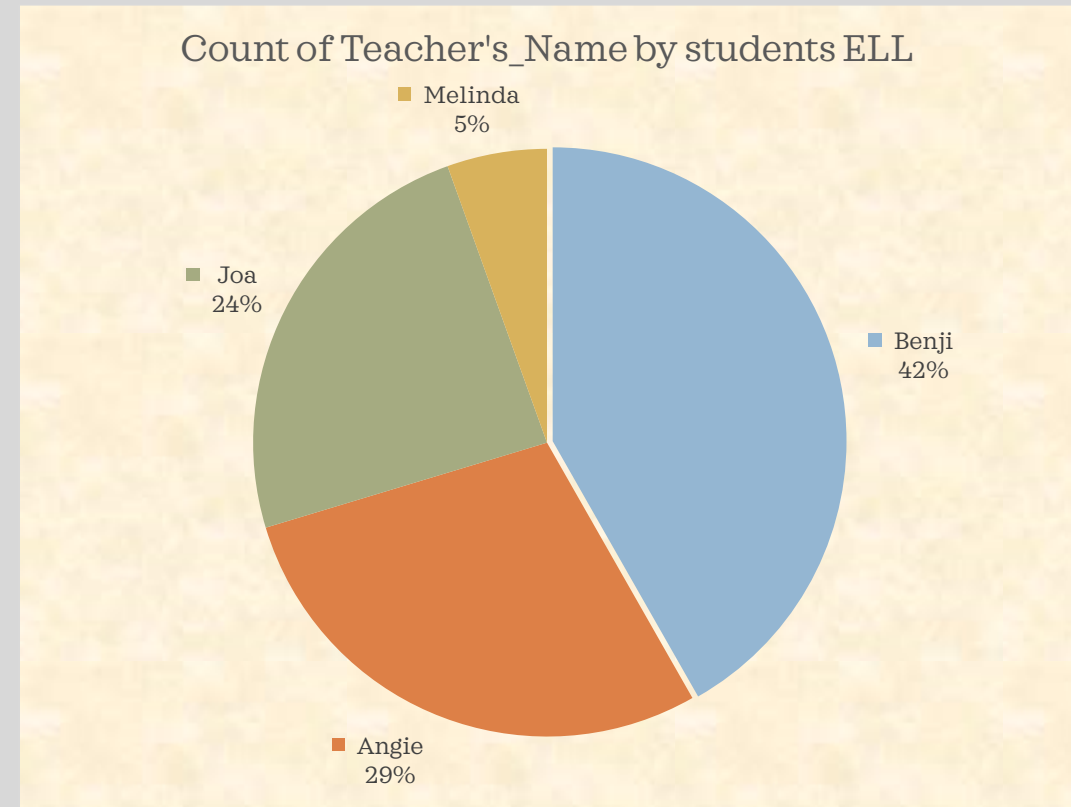
We have more students in evenings/online than mornings.





