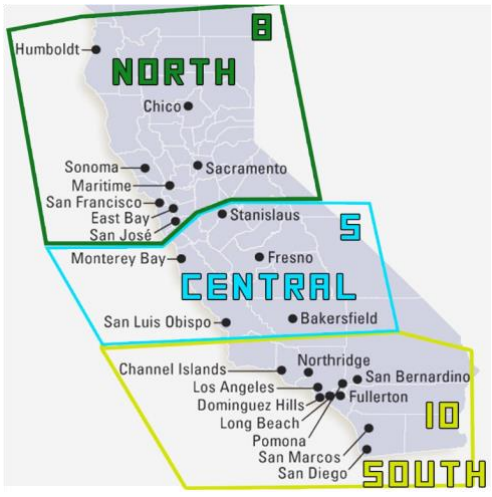
The Civil War of California State Colleges

CONTEXT

As a lifelong native and frequent traveler of all over California, I've noticed one thing about the culture of this great state. It acts as a country of its own and has competing subcultures within the overall California-state culture. It might just be a "California thing" that only natives will understand, but in just about every possible topic of discussion, the comparison of the different regions of California are made. This concept birthed the now-famous competition of "NorCal" vs "SoCal", which has recently grown to include "Central Cal". From sports, to food, to outdoor scenery, there are so many unique aspects these great three regions of California have to offer, which is why the comparisons between them will forever continue to be made.

INTRODUCTION

Today, to connect this perpetual debate into the parameters of this final project, I will be comparing the overall greatness of the central, northern, and southern universities in the California State University system (CSU). When it comes to education, California is among the best, hosting both the UC (University of California) and CSU systems. The CSU system is the larger, yet less prestigious version of the two with 23 campuses all across California. Due to the easy access of their data and as a way to control variation in our study, we will only be studying the 23 CSU schools. To fully understand the regional differences of these CSU schools, I will first examine their desire and appeal to students by looking at their overall yield rate (including first time freshmen and transfers). Then, to follow up, I will study the universities' ability to live up to their name and their degree quality by looking at their most recent 4-year undergraduate graduation rates and most recent graduates' average salary 2 years after graduation. These 3 measures should provide an appropriate scope of the overall success and value of these California State Universities, differing among the 3 great regions of California. Every schools' data used in this report is as of Fall 2020 and for the most recent inaugurating and graduating classes, for each measure respectively. The CSU schools are divided by the regions as shown in Map 1 below:



Map 1. All 23 CSU campuses divided into their respective region

RESEARCH QUESTION 1 ANALYSIS

Which region of California has the most irresistible schools?

To answer this question, we will compare the yield rate between the CSU schools of the central, north, and south regions of California. Yield rate is the percentage of accepted students who enroll in that college in which they were accepted. The CSU database provides separate numbers of acceptances and enrollments for first-time freshmen and transfers. However, I combined the data so we can assess each schools' overall yield rates among both first-time freshmen and transfers together. Higher yield rates means the school is more desirable and people are less likely to turn down a chance to attend the school. Further, if our goal from examining yield rate is to assess universities' appeal and popularity, we can also simply look at the number of applications, as that too shows desire and popularity. The summary statistics of the different regions of CSU schools' yield rates and more seem to demonstrate the existence of a more-popular, favorite region, as can be found in Table 1 below:

Region	Number of Universities	Mean (%)	SD (%)	Median (%)	IQR	Average Number of Applicants	95% Confidence Interval of Mean
Central	5	23.88	7.66	25.48	6.09	25823.40	(14.37, 33.39)
North	8	18.18	5.72	17.53	9.79	27494.88	(13.40, 22.97)
South	10	24.39	5.48	23.71	6.28	51407.30	(20.47, 28.31)

