

Moving Towards Equity in AP Computer Science

Emily Shepherd

1. Introduction: In Fall 2016, The College Board introduced AP Computer Science Principles. This course was promoted as a computer science course that did not need to be taught by a computer science teacher. Additionally, AP CSP was introduced to increase the number of students studying computer science, and also, increase the number of traditionally underrepresented students studying computer science. By analyzing the test data released by The College Board for the AP CSA test from 2010 to 2019 and the AP CSP test from 2017 to 2019, the following questions will be examined:

- Did the number of AP CSA and AP CSP tests increase from over time?
- Was there a significant jump in AP CSA participation after the introduction of AP CSP?
- Did the participation of traditionally underrepresented populations, such as girls and non-white students, increase with the introduction of AP CSP?
- Are female, Black, and Hispanic students underrepresented in AP Computer Science classes?
- Is there a significant difference in pass rates between white male students and traditionally underrepresented students?
- Do pass rates differ among geographical regions?

2. Initial Exploration of Data: The data available consisted of state level data for AP CSA and AP CSP. Variables included in the datasets included number of tests administered, including desegregated data based on sex and race. The data also included both the number of passing tests and the percent passing, also desegregated by sex and race. The AP CSA data included the data for years 2010 through 2019, and the AP CSP data included data for years 2017 to 2019.

Initially, by looking at the data over time, a pattern can be observed that there is an increase in both number of students taking AP computer science tests and the number of students passing AP computer science tests.



Fig. 1: Plot of Number of CSA tests, 2010-19

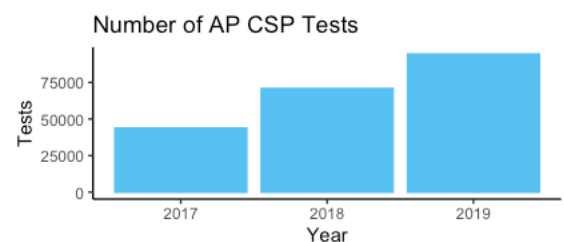


Fig. 2: Plot of Number of CSP tests, 2017-19

From the plots above, we can observe an increase in the number of tests administered for both AP CSA tests and AP CSP tests. However, the increases from in number of AP CSA tests appear to be minimal from 2017 to 2018 and 2018 to 2019, which is after the introduction of the AP CSP tests.

Moving Towards Equity in AP Computer Science

Emily Shepherd

Looking at the number passing, we see there is also a general increase.

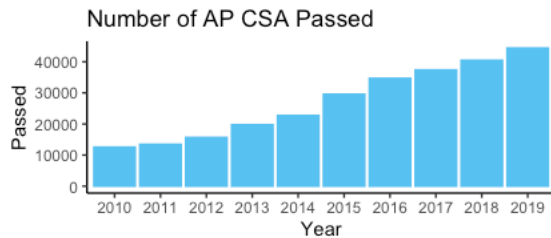


Fig. 3: Plot of Number of Passing CSA tests, 2010-19

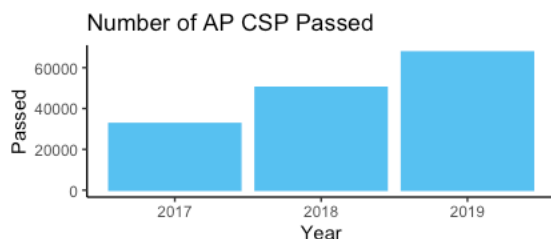


Fig. 4: Plot of Number of Passing CSP tests, 2017-19

The plots for passing tests and total tests look similar for each course. Looking at a plot for the percent passing for each test, we observe that the pass rates stay constant over the years.