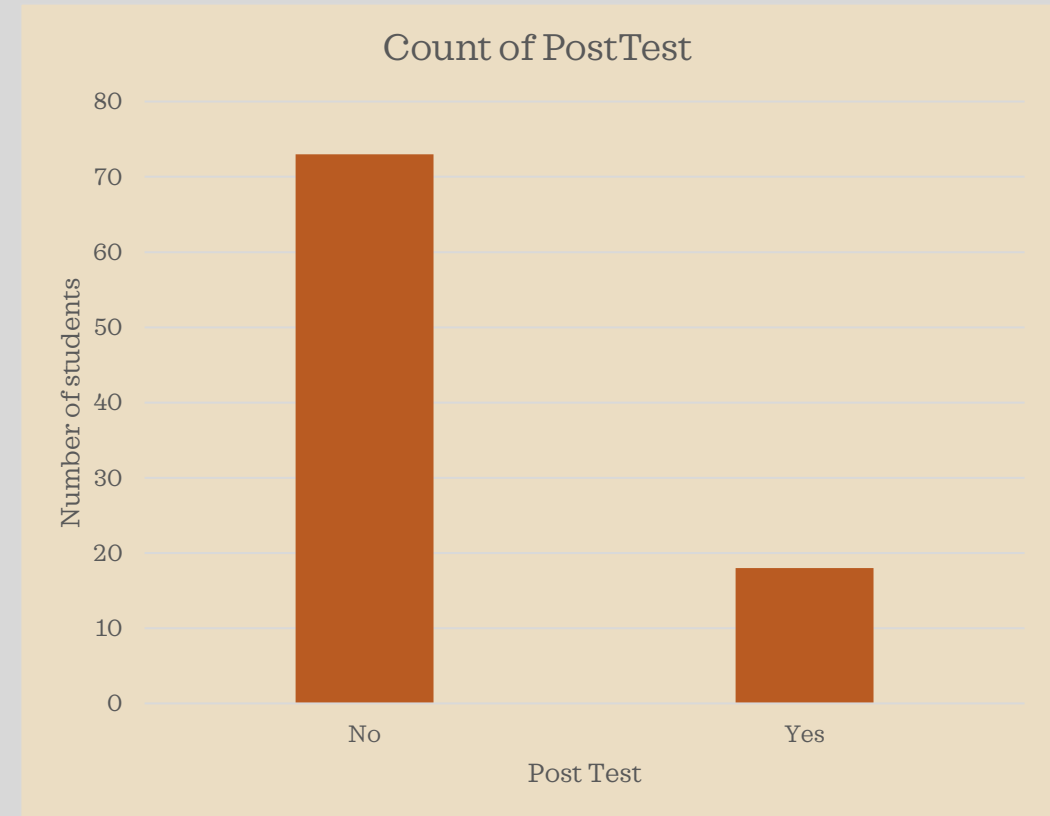
## Which level of students are most in ELL?

- Here is the pie chart of based on student's level. We can see most of the students are in the intermediate level than beginner and advance level.



## How many students have taken post test--Ell

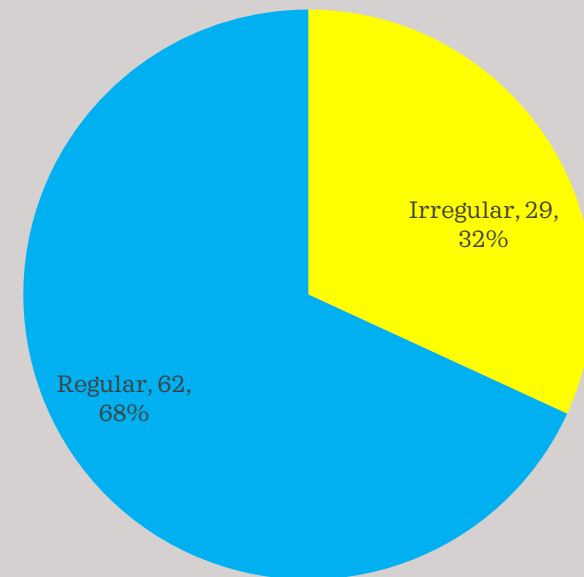
- Here is the bar chart of students who have taken post test. Majority of Ell students haven't taken post test yet. The likely reason is that they didn't complete enough hours in the class for the post test.



How many students are regular/irregular—Ell?