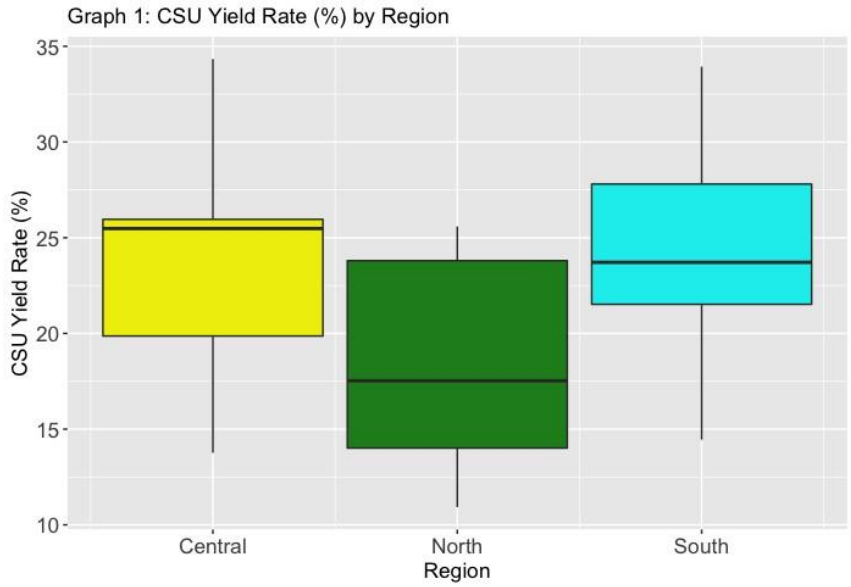
Table 1. Descriptive statistics on CSU schools' yield rate (%), by region

Although it may not be outwardly obvious, these numbers do tell us a lot. Firstly, the south region CSU schools have the highest mean yield rate at 24.39%, much higher than the north's region of only 18.18%. From this, applicants tend to place a higher value and opportunity cost on attending south region CSU schools, certainly over north region schools, and marginally over central region schools. Notice, because of the small sample size of schools in each region, there is lots of variation within each group, as demonstrated by the large standard deviations (SD) and wide confidence intervals among all the regions. Due to these completely overlapping confidence intervals, we cannot conclusively say one region's yield rate is statistically significant from the other. This conclusion is consistent with the ANOVA test that produced a p-value of 0.0989. To even further prove this, I ran a Tukey's multiple comparison test at the 5% significance level, which yielded the letter's report in Output 1 below:

```
        TotalYieldRate groups
south         24.39259      a
central       23.87999      a
north         18.18433      a
```

Output 1. Tukey's multiple comparison HSD test letter report, computed in RStudio

As all three regions have the same letter in the connecting letter's report, it is now very obvious that the three different regions of CSU schools do not have a statistically significant difference in their yield rates. For this very reason, I also included the average number of applicants for each region in Table 1 above, as a naive way to measure popularity and preference. Referring back to Table 1, we can see that the average number of applicants for south region CSU schools (51,407) is almost double that of central region schools (25,823). Across the 10 total CSU campuses in the south region of California, over half a million students applied for enrollment! Considering this all, it is apparent, yet maybe not exactly statistically significant, that the CSU schools in the south region of California are more desirable than those in the central and northern regions. To further explore this notion, we can visually observe the relationship of CSU schools' yield rate by region using a boxplot, as seen in Graph 1 below:



Graph 1. Boxplot of CSU schools' yield rate (%), by region

This graph provides further evidence of our verbal claim made above. As demonstrated by its vertically-highest box, the south region CSU schools are more desirable and, in general, students pass up on the opportunity to attend south region CSU schools the least, among the three regions. The central region of CSU schools appears to be next in ranking of appeal, just below the south region, yet hosting the highest yield rate campus in the entire CSU system with Fresno State (34.33% yield). In dead last place in the CSU ranking of appeal is the north region, which hosts the two lowest yield rate campuses in the entire CSU system, with Humboldt State (10.93% yield) and Sonoma State (11.41% yield). Now that we know which regions of California host the most desirable CSU campuses, let's examine which region of California has the most effective CSU schools.

RESEARCH QUESTION 2 ANALYSIS

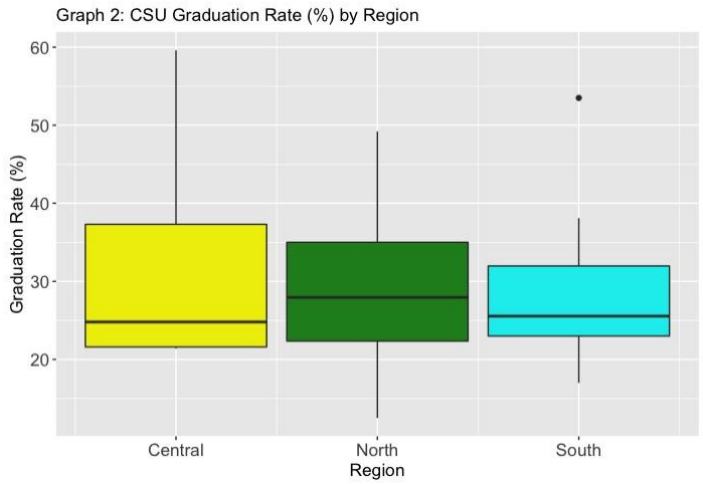
Which region of CSU schools has the most effective schools?

