

*GIS Analysis of Potential Factors that Influence Students' Selection Over Universities in California*



*Yue Qin*

*PPD 631 Geographic Information System*

*Professor Barry Waite & Bonnie Shrewsbury*

*Proofread by Yangguang Yao*

*4/20/2020*

## Introduction

Benjamin Franklin had described education as an investment that pays the best interest, and nowadays, most of the countries in the world put education in the priority position. For students and their parents, they will face a difficult selection among multiple Universities sooner or later. Choosing a college is about more than the name on the diploma. Where a student goes to school touches numerous aspects of his or her life, from academic studies to social activities and beyond. (Katy, Josh) So, a single factor such as national ranking or geographic location will no longer be the only conclusive factor while choosing the right university. Then what factors may influence students' decisions over different universities? And what are the relationships between students and their decisions over universities?

Based on author Seeta's survey among 62,366 students from 61 universities around the world, she found out that the top 5 factors that affect students' decisions were high-quality teaching, scholarships, university ranking, rate of international student and prestigious brand. (Seeta) The method used in this project will be different from Seeta's survey and the research will be only limited in California state. Four universities are selected in this project, including USC in Los Angeles County, UC Berkeley in Alameda County, UC San Diego in San Diego County and UC Irvine in Orange County. Then four potential factors that will be analyzed are population distribution in different County, Public transportation in different County, the unemployment rate for bachelor's degree or higher in different County, and crime rate in the university area. By analyzing and comparing those factors with the number of students applying for those four universities, we will successfully answer the questions put forward previously. And also, I will score each factor at the end of the analysis and then add them up to see which university become the winner under the selected factors.

## Data Collection

In order to illustrate and analyze the issues we have put forward above, there are five types of data we need to collect. First, the number of applicants to those four universities in 2018. Based on *The classroom database*, there were 64,000 applicants applied for USC in Los Angeles County, 88,466 applied for US San Diego, 85,092 applied for UC Irvine, and 85,044 applied for UC Berkeley. Second, population density in those four Counties, third, Public transportation stops in those four Counties, fourth, unemployment rate for bachelor's degree or higher in those four counties and the number of offenses in those four universities' areas.

*Number of applicants:*

Colleges with the most applicants (2018), Retrieved April 1, 2020 from The classroom database.

<https://www.theclassroom.com/colleges-with-the-most-applications-data-study.html>

*Population distribution:*

California Open Data Portal (2018), Retrieved March 29, 2020 from CA Geographic Boundaries.

<https://data.ca.gov/dataset/ca-geographic-boundaries>

*Public transportation stops:*

Bus Stops Zip file (2018). Metro GIS Database, BusStops1218 of Los Angeles County. Retrieved March 29,2020

Bus Stops Zip file (2019), The Orange County Transportation Authority (OCTA), Bus Stops Oct 2019 of Orange County, Retrieved March 29,2020

Bus Stops Zip file (2012), Alameda-Contra Costa Transit District, Bus Stops March 2012 of Alameda County, Retrieved March 29,2020

Bus Stops Zip file (2018), GTFS Transit, Public transit route of San Diego Sep 2018, Retrieved March 29,2020

*Unemployment rate for bachelor's degree or higher:*

US Census Bureau, Retrieved March 29, 2020 from Employment Status 2013-2017 American Community Survey 5 year estimates

[https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS\\_17\\_5YR\\_S2301&prodType=table](https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_17_5YR_S2301&prodType=table)

