Equal Accessibility to Computer Science Classes in High Schools in California

Jeffrey Miller, Courtney-Lynn Presto
Department of Computer Science
University of Southern California
{jeffrey.miller, cpresto}@usc.edu

Abstract - In this paper, we provide a statistical analysis of data from the California Department of Education to determine whether students of different ethnic and gender backgrounds are equally represented in computer science classes in California. Contrary to public perception, we found that every ethnic demographic was accurately represented in the classroom. The percentage of students of each demographic matched in the classroom the percentage of students in each demographic for the state. We found that this representation does not depend on gender. The ethnic distribution of male students in computer science match the ethnic distribution of the state of California. Similarly, the ethnic distribution of female students in computer science match the ethnic distribution of the state of California. We found that, despite inclusion efforts, female students are still a minority in these classes. Only 15% of schools with computer science programs have enrollment of at least 50% female. Related research suggests that the most effective way to include female students is to integrate STEM programs into the school curriculum at an early age before girls start to associate with gender stereotypes.

I. INTRODUCTION

Recently, schools across the country have been pushing to offer more STEM programs to their students. One of the key subjects included in this initiative is computer science. Information Technology is one of the fastest growing industries in the country in 2017 [1] which explains why schools across the country have rushed to offer computer science classes to their students. Schools have gone so far as to debate whether computer science education should be mandatory in high schools. According to Forbes, mandatory coding classes could exist within the decade since thousands of new teachers are trained on this topic every year [2]. The increase in teachers reflects the increased demand for these classes. In fact, the number of students taking the AP Computer Science exam doubled from 2016 to 2017[3]. This increase in high school students choosing to take AP Computer Science reflects an increase in college students choosing to pursue computer science as a career; there was a 161% increase in students choosing to major in computer science in 2010 to 2015 [4]. And yet, the percentage of women in computer science far from represents the percentage of women in the general population. As of 2014, women composed only 18% [5] of the computer science workforce and only 19% of students who took the AP Computer Science exam in 2011 were women [6]. This proposes several questions: Is the gender disparity a result of lack of accessibility to computer science classes in primary and secondary education? Are female computer science students equally represented within their ethnic demographic?

One way to try to increase the number of women in the field of computer science is to encourage them to code at an early age. For example, Girls Who Code has inspired over 40,000 girls to attend after school programming classes [7]. Girl Develop It is a similar program aimed at providing affordable coding classes to adult women interested in pursuing a career in the tech industry [8]. Even with these programs, the percentage of women in computer science classes in primary and secondary education is much lower than the percentage of women in the schools that offer them. In this paper, we will provide a study of the California Board of Education’s demographic data regarding the demographic breakdown of both ethnicity and gender of students taking computer science classes in primary and secondary schools in California. To gauge equal representation, we
will explore: (1) the percentage of ethnic demographics in the school/district versus the percentage of ethnic demographics represented in computer science; (2) the ethnic distribution of females in computer science classes and the ethnic distribution of males in computer science classes; and (3) the percentage of females in the school versus the percentage of females in computer science classes. Section II offers an overview of existing studies that have conducted similar research. Section III discusses our research methodologies. Finally, Section IV presents our findings, provides a conclusion and plan for future research.

II. RELATED WORK

There have been many studies that have examined the gender and ethnicity gap in the computer science field as it is an epidemic in many colleges nationwide. In 2001-2002, just 7% of computer science bachelor's degrees were conferred to African Americans and Latinos of both genders and 18% to women. In a direct reflection, only 6% of students taking the AP Computer Science exam in 2002 were African Americans or Latinos and only 17% were women [9]. This suggests a direct correlation between gender and ethnicity representation in undergraduate computer science programs and gender and ethnicity representation in high school computer science programs.