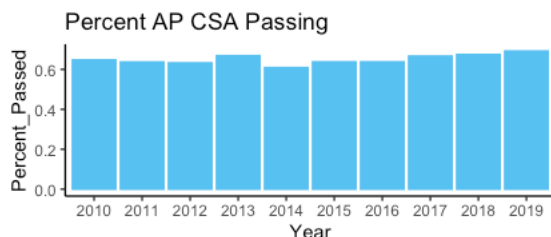


Fig. 5: Plot of AP CSA Passing Rate, 2010-19

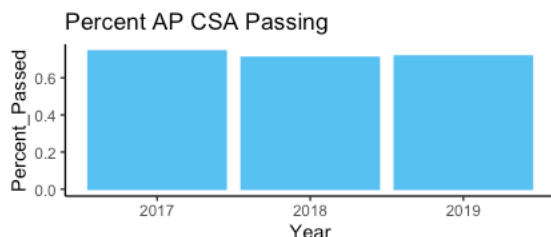


Fig. 6: Plot of AP CSP Passing Rate, 2017-19

The consistency of the pass rates is not that interesting because the advanced placement tests are norm-referenced. The cut scores are determined after the tests

are graded to keep the pass rates relatively constant.

Next, one-way ANOVA analysis was conducted on both the AP CSP total tests and AP CSA total tests, using a factor of years. For the AP CSP test, the p-value equal to 0.09582, indicating there is not a significant association between the number of tests and the year. For the AP CSA test, there was a significant association between year and number of tests, with a p-value equal to 0.00413. However, after completing TukeyHSD, there was no significant interactions among the years 2017-2019, when the AP CSP test was implemented.

3. Analysis of Participation by Sex: Girls are underrepresented in computer science courses and the field of computer science overall. One of the objectives for the introduction of AP Computer Science Principles is to increase the number of girls in computer science courses and eventually in the field of computer science.

For AP CSA, there is a huge difference in the number of boys enrolled and the number of girls enrolled.

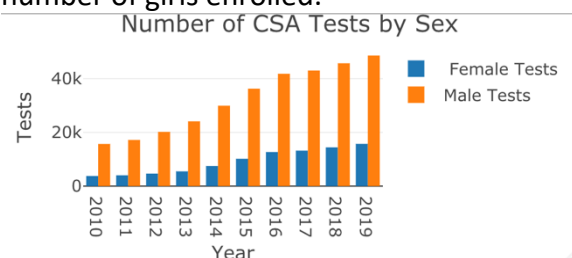


Fig. 7: Grouped Bar Graph of Number of Female and Males AP CSA Tests, 2010-19

For AP CSP, we see a difference still exists in number of females tested and number of males tested. However, the differences do not seem as large.

Moving Towards Equity in AP Computer Science

Emily Shepherd

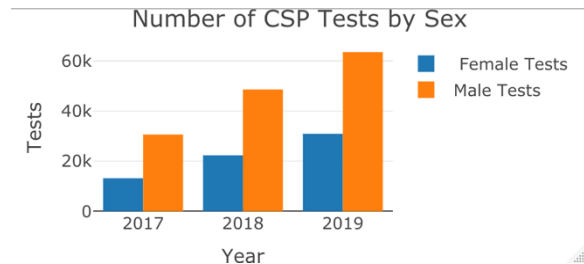


Fig. 8: Grouped Bar Graph of Number of Female and Males AP CSP Tests, 2017-19

It appears from Figure 8, that by 2019, the number of females who take the AP CSP is about half the number of males who take the test. By completing the Chi-Square Goodness of Fit Test, we can test whether the observed participation matches the expected participation based on sex for AP CSP. According to the National Clearinghouse of Educational Statistics, the U.S. high school population consists of 51% male and 49% female.

Table 1 provides the number of observed and expected tests based on sex from 2019.

Sex	Observed	Expected
Female	30907	46254
Male	63489	48142

Table 1: Observed and Expected Distributions for 2019 AP CSP Test Based on Sex

This test resulted in a test statistic $\chi^2 = 9984.52$ and a p-value rounded to 0. This leads us to the conclusion that the observed number tests grouped by sex do not match the expected distribution of tests.

Additionally, we can test to see if the number of female and male students who participated matches the expected for the AP CSA test by 2019.

Sex	Observed	Expected
Female	15699	31457
Male	48498	32740

Table 2: Observed and Expected Distributions for 2019 AP CSA Test Based on Sex

The results of the test, with a degree of freedom of 1, yielded a test statistic

$\chi^2 = 15478.22$ and a p-value approximately equal to 0. This is sufficient evidence to reject the null hypothesis that the expected number of tests match the observed number of tests. This leads to the conclusion that there is significant evidence that the participation in AP CSA is different for males and females, which is not surprising based on figure 7.

4. Analysis of Participation by Race:

Another concern in computer science education is the lack of Black and Hispanic students enrolled in computer science classes and also, eventually pursue careers in technology. I first explored this issue as I did before by plotting the number of CSA and CSP tests over the years for Black, Hispanic, and white students.

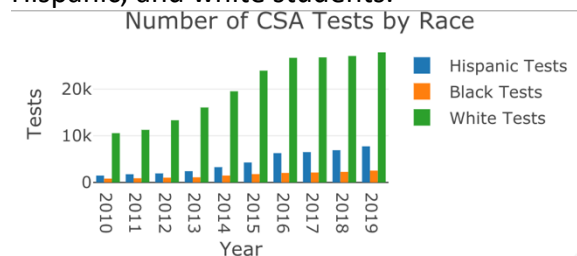


Fig. 9: Grouped Bar Graph of Number of Hispanic, Black, and White AP CSA Tests, 2010-19

From figure above, we can see there was a general increase for all three racial groups through 2017, but then the number of tests leveled off. However, looking at the companion plot for AP CSP tests, we see that the number of tests for all three groups increased from 2017 to 2019.

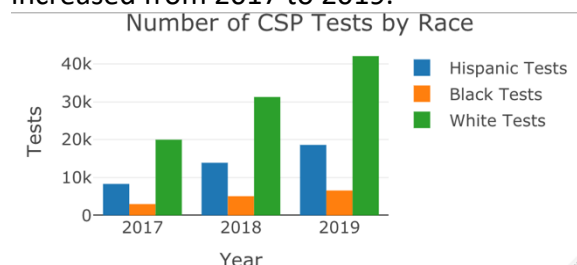


Fig. 10: Grouped Bar Graph of Number of Hispanic, Black, and White AP CSP Tests, 2017-19

Moving Towards Equity in AP Computer Science

Emily Shepherd

According to the Center for Educational Statistics, in 2020, 15% of students enrolled in public schools were Black and 28% were Hispanic and 46% of students were white. Using these percentages as a proportion of 89%, the expected number of AP CSP tests can be calculated for 2019 as seen in Table 3 below.

Race	Observed	Expected
Black	6559	11329
Hispanic	18601	21153
White	42077	34755

Table 3: Observed and Expected Distributions for 2019 AP CSP Test Based on Race

A Chi-Square Goodness of Fit test, with degree of freedom 2, resulted in a test statistic $X^2 = 3858.82$ and a p-value that rounds to 0. This leads us to the conclusion that the observed distribution does not match the expected distribution of AP CSP tests based on race.

Similarly, we can complete the same analysis on 2019 CSA tests.

Race	Observed	Expected
Black	2521	6420
Hispanic	7728	11987
White	27853	19695

Table 4: Observed and Expected Distributions for 2019 AP CSA Test Based on Race

The Chi-Square Goodness of Fit test leads us to a similar conclusion that the observed number of CSA tests by race differs significantly from the expected. The test yields a test statistic $X^2 = 7260.35$ and a p-value that once again rounds to 0.

Based on the above analysis, while the number of tests for Black, Hispanic, and white students have increased, the proportion of tests by race differ significantly from the overall population of high school students, with white students being overrepresented in AP Computer

Science courses and Black and Hispanic students being underrepresented.

5. Analysis of Passing Rates for

Underrepresented Populations: While we have found that participation has increased in both AP CSP and AP CSA for females, Black, and Hispanic students, there is still concern that they are not as successful as their white or male counterparts. Since the AP CSP test is a norm-referenced test, the overall pass rate stays the same. However, pass rates may vary among subpopulations either based on race, sex, or geographic region.

A. Analysis of Pass Rates Based on Gender:

As observed in Figure 11 and Figure 12, for both the AP CSP and AP CSA tests, the pass rate for girls is less than the pass rate for boys, but not by much. Additionally, the difference in pass rates does not seem to be decreasing over time for either test.

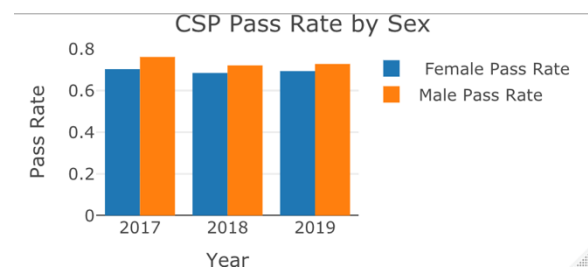


Fig. 11: Grouped Bar Graph of AP CSP Test Pass Rate by Sex, 2017-19

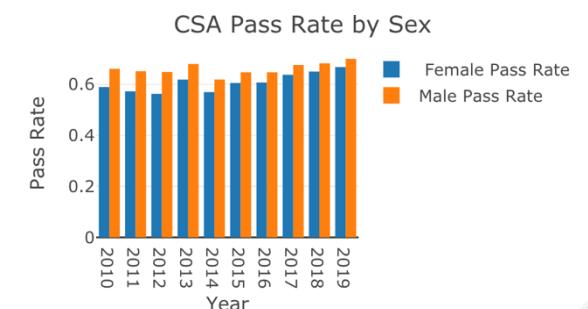


Fig. 12: Grouped Bar Graph of AP CSA Test Pass Rate by Sex, 2017-19

Moving Towards Equity in AP Computer Science

Emily Shepherd

B. Analysis of Pass Rates Based on Race:

Pass rates for the AP CSP test appeared be relatively constant for all racial groups from 2017 to 2019. As observed in Figure 13, the pass rates for Black and Hispanic students were substantially less than their white counterparts.

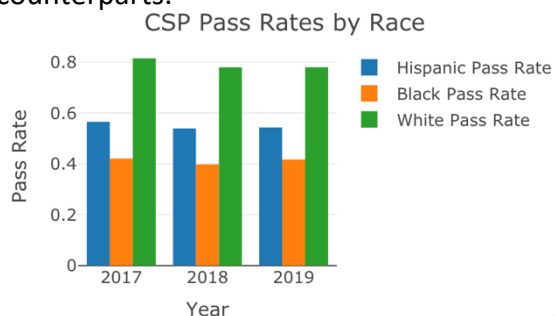


Fig. 13: Grouped Bar Graph of AP CSP Test Pass Rate by Race, 2017-19

Additionally for the AP CSA test, there is a substantial difference between the pass rate for white students and for both Black and Hispanic students, as shown in Figure 14. Additionally, the difference among the groups seems to stay constant over the years, meaning, the achievement gap is not closing.

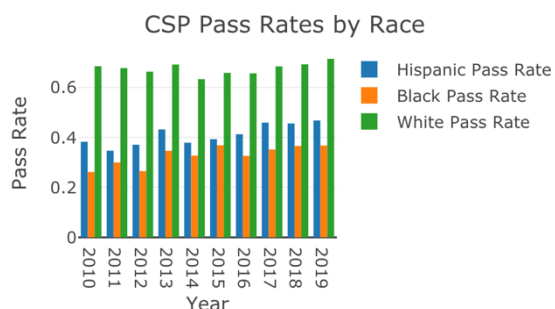


Fig. 14: Grouped Bar Graph of AP CSA Test Pass Rate by Sex, 2017-19

C. Analysis of Pass Rates Based on

Geographical Region: From observing Figure 15, it appears that pass rates are lower throughout the south than other regions of the United States.