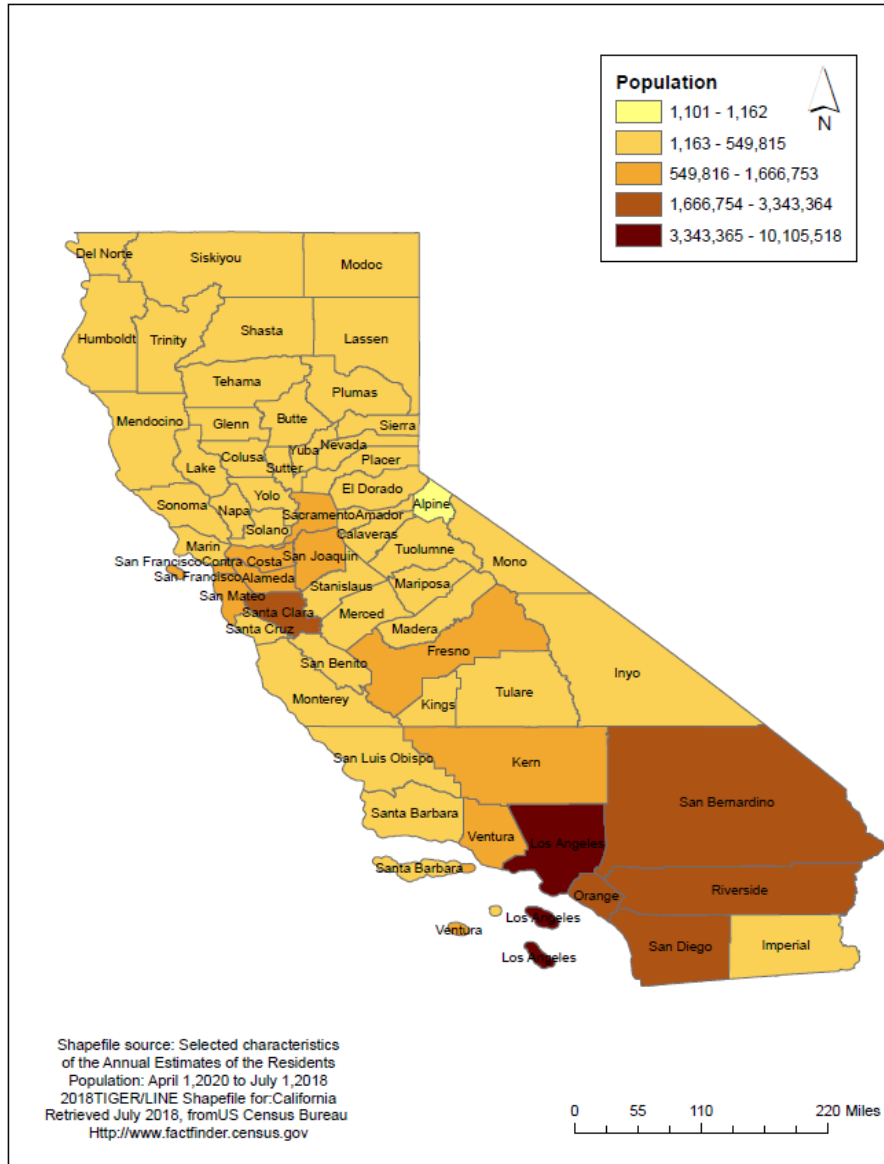
*The number of offenses in those four universities' area:*

U.S Department of Education, Retrieved March 29, 2020 from CSS Campus Safety and Security.