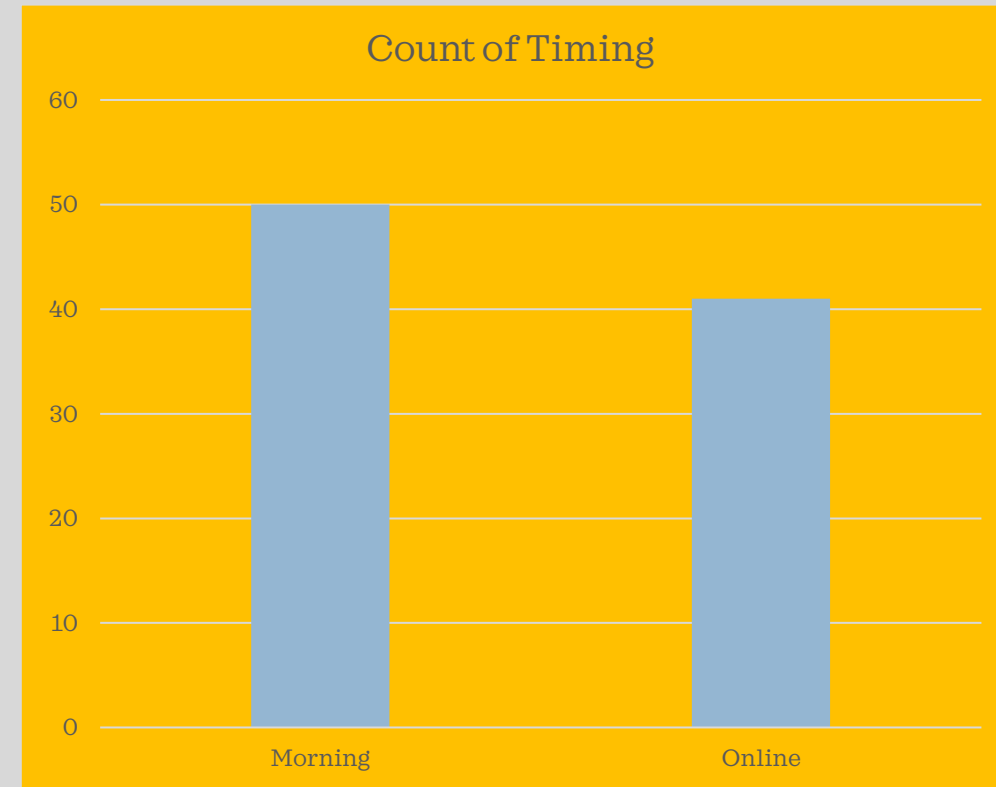
- 68% of students are regular and 32% of students are not regular.

Count of Attendance--Ell



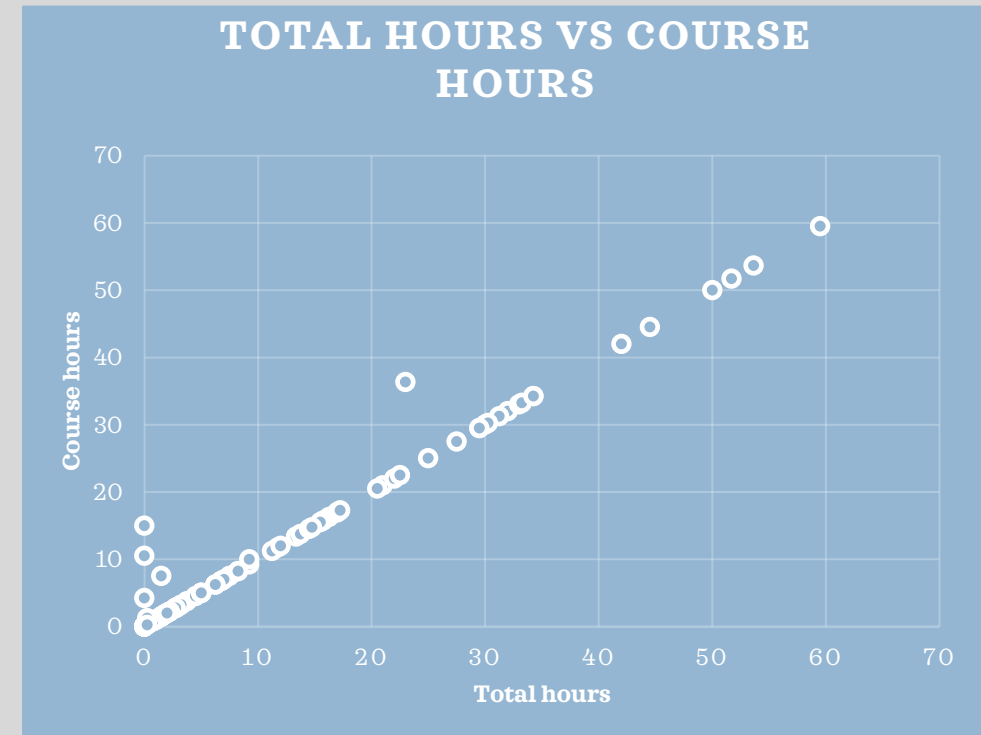
Which schedule are most students—Ell?