To answer the question above, we will examine the 4-year undergraduate graduation rate among the three regions of CSU colleges. Although every college student may have their own specific reason for attending college, I think most all would agree that their overarching goal is to graduate. Moreso, I think a university's main priority is for every enrolled student to succeed, and complete their degree within 4 years. The main point is that graduation is the main goal of college, so graduation rate is an appropriate measure of how 'good', or effective, a university is at providing their education. Of course, there are many factors to each student's success and eventual graduation. To explore this idea, we can look at a basic numeric summary of undergraduate 4-year graduation rate statistics for each of the CSU regions, as seen in Table 2 below:

Region	Number of Universities	Mean (%)	SD (%)	Median (%)	IQR	95% Confidence Interval of Mean
Central	5	32.94	16.26	24.80	15.70	(12.75, 53.13)
North	8	29.38	11.37	27.95	12.65	(19.87, 38.88)
South	10	28.60	10.90	25.55	8.98	(20.80, 36.40)

Table 2. Descriptive statistics on CSU schools' graduation rate (%), by region

As expected, no region of CSU schools appears immediately outwardly different in terms of graduation rate. While the central region does have the largest mean, it also has the largest standard deviation and smallest sample size. This is important because Cal Poly, San Luis Obispo is a very influential leverage point in the central region data, with a graduation rate of 59.6%. Without San Luis Obispo, the central region's mean graduation rate would be only 26.28%, closer to but dead last among the other regions' current graduation rates. With this considered, along with the fact that all of the 95% confidence intervals are once again entirely overlapping, there is no convincing evidence that there is a statistically significant difference of 4 year undergraduate graduation rates among the three different regions of CSU schools. Once again, this result is consistent with an ANOVA test which yielded a very large p-value of 0.8086. Although the statistical result is the same as in the question 1 analysis above, the least to greatest rankings of the means is actually in a completely different order. To better understand this data and the relationship between region and graduation rate for CSU schools, we can visualize it using a boxplot, as shown in Graph 2 below:



Graph 2. Boxplot of CSU schools' graduation rate (%), by region

Looking at this graph, we can confirm our previously held notion. No region of CSU schools visually appears very different from the other regions. What one region's boxplot may lack in IQR density, it makes up for in range, and vice versa. This visual graphic of Graph 2 is just as inconclusive as the table of data in Table 2, so we will not change our conclusion that the regions are not significantly different. Observing, notice the top whisker for the central region box is very outstretched, representing the data point for Cal Poly, San Luis Obispo. In fact, Cal Poly, San Luis Obispo has the number one highest graduation rate in the entire CSU system at 59.6%, with the south region's San Diego State University coming in at number two, with a 53.5% graduation rate. The San Diego State University data point is the black outlier point on top of the south region box. By this point, we know which CSU universities are the most popular and most 'effective'. Now let's examine the difference in graduates' real life success, among the different regions of CSU schools.

RESEARCH QUESTION 3 ANALYSIS

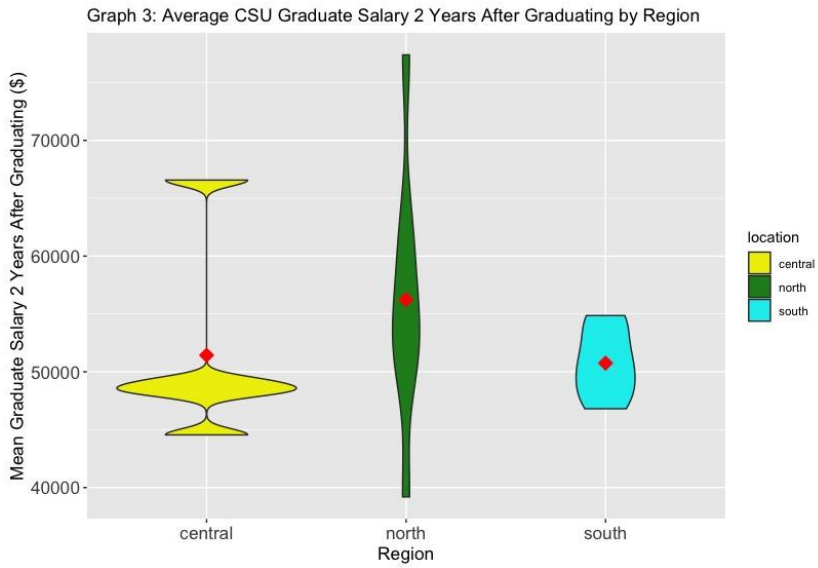
Which region of CSU schools produces the most successful graduates?

Finally, to answer the question above, we will look at the average salary of CSU schools' undergraduate and graduate degree earners, 2 years after completing their degree. In generalized terms, the objective of every college student is to not only earn their degree within 4 years, but then go and find a high-paying job as soon as possible. Of course, success isn't measured just in terms of money – but in this case, it is. If college grads were historically paid the exact same as non-college grads, nobody would attend college. Therefore, your salary after getting your degree is a key part of college and a direct reflection of the institution that educated you. In order to see which region of CSU institutions prepared their graduates for immediate financial success the best, we can first look at the summary statistics of their graduates salary 2 years after completing their degree, as displayed in Table 3 below:

Region	Number of Universities	Mean ($)	SD ($)	Median ($)	IQR	95% Confidence Interval of Mean
Central	5	51434.60	8656.48	48594.00	1088.00	(40686.16, 62183.04)
North	8	56235.62	11037.08	54370.50	8941.00	(47008.39, 65462.86)
South	10	50749.30	2937.85	50686.00	4334.00	(48647.69, 52850.91)

Table 3. Descriptive statistics on CSU schools' graduates' salary ($) 2 years after degree completion, by region

Upon looking at these numbers, most would be surprised by the results. Although we learned from the question 1 analysis that the south region of CSU schools were the most appealing and applied too, they seem to have the worst return in terms of immediate career salary. The 8 CSU campuses in the north region produced the graduates with the highest average salary 2 years after completing their degree, at $56,235.62, while the south region produced the lowest earning average, at $50,749,30. The central region of CSU schools ranked in the middle with an average graduate salary 2 years after degree completion of $51,424.60. However, the central region also had the smallest median salary, at $48,594, due to its small sample size, once again. To better visualize how sample size factors into CSU graduates' average salary 2 year after degree completion by region, we can study the violin plot in Graph 3 below:



Graph 3. Violin plot of CSU schools' graduates' salary ($) 2 years after degree completion, by region

In the graph above, the red studs represent the mean graduate salary value 2 years after degree completion for each region of CSU schools. This graph may lead us to backtrack on our original observations from Table 3 alone. While the south region did indeed have the lowest mean graduate salary value 2 years after degree completion, we can see that the mean is inflated by one very high outlier, while the majority of the data points fall below the mean value. As one would expect from the findings in the question 2 analysis above, this outlier campus is Cal Poly, San Luis

Obispo, achieving a mean graduate salary value 2 years after degree completion of $66,569, the second highest among all CSU schools. Also of note, the north region of CSU schools host the campuses with the highest and lowest mean graduate salary value 2 years after degree completion, with Maritime Academy at $77,390 and Humboldt State at $39,192, respectively. This explains the north region's long, stretched shape in Graph 3 above. Meanwhile, the south has an extremely compact and wide shape surrounding the mean. This is reflected in the fact that the south region's standard deviation is significantly smaller than the others, as seen in Table 3 above. Now, we have a strong understanding of the immediate financial benefits of obtaining a college degree from a California State University school. This data has demonstrated the CSU system's amazing potential and return for students, at all of these great CSU campuses.

CONCLUSION

At this point, it is still not obvious who wins the "Civil War of California Colleges". A big fault in our data was the small sample size, which led to indecisive statistical results, but that was unavoidable with the small set of 23 CSU colleges to begin with. Due to this setback, no region of CSU schools was able to prove itself as a statistically significant better region of academia. However, if we were to reward the regions with simple gold, silver, and bronze awards for each of the three analysis questions, the rankings would go 1) Central 2) North 3) South. Keep in mind though, the Central region contained only 5 campuses, with its data averages largely being inflated by Cal Poly, San Luis Obispo, perhaps the most prestigious of all CSU institutions. So as a Californiana, I am very unsatisfied to say that there is no clear winner or loser among the different regions of CSU colleges – California is equal! While the results discovered may be statistically frustrating due to the small sample sizes, maybe the answers we discovered today are a symbolic and metaphoric message. In an amazing world, country, and state divided more than ever before, where everyone wants to be better than their fellow man, it is refreshing to be reminded of the concept of equality. There is beauty in the belief that no one or no group is intrinsically better than another, but just different in nature. Every group, fit by whatever criteria, is just as equal and important compared to all others. As I stated in the beginning Context section, each one of these great three regions of California have so much to offer. Rather than comparing them, we should simply just enjoy them all while we still can.