<https://ope.ed.gov/campussafety/#/>

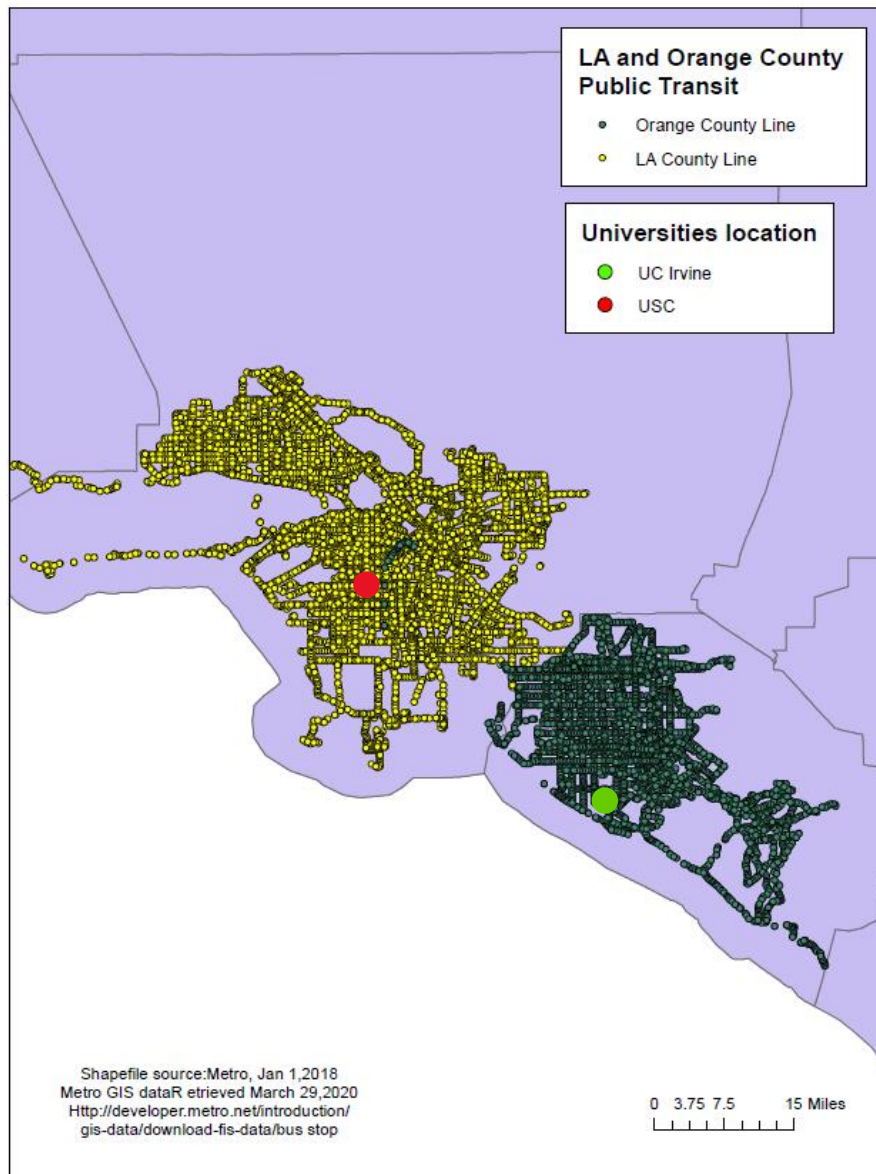
## Layout Distribution and Analysis

Figure 1. The population distribution by county in California



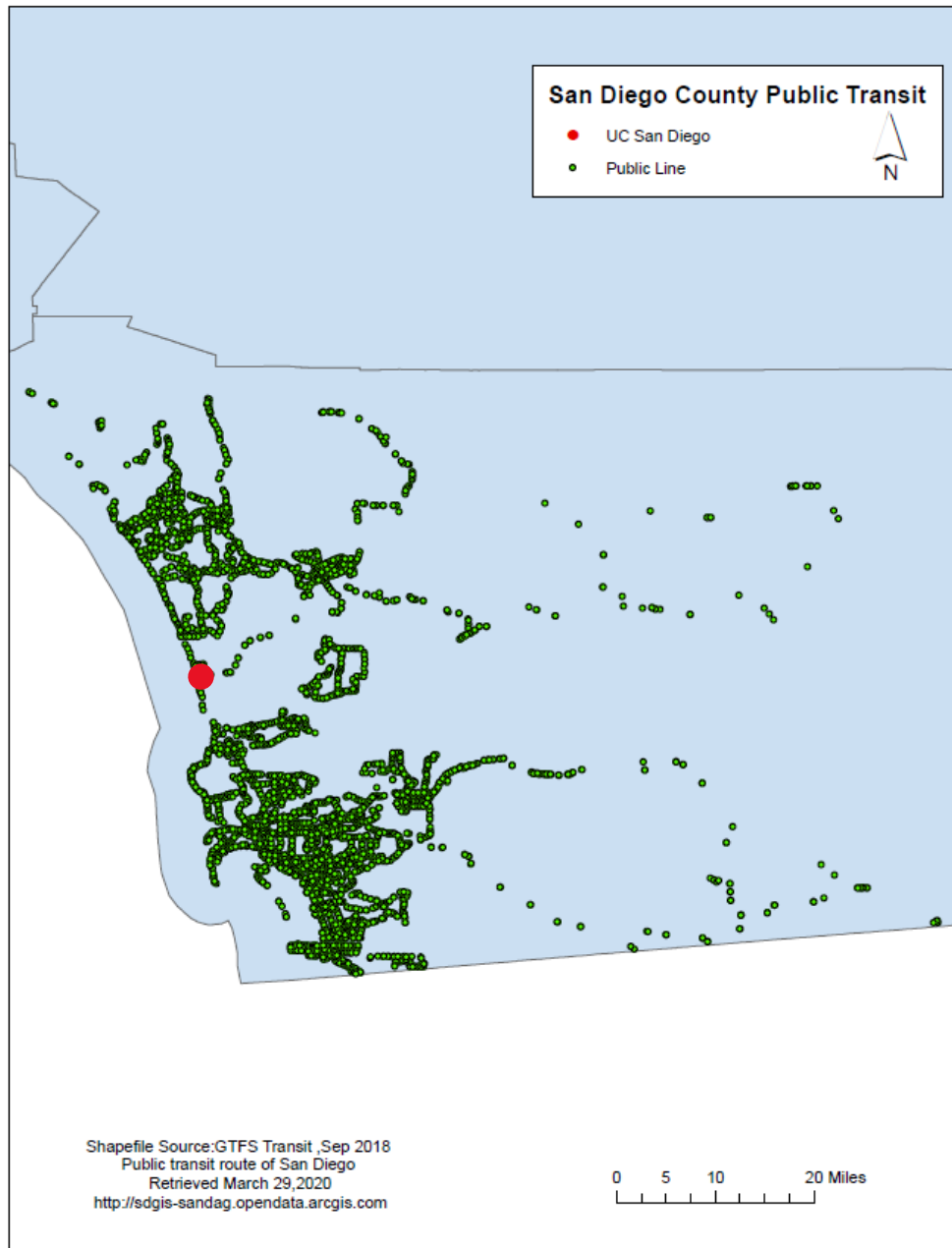
According to the data we retrieved from California Open Data Portal (2019) and the maps shown above, the deeper color represents the larger population. So, Los Angeles County has the deepest color, and it matches the actual data that the LA county has the largest number of populations, which is 10,105,518. The second largest number of populations belongs to San Diego county of 3,343,364. The third largest number of populations is Orange county, which has a 3,185,968 population. Then, Alameda county has the smallest number of populations, and it is 1,666,753. We will analysis the population distribution to see how the population in each county influence future students' decision over universities.