Research suggests that girls and students of color lose interest in STEM subjects by middle school [10], the age when students begin to identify with race and gender stereotypes. Because of this, male students choosing a college major are more likely to be exposed to computer science academically or recreationally [11]. Furthermore, according to the survey conducted by Lori Carter of Point Loma Nazarene University, “The proportion of men citing the amount of experience as a factor against choosing a CS major was significantly different from the proportion of women citing this reason, with women seeing it as a deterrent much more often (z=6.245, p<0.0001)[12].” This research experiment concluded that a possible way to increase the amount of women in computer science was to increase girl’s exposure and experience with CS early on.

Other schools have tried to bridge the gender gap in a different way. For example, East River High School, which according to Cohoon is a “predominantly Latino/a high school located in a low-income neighborhood on the eastside of Los Angeles” [9] uses a partnership program to teach students computer science concepts. In all but two cases, the partner set consisted of an older “tech savvy” male student and a younger more amateur female student. The male student adopted a “big brother” role. According to the study, “Despite the girls’ high standing in high-level mathematics classes, the students quickly fell into ‘traditional’ roles, with girls taking on more ‘secretarial’ tasks [9].” Situations like these can cause girls to feel insecure or unempowered and even inhibit their ability to learn. Another interesting point is that girls lack the same technology-oriented social networks that male students take advantage of, for example, video game chat-rooms, technology blogs, and tech magazines labeled “For Boys.”[9] Therefore, girls have less background knowledge of the subject and feel like they are already at a disadvantage. Because of the reinforcement of gender stereotypes, girls feel less inclined to pursue computer science, as exemplified by the low proportion of women to men in computer science classes and jobs.

III. RESEARCH METHODOLOGIES

For this research project, we chose to focus on schools in California since our non-profit STEM education foundation is based in California. We first downloaded relevant data from the California Board of Education website, found at: http://www.cde.ca.gov/ds/sd/df/filesassign.asp. We pulled assignment code data, which is a master document that maps unique class identification codes to information regarding that class. Once downloaded, we wrote a java program to parse the file into a database, exactly mimicking the structure of the file. We used this database table to identify classes with subjects of: “Computer Education” or “Information and Communication Technologies.”
Next, we pulled Class Enrollment documents from 2012-2017. Class Enrollment documents are provided every year. They list every class taught that year in the state of California along with the gender and ethnic distribution of students enrolled in that class. This provided us with a demographic breakdown for every class in every school in California. Similarly, we parsed these files into the MySQL database as well. The database structure exactly mimics the fields of the file.

In order to compare this data with district data, we also obtained a file that provided us with the demographics of enrollment for the school. This can be found at: https://www.cde.ca.gov/ds/sd/sd/filesenr.asp. We pulled enr12-enr16. We parsed this data into the MySQL Database and used this data to compare the gender and ethnicity breakdown of the school/district with the gender and ethnicity breakdown of computer science classes.

We analyzed the school district data to determine whether female students and students of minorities are equally represented in Californian Computer Science (CS) classes. Our hypothesis was that CS classes were comprised overwhelmingly of male students. We suspected that the ethnic composition of students in CS classes were not representative of the ethnic composition of the school offering the class, favoring Asian and White students. We found that our hypothesis was true regarding gender but was false regarding ethnic groups. In general, all ethnicities are well represented in CS classes, but female students are underrepresented in CS classes.

First, let’s examine the gender composition. In general, we found that most CS classes had a higher percentage of male students than female students. Figure 1. shows the distribution of gender in general for all schools in California compared to the distribution of gender in general in all CS classes in California. Female students make up 48.7% of the student population in California for schools that offer CS classes. On the other hand, female students make up only 38% of the population of CS students in California. This is a decrease of 10.7%. On average, significantly more male students are enrolled in CS classes than women.

Additionally, out of the 8476 schools in California, only 1853 of them offer CS classes. Out of those 1853 schools, only 277 (or 15%)
have female students comprising at least half of the CS student population of the school. This means that 85% of schools have male-dominant CS populations. This is another measure of the inequality between female and male students in the CS classroom in California.