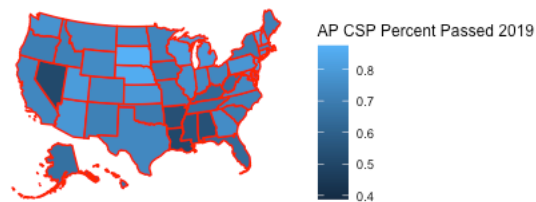


Fig. 15: Percent Passed AP CSP Test by State, 2019

A one-way ANOVA test was run to determine whether regional differences were significant. Each state was categorized into one of four regions, Northeast, North Central, West, and South. With a p-value of 0.000942, we can conclude that there are significant differences in pass rates based on region. The results of the TukeyHSD post-hoc test can be viewed in Table 5. Based on these results, there is a significant difference in pass rates between the South and Northeast regions and also, the South and North Central regions. In both cases the pass rate in the South is significantly less.

Regions	Difference	P-Value
S-NE	-0.0956	0.0219
NC – NE	0.0256	0.8735
W – NE	-0.0299	0.7985
NC – S	0.1212	0.0008
W – S	0.0658	0.1024
W - NC	-0.0554	0.2687

Table 5: TukeyHSD Results 2019 AP CSP Pass Rate with Region as Factor

In the case of the 2019 AP CSA test, the pass rate also seemed to vary by geographical region. In Figure 16, it once again appears that the pass rate may be significantly lower in the South than the other regions.

Moving Towards Equity in AP Computer Science

Emily Shepherd

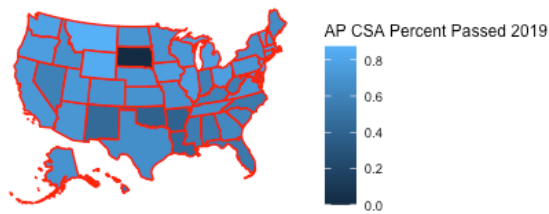


Fig. 16: Percent Passed AP CSA Test by State, 2019

The results of the ANOVA analysis indicated there was significant difference in pass rates among geographic regions, with p -value = 0.0325. The TukeyHSD results are provided in Table 6. However, the post-hoc test did not indicate significant differences in pass rates among any pair of regions for the AP CSA test in 2019.

Regions	Difference	P-Value
S-NE	-0.1413	0.0801
NC – NE	-0.0413	0.9042
W – NE	-0.0045	0.9998
NC – S	0.0100	0.2430
W – S	0.1368	0.0511
W - NC	-0.0368	0.9090

Table 6: TukeyHSD Results 2019 AP CSA Pass Rate with Region as Factor

6. Conclusions: There is both good news and bad news to be gleaned from the analysis. The good news is that more students than ever are enrolling in AP Computer Science courses. This is important for the workforce to keep pace with the needs of the tech sector. Additionally, the achievement gap between female and male students is small and the difference between the number of female students and male students participating in AP Computer Science courses is decreasing.

Unfortunately, the same cannot be concluded when comparing Black students and Hispanic students with white students.

The pass rates for Hispanic and Black students are much lower on both tests than white students. Additionally, the number of Black and Hispanic students enrolled in computer science have not been increasing as quickly as many would like.

Lastly, there seems to be some differences in both access and achievement based on geographical region. However, more analysis would need to be conducted to be able to make solid conclusions.

7. References:

Erickson, Barbara. "Advanced Placement Data – CSA and CSP for 2019." *Computing for Everyone*, 6 Dec. 2019, <https://cs4all.home.blog/2019/11/11/advanced-placement-data-csa-and-csp-for-2019/>.

National Center for Education Statistics. (2022). Racial/Ethnic Enrollment in Public Schools. *Condition of Education*. U.S. Department of Education, Institute of Education Sciences. Retrieved [Nov 27, 2022], from <https://nces.ed.gov/programs/coe/indicator/cge>.