Figure 2. Public transportation stops in Los Angeles and Orange County



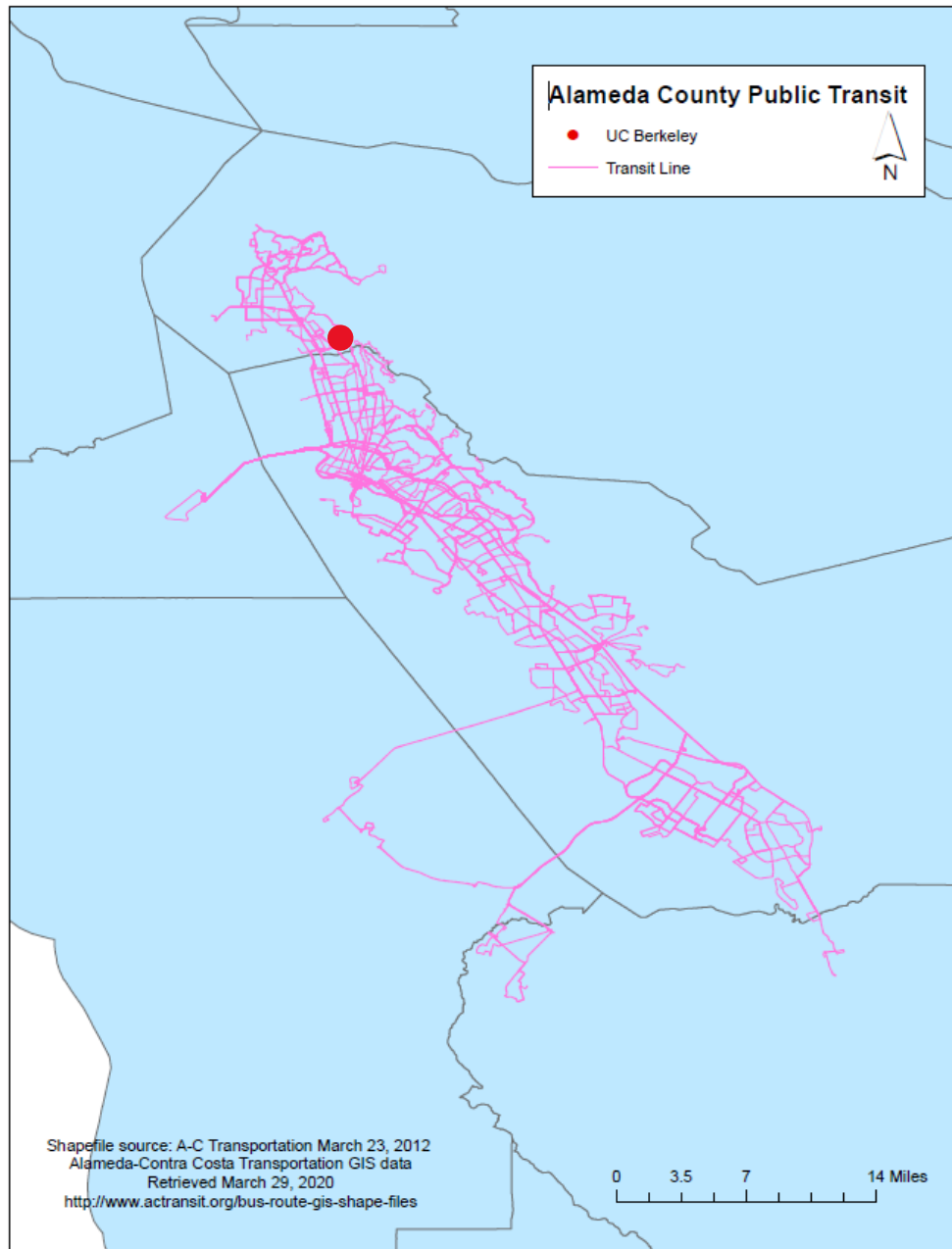
Please see the map of Los Angeles and Orange Counties' public transit stops above. The stops in LA county were represented by deep blue points and stops in Orange County were represented by purple points. According to the data from Metro GIS Database and The Orange County Transportation Authority (OCTA), Los Angeles County and Orange County have largest and third largest number of public transportations stops, which are 19,066 and 5,479. The analysis of public transportation will let us understand if the convenience of public transportation will influence future students' university selection.

Figure 3. Public transportation stops in San Diego County



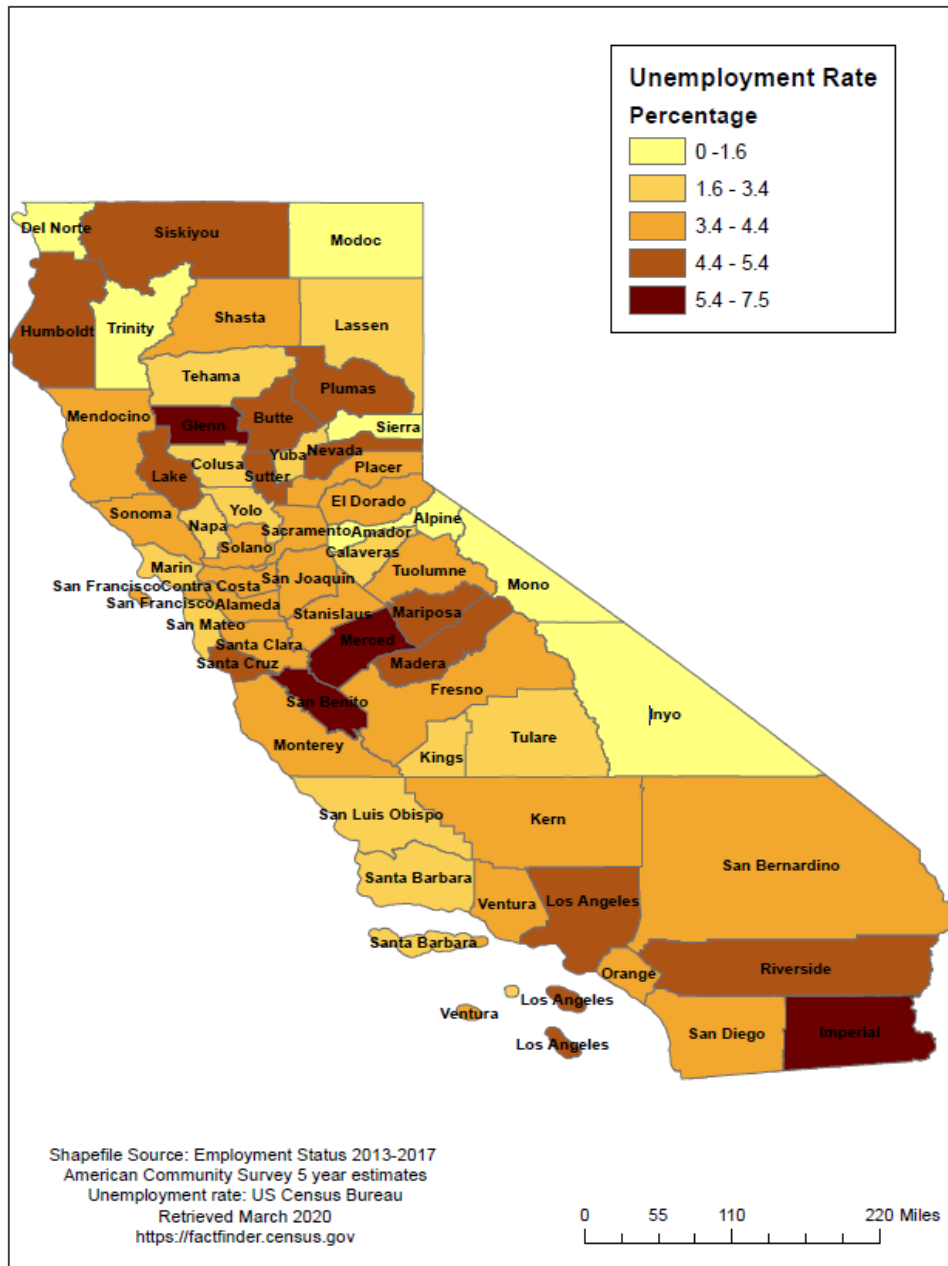
The map above shows the public transportation stops in San Diego County, and the actual stops were represented by red points. According to GTFS Transit, Public transit route of San Diego, there are totally 6,217 public transportation stops in San Diego County, and it has second largest number of public transportations stops in all four countries we studied.

Figure 4. Public transportation stops in Alameda County



The map above shows the public transportation line in Alameda County, which is located in Northern California. According to the data from Alameda-Contra Costa Transit District, there are a total of only 173 public transportation stops in Alameda County, which is the smallest number among all four counties we studied.

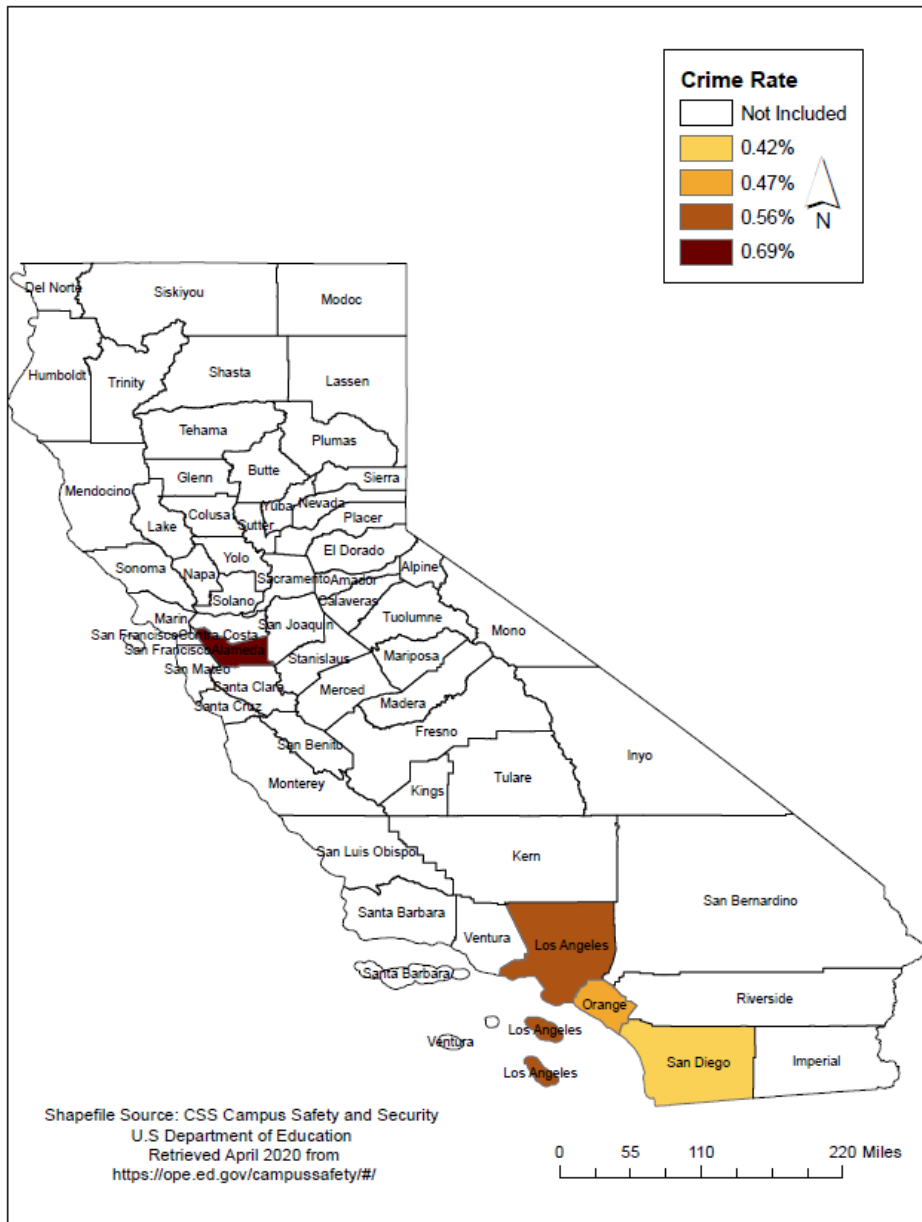
Figure 5. The distribution of unemployment rate for bachelor’s degree or higher in California



The fifth map showed the distribution of unemployment rate for bachelor’s degree or higher in California, we analysis this type of unemployment to see if well-educated students can easily find jobs in those four counties we studied or not. This factor may influence future students’ decisions while they do not only consider their student career but also their jobs. According to the data we retrieved from US Census database, Los Angeles County has the largest unemployment rate for people with bachelor’s degree and higher, which is 4.7 percent, then San Diego county has the second largest index of 4.2 percent, and Orange County and Alameda County has the smallest rate of 3.8 percent.