Next, let’s examine ethnic representation. Figure 2. depicts the percentage of ethnic groups in general in schools in the state of California compared with the percentage of ethnic groups in CS classes. Groups with percent changes of more than 1% are called out in orange. We can see that White students make up 25.1% of the Californian student population but only 23.2% of the CS student population. This is a decrease of 1.91%. Conversely, Asian students make up 8.9% of the Californian student population but 10.8% of the CS student population. This is an increase of 1.96%. These two groups have the most variation from state population to CS population. All other ethnic groups have a percent change of less than 1% from Californian student population to CS student population. See the following list for variations in ethnic groups with less than a .4% deviation from the Californian student population.

In the groups that were under or over represented (American Indian, Not Reported/No Ethnicity Represented, African American, Asian, White, and Hispanic) the difference between the Californian population and the population of CS students was no greater than 2% which means that the population of CS students is a good sample of the total population of students in California.
Another concern we had was whether gender affected the ethnic representation of students in CS classes. In general, we found that the ethnic distribution of males and females in CS was a good sampling of students in California.

Figure 3 shows the ethnic distribution of all students in California compared with the ethnic distribution of male students in CS classes in California. Groups with percent changes of more than 1% are called out in orange. We see that there are 2.5% less Hispanic males in CS classes in California than there are Hispanic students in all classes in California. Conversely, there are 2.6% more Asian males in CS classes than there are Asian students in all classes in California. These two ethnic groups see the most variation. All other groups observe a difference of .3% or less. See the following list for variations in ethnic groups with less than a .3% deviation from the Californian student population.

<table>
<thead>
<tr>
<th>Percent Increase/Decrease from Californian Student Population to CS Student Population, .3% change or less, MALE</th>
</tr>
</thead>
<tbody>
<tr>
<td>▼.3% Decrease in African American Males</td>
</tr>
<tr>
<td>▼.2% Decrease in White Males</td>
</tr>
<tr>
<td>▲.2% Increase in Not Reported</td>
</tr>
<tr>
<td>▲.1% Increase in Filipino</td>
</tr>
<tr>
<td>▲.1% Increase in Two or More Races</td>
</tr>
<tr>
<td>No change in American Indian</td>
</tr>
<tr>
<td>No change in Pacific Islander</td>
</tr>
</tbody>
</table>

In the groups that were under or over represented (Hispanic, White, Asian, African American, Filipino, Two or More Races, Not Reported) the difference between the Californian population and the population of CS male students was no greater than 2.6% which means that the population of CS male students is a good sample of the total population of students in California.
The female CS student population follows the general student population with slightly more variation than the male population. Figure 4. shows the ethnic distribution of students in California compared with the ethnic distribution of female students in CS classes in California. Groups with percent changes of more than 1% are called out in orange. We see that 25.1% of the student population in California is White. However, only 20.3% of the female student population in CS classes in California is White. This is a decrease of 4.8%. Conversely, Hispanic students make up 52.7% of the total population yet 55.9% of female students in CS classes are Hispanic. This is an increase of 3.2% all other groups have variations of .8% or less. See the following list for variations in ethnic groups with.8% or less deviation from the Californian student population.

In the groups that were under or over represented (White, Hispanic, Asian, African American, Not Reported, Filipino, Two or More Races) the difference between the Californian population and the population of CS students was no greater than 4.8% which means that the population of CS students is a good sample of the total population of students in California.

In conclusion, the gender distribution of students in CS classes in California does not approximate the general student population in California. However, the ethnic distribution of students in CS classes does approximate the general student population in California. Additionally, both genders are ethnically represented in CS classes; the male CS student
population mimics the general Californian student population and the female CS student population mimics the general Californian student population. We recommend increased inclusion efforts targeted toward female students to raise the percentage of female students in CS classes in California.

V. REFERENCