Student Representation in the Lee College Service Area

By: Douglas Walcerz, Provost
April 10, 2022

Charge

The Board of Regents charged the administration with conducting an analysis of enrollment disaggregated by age, gender and race/ethnicity to determine if the college is enrolling proportional numbers of students throughout the service area. This report presents the results of the analysis.
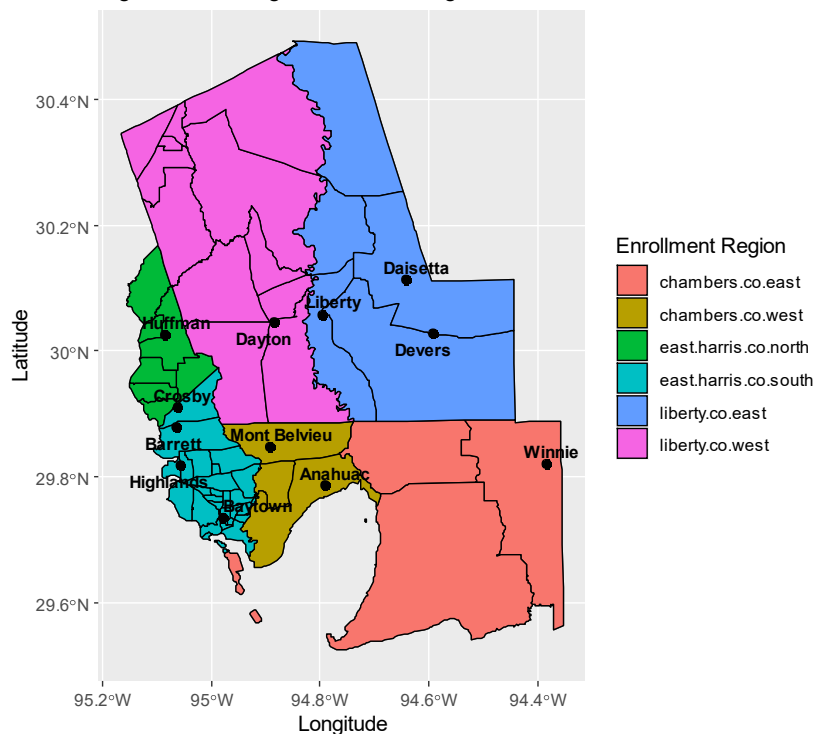
Refinement of the Question

The question we want to answer is this: if we consider a particular gender and race/ethnic group, such as White (non-Hispanic) males, is that group proportionally represented at Lee College? To answer the question, we need to know what percentage of the population are White (non-Hispanic) males, and we need to know what percentage of the Lee College student body are White (non-Hispanic) males. If the two percentages are equal, then White (non-Hispanic) males are proportionally represented at Lee College. Key to this analysis is the identification of the population and the identification of the gender/race/ethnic group within the population.

Populations

The Lee College service area covers 2,000 square miles with a population of about 275,000 and includes urban, suburban and rural communities east of the Houston Ship Channel. Different parts of the service area have very different demographics; for example, Baytown is about 47% Hispanic while Winnie is less than 5% Hispanic. If we simply analyzed the service area as a single population we would probably lose important information about representation in different parts of our service area. Therefore, the service area is divided into six enrollment regions that are relatively homogeneous as shown in Fig 1.

Fig. 1: Lee College Enrollment Regions

Within each enrollment region there are people of all ages, but we are most interested in people age 18 to 24 and age 25 to 34 because they account for almost 90% of our enrollment. We divide the overall population into six regions and two ages, thus we have twelve populations:

1. Age 18-24 in East Harris County South
2. Age 18-24 in East Harris County North
3. Age 18-24 in Chambers County West
4. Age 18-24 in Chambers County East
5. Age 18-24 in Liberty County West
6. Age 18-24 in Liberty County East
7. Age 25-34 in East Harris County South
8. Age 25-34 in East Harris County North
9. Age 25-34 in Chambers County West
10. Age 25-34 in Chambers County East
11. Age 25-34 in Liberty County West
12. Age 25-34 in Liberty County East

Groups Within Each Population

In this analysis we examine two genders: male and female, and three race/ethnic groups: White, Black and Hispanic, so we have six groups:

1. White (non-Hispanic) males
2. White (non-Hispanic) females
3. Black (non-Hispanic) males
4. Black (non-Hispanic) females
5. Hispanic males
6. Hispanic females

For each group, we need to know their percentage in each of the twelve populations and their percentage in the Lee College student body to see if each group is proportionally represented. For example:

- Assume we are looking at the population defined as: Age 18-24 in East Harris County South
- Assume we are looking at the group defined as: White (non-Hispanic) males
- We calculate that 9% of the population belong to the group. In other words, 9% of the 18-24-year-olds living in the East Harris County South region are White (non-Hispanic) males.
- We calculate that 8% of the Lee College student body belong to the group. In other words, 8% of the 18-24-year-olds living in the East Harris County South region and attending Lee College are White (non-Hispanic) males.
- We conclude that White (non-Hispanic) males in the 18-24 age group and living in the East Harris County South region are slightly underrepresented.

Results

Results are shown in Table 1, which has twelve dark-bordered boxes. Each box represents one of the twelve populations previously defined. Within each dark-bordered box are six white cells corresponding to the six gender/race/ethnic groups within each population. Within each white cell are two percentages. The first is the percentage of the gender/race/ethnic group in the given population, and the second is the percentage of the gender/race/ethnic group enrolled at Lee College in the given population. We do not report percentages when the number enrolled at Lee College is less than ten.

Table 1: Representation of 18-to-34-year-olds at Lee College by Gender and Race/Ethnicity					
Within each cell are two percentages. The first is the percentage of the age group in the population; the second is the percentage of the age group that is enrolled at Lee College.					
Region: East Harris County South (pop: 110,350; enrl: 2,942)					
	Age 18-24 (pop: 10,460; enrl: 1,835)			Age 25-34 (pop: 15,433; enrl: 724)		
	White pop: 2,275 enrl: 330	Black pop: 1,783 enrl: 220	Hispanic pop: 6,258 enrl: 1,232	White pop: 4,343 enrl: 174	Black pop: 2,458 enrl: 144	Hispanic pop: 8,448 enrl: 393
Male	9% / 8%	9% / 5%	30% / 30%	14% / 9%	8% / 6%	28% / 18%
Female	12% / 10%	8% / 7%	30% / 37%	14% / 15%	8% / 14%	27% / 36%

Region: East Harris County North (pop: 35,389; enrl: 158)

	Age 18-24 (pop: 3,396; enrl: 105)			Age 25-34 (pop: 4,734 ; enrl: 32)		
	White pop: 1,741 enrl: 75	Black pop: 133 enrl: 4	Hispanic pop: 1,522 enrl: 25	White pop: 3,370 enrl: 24	Black pop: 105 enrl: 1	Hispanic pop: 1,187 enrl: 6
Male	31% / 37%	NA	15% / 13%	34% / 31%	NA	NA
Female	21% / 34%	NA	30% / 10%	38% / 44%	NA	NA

Region: Chambers County West (pop: 29,866; enrl: 707)

	Age 18-24 (pop: 2,382; enrl: 502)			Age 25-34 (pop: 4,002; enrl: 128)		
	White pop: 1,266 enrl: 271	Black pop: 150 enrl: 33	Hispanic pop: 957 enrl: 188	White pop: 2,435 enrl: 67	Black pop: 409 enrl: 12	Hispanic pop: 1,125 enrl: 47
Male	29% / 29%	0% / 2%	27% / 18%	27% / 23%	7% / 2%	11% / 13%
Female	24% / 25%	6% / 5%	13% / 19%	34% / 29%	3% / 7%	17% / 24%

Region: Chambers County East (pop: 12,705; enrl: 88)

	Age 18-24 (pop: 1,196; enrl: 59)			Age 25-34 (pop: 1,545; enrl: 18)		
	White pop: 826 enrl: 28	Black pop: 110 enrl: 4	Hispanic pop: 260 enrl: 26	White pop: 1,068 enrl: 12	Black pop: 27 enrl: 3	Hispanic pop: 435 enrl: 3
Male	32% / 24%	NA	10% / 19%	37% / 17%	NA	NA
Female	37% / 24%	NA	12% / 25%	32% / 50%	NA	NA

Region: Liberty County West (pop: 64,263; enrl: 261)

	Age 18-24 (pop: 5,780; enrl: 177)			Age 25-34 (pop: 9,702; enrl: 59)		
	White pop: 3,123 enrl: 87	Black pop: 471 enrl: 8	Hispanic pop: 2,151 enrl: 82	White pop: 5,330 enrl: 37	Black pop: 1,136 enrl: 6	Hispanic pop: 3,128 enrl: 15
Male	26% / 25%	NA	20% / 22%	27% / 25%	NA	16% / 8%
Female	28% / 24%	NA	18% / 25%	28% / 37%	NA	17% / 17%

Region: Liberty County East (pop: 21,910; enrl: 156)

	Age 18-24 (pop: 1,976; enrl: 107)			Age 25-34 (pop: 2,751; enrl: 34)		
	White pop: 923 enrl: 56	Black pop: 469 enrl: 13	Hispanic pop: 546 enrl: 38	White pop: 1637 enrl: 21	Black pop: 551 enrl: 5	Hispanic pop: 418 enrl: 8
Male	27% / 29%	15% / 6%	13% / 16%	29% / 15%	NA	NA
Female	19% / 23%	9% / 7%	15% / 20%	31% / 47%	NA	NA

Discussion

About 90% of the students who live in our service area come from three enrollment regions: East Harris County South, Chambers County West and Liberty County West. Looking primarily at these three enrollment regions, we see that Age 18-24 Hispanic females are uniformly overrepresented, and Age 25-34 Hispanic females are overrepresented in two regions and evenly represented in the third. Age 18-24 Hispanic males are evenly represented in one region, underrepresented in another, and slightly overrepresented in the third, and Age 25-34 Hispanic males are underrepresented in two regions and slightly overrepresented in the third.

Age 18-24 Black females are very slightly underrepresented in two regions and have insufficient population in the third to estimate representation. Age 25-34 Black females overrepresented in two regions and have insufficient population in the third to estimate representation. Age 18-24 Black males underrepresented in one region, slightly overrepresented in another, and have insufficient population in the third to estimate representation. Age 25-34 Black males are underrepresented in two regions and have insufficient population in the third to estimate representation.

Age 18-24 White females are underrepresented in one region and almost evenly represented in two regions, and Age 25-34 White females are overrepresented in one region evenly represented in another, and underrepresented in the third. Age 18-24 White males are evenly represented in all three regions, and Age 25-34 White males are underrepresented in all three regions.

Methodological Notes

Enrollment Data

Student enrollment data were extracted from PeopleSoft for the fall and spring semesters from three academic years: 2019-2020, 2020-2021 & 2021-2022. Data for the three years was combined and then divided by three to produce a three-year average. The use of a three-year average should reduce year-to-year variation, especially since the last two years have been impacted by the COVID-19 pandemic.

Enrollment data included gender, race/ethnicity date of birth, and address. Date of birth was converted to age by calculating the number of years from the date of birth to the first day of the relevant semester. Address was converted to longitude and latitude using an online commercial application[1]. The boundaries of each enrollment region were obtained from the Bureau[2] as a set of points (longitude and latitude) defining the vertices of a polygon. The determination of which region each address belonged to was made using a winding number algorithm[3]. Thus, for each student, we identified gender, race/ethnicity, age, and enrollment region. Using this data, we determined the average number of enrolled students from each region disaggregated by age, gender and race/ethnicity.

US Census Data

We used five-year averages from the 2016-2020 American Community Survey. The US Census Bureau provides estimates of the number of people in each census tract by age, gender and race/ethnicity; however, the categories used by the Bureau are not exactly the same as the categories that are used in this analysis. The following paragraphs explain how we convert the data for use in this analysis

Age Categories

This analysis divides students into two age categories: Age 18-24, which accounts for about 65% of total enrollment, and Age 25-34, which accounts for about 23% of total enrollment. Together, the two groups account for 88% of total enrollment. Other age groups have small population, and when those small populations are disaggregated by gender and race/ethnicity they create very small groups that are difficult to interpret because random changes of one or two students can produce large percentage changes.