Figure 6. Criminal offense near University area

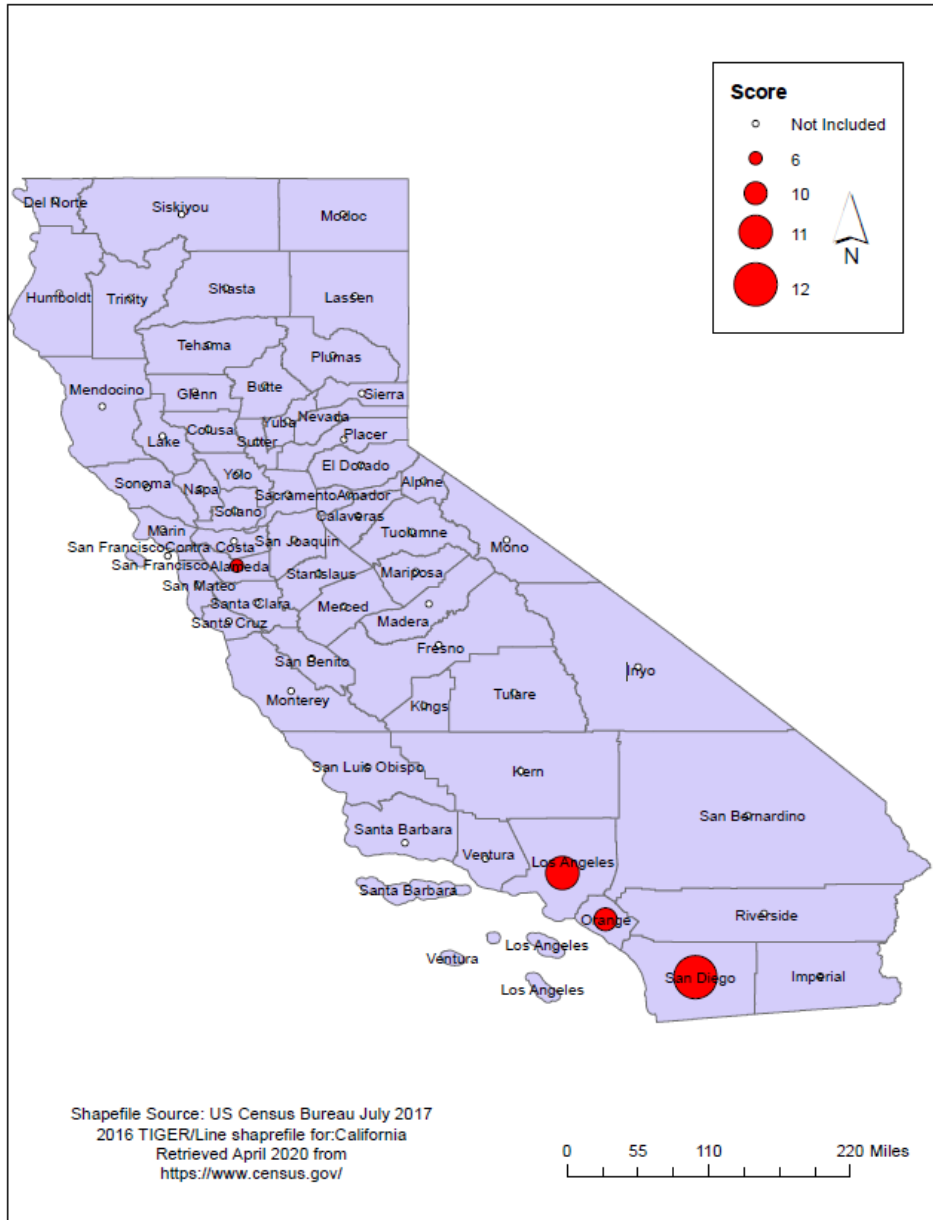


According to the data we retrieved from the U.S Department of Education, CSS Campus Safety and Security, the map above shows the crime rate including robbery, sexual assault and some other types. Since we only study the four counties where targeted universities locate, the rest of the Counties in California are not included in the map and they will be blank. So, UC Berkeley, represented by Alameda County in the map has the largest crime rate, which has a crime rate of 690 per 100,000 population in the year of 2018. Then the area of USC represented by Los Angeles County has the second largest rate of 560 per 100,000 population. UC Irvine represented by Orange County demonstrated crime rate of 470 per 100,000 in 2018. UC San

Diego represented by San Diego County on the map has the smallest crime rate, which is 420 per 100,000 in 2018. The crime rate is calculated by divide the crime number by the student number enrolled in 2018. It's very hard to accurately calculate the entire number of students, faculties and residents within the university area, so the best choice is to only include enrolled students since students are the main target of criminal offenses.

**Score earned**

Figure 7. Score earned by sample universities



The rule to score those four sample universities is that, the university which performances best in each factor will earn 4 points, the second will earn 3 points, and so on. So, based on our

analysis, UC San Diego earns 3 points in both population distribution and public transportation, then 2 points in unemployment rate, and 4 points in crime rate, which means UCSD has totally 12 points. USC has the best population distribution and public transportation, and earn 4 points each category, then 1 point in unemployment rate and 2 points in crime rate, therefore UCS gains 10 points in total. For UC Irvine, there are 2 points for both population distribution and public transportation, and 3 points for both unemployment rate and crime rate, so totally 10 points been earned. UC Berkeley wins 1 point for three categories and 3 points for unemployment rate, therefore totally only 6 points.

Therefore, after score calculation, UC San Diego becomes the winner by using the factors I selected, the second best is USC, then UC Irvine and UC Berkeley.

## **Conclusion**

Based on the GIS analysis shown above, the number of university applicants shows a roughly positive relationship with population distribution and the number of public transportations stops. The larger population in each county, the more students would apply for the university located in that county. The same theory for public transportation, the county with advanced public transportation network would attract more students. But the only exception is Los Angeles County in our study, which does not follow the rule. Los Angeles County has the largest population and public transportation stops compare with the rest of the three sample counties, however, USC has the smallest number of applicants in 2018. I would explain this unusual phenomenon caused by multiple other factors, such as competition from other universities, since we know LA County has the largest number of 230 colleges and universities. Or this situation could be caused by air pollution or a serious traffic jam.

However, the number of university applicants demonstrates a roughly negative relationship with the unemployment rate for people with bachelor's degree or higher based on our GIS analysis. The analysis result shows that Los Angeles County has the highest rate of unemployment, which matches the smallest number of applicants. This negative relationship also followed by UC Irvine and UC Berkeley, because Orange County and Alameda county have slightly lower unemployment rate, and the number of applicants is rising in response. However, UC San Diego shows different result, as San Diego County has second highest unemployment rate, but UC San Diego still attracts the largest number of students. One of the reasons to explain that is the acceptance rate of UC San Diego has risen in recent years and it attracts more students with diverse background. Based on the report, UC San Diego saw a 10 percent increase among freshmen and 6 percent rise among transfers, compared to last year. UC San Diego has the second highest number of applicants among the University of California campuses. (Clark) On the other hand, the number of university applicants shows us the absolute negative relationship with our last index, the crime rate near the university area. Every university follows the theory that the more criminal offense reported, the fewer candidates applied.

Finally, the score earned part has already done the math to score all the factors, UC San Diego becomes the winner, and theoretically becomes the most popular university among the four samples. However, we must admit that there are so many other factors that may influence students' decision, if we talk about the degree of significance for the factors analyzed, I would say the public transportation has the greatest influence, even though every factor play important

part in affecting people's mind. And then crime rate plays the second significant role, then population distribution and unemployment rate. Because public transportation affects the population distribution and unemployment rate directly or indirectly, and crime rate is an important index to diagnose if the university is safe or not.

So, we can make a final conclusion based on the GIS analysis that, students would like to select the university located on the county with advanced public transportation network, less criminal offense, large number of population and lower unemployment rate for well-educated people.

### **Data limitation**

When I select data to analyze the crime rate in the university area, it's hard to determine boundary of the university area, some campus have very clear boundaries, but others do not. So, I have to do the difficult decision to determine the boundaries for each university, since the data appears a huge fluctuation between in campus and around campus area. Secondly, the number of criminal offenses for each university may not be accurate, because the database does not collect all types of criminals. Thirdly, as I have already stated previously, the calculation of the entire number of people within university area is extremely hard, because we cannot accurately locate every student, faculty and resident within the area, so the only option is to include enrolled students. Therefore, the crime rate may not be 100 percent accurate.

## **Reference**

Katy Hopkins, Josh Moody, *"10 Steps to Choosing the Right College"*, US News Education, Nov 6, 2019

Seeta Bhardwa, *"Why do students go to university and how do they choose which one?"*, The World University Rankings, June 6, 2017

Christine Clark, *"A Record Number Of Students Apply to UC San Diego for Fall 2018"*, Triton, 15 December 2018

## **Data Reference**

Colleges with the most applicants (2018), Retrieved April 1, 2020 from The classroom database.

California Open Data Portal (2018), Retrieved March 29, 2020 from CA Geographic Boundaries.

Bus Stops Zip file (2018). Metro GIS Database, BusStops1218 of Los Angeles County. Retrieved March 29,2020

Bus Stops Zip file (2019), The Orange County Transportation Authority (OCTA), Bus Stops Oct 2019 of Orange County, Retrieved March 29,2020

Bus Stops Zip file (2012), Alameda-Contra Costa Transit District, Bus Stops March 2012 of Alameda County, Retrieved March 29,2020

Bus Stops Zip file (2018), GTFS Transit, Public transit route of San Diego Sep 2018, Retrieved March 29,2020

US Census Bureau, Retrieved March 29, 2020 from Employment Status 2013-2017 American Community Survey 5 year estimates

U.S Department of Education, Retrieved March 29, 2020 from CSS Campus Safety and Security.