- Here is the clustered column chart of student's schedule. Students are more in the mornings than online. From the data we can say that students prefer in person classes rather than online.



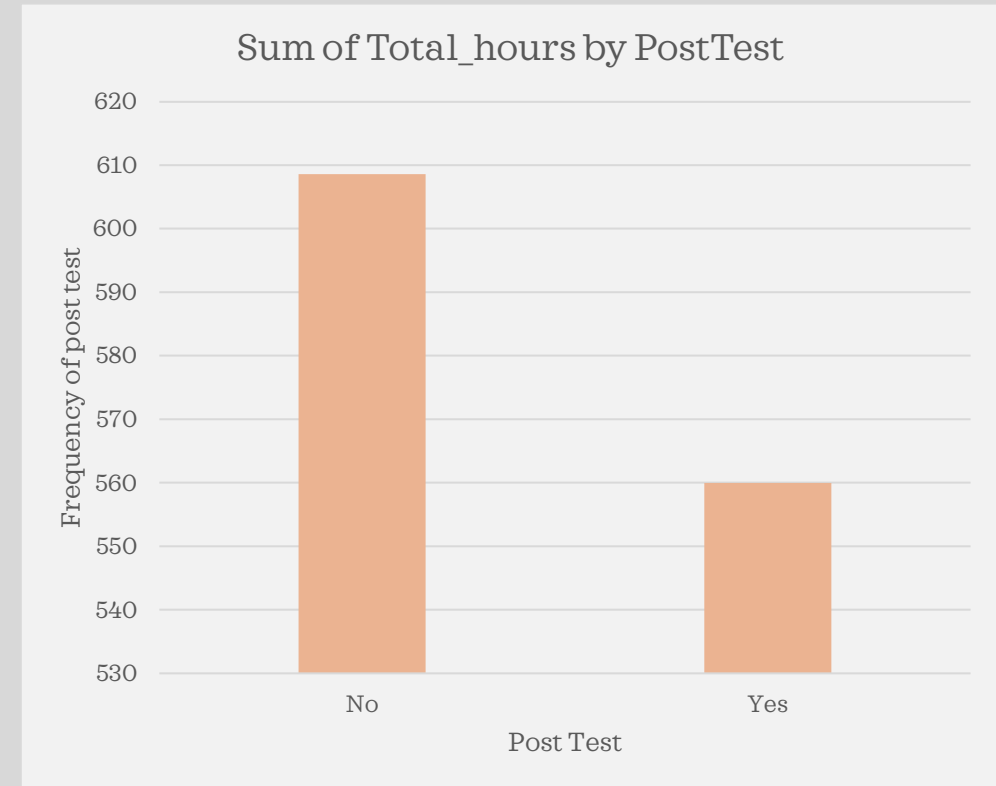
Are the course hours and total hours normally distributed—Ell?

- Here is the scatter chart of total hours and course hours. We can see both attributes are positive correlated with each other.



How many students take test based on their total hours—Ell?

- Here is the clustered column chart of post testers by total hours. Frequency of no post test are greater than yes post test.



## Conclusion:

- Here is conclusion of analysis of a small amount of data which is gathered from the teachers and from the InTERs. After comparing both data manually, I used clustered column charts, pie charts and scatter plots for the analysis. Most of the attributes are object data type, however, chart types were limited based on the data type. I am still not satisfied from the whole data what I collected from the sources because some of information was not accurate and we do not have all data in one place. Students' attendance, and post test data are in different-2 sources, and it is very hard to compare all the data manually, I tried my best to explore insight of data.
- For HSC students: Most of the students have taken TASC test, other students will take the TASC test soon while Ell students have less post tests. Total hours and course hours are positively correlated in both HSC and Ell which means total hours and course hours have the same variation. Cheryl has most HSC/GED students so far from that data. I didn't get data information from Dave's class. If we have that data probably there could be a different result of my analysis.
- Non graduated students' frequency are more than graduated students. I assumed that from the lack of data information, but hopefully those students will graduate soon. Everyone has the same objective at LARA to make the community educated.