The age categories used by the Bureau are shown below. We combined the categories for "18-19 years" and "20-24 years" to get the "Age 18-24" category used in this analysis. The "25-34 years" category was used as provided and did not require any conversion.

Table 2: Age Categories Used by the US Census Bureau				
Under 5 years	15-17 years	25-29 years	45-54 years	75-84 years
5-9 years	18-19 years	30-34 years	55-64 years	85+ years
10-14 years	20-24 years	35-44 years	65-74 years	

---

[1] Address verification and geocoding: https://www.smarty.com/
[2] Shapes of census tracts: https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.html
[3] Winding number algorithm: https://en.wikipedia.org/wiki/Point_in_polygon

This analysis divides students into two gender categories: Male and Female. The gender categories used by the Bureau are also "male" and "female" so no conversion was needed.

This analysis divides students into six race/ethnicity categories following the convention used by the Integrated Postsecondary Education Data System (IPEDS):

1. White (non-Hispanic)
2. Black / African American (non-Hispanic)
3. American Indian and Alaska Native (non-Hispanic)
4. Asian (non-Hispanic)
5. Native Hawaiian and Other Pacific Islander (non-Hispanic)
6. Hispanic / Latino

The race categories used by the Bureau (Table 3) differ from the categories used by IPEDS.

Table 3: Race/Ethnicity Categories Used by the US Census Bureau	
A: White Alone	F: Some Other Race Alone
B: Black / African American Alone	G: Two or More Races
C: American Indian and Alaska Native Alone	H: White Alone, Not Hispanic or Latino
D: Asian Alone	I: Hispanic or Latino Population
E: Native Hawaiian and Other Pacific Islander Alone	

Race is complicated. The Bureau allows a person to select up to six race categories and also indicate if they are Hispanic/Latino or not. The first six categories used by the Bureau (A-F) are for people who indicate only one race category, but these six categories include both Hispanic/Latino and non-Hispanic/Latino people. The seventh category (G) is for people who pick more than one of the six categories, and also includes both Hispanic/Latino and non-Hispanic/Latino people. The seventh category (H) is a the non-Hispanic/Latino subset of category (A). The last category is the number of people who indicate they are Hispanic/Latino, and they can come from any of the first seven categories.

The IPEDS-based race/ethnicity categories used in this analysis are derived from Bureau categories as follows:

1. IPEDS White (Non-Hispanic) is the same as Bureau Category H.
2. IPEDS Black / African American (non-Hispanic) is determined by estimating the number of Black Hispanic people and subtracting it from the total number of Black people.
   a. The number of White Hispanic people is Category A minus Category H.
   b. The number of non-White Hispanic people is Category I minus the number of White Hispanic people.
   c. We assume that all people indicating "Some Other Race" are Hispanic/Latino. The Bureau notes that 25% of all Hispanic/Latino people select this race category, and only a fraction of a percent of non-Hispanic/Latino people select it.
   d. We assume that all people indicating "Two or More Races" are Hispanic/Latino.
   e. We assume that none of the people indicating American Indian and Alaskan Native or Asian or Native Hawaiian and Other Pacific Islander are Hispanic/Latino.
   f. Based on our assumptions, the number of non-White Hispanic people is composed of three groups: Black Hispanics, Some Other Race Hispanics, and Two or More Race Hispanics. We subtract Column F and Column G from the number of non-White Hispanics to find the number of Black Hispanics.
   g. We subtract the number of Black Hispanics from the number of Black/African American people to get the number of Black/African American (non-Hispanic) people.
3. American Indian and Alaska Native (non-Hispanic) is the same as Category C based on our assumption that none are Hispanic.
4. Asian non-Hispanic is the same as Category D based on our assumption that none are Hispanic.

5. Native Hawaiian and Other Pacific Islander non-Hispanic is the same as Category E based on our assumption that none are Hispanic.
6. Hispanic / Latino is the same as Category I.

Margin of Error

Populations in our service area are relatively small from the perspective of the US Census Bureau, so margins of error in their population estimates are relatively large, sometimes +-50% or more.

End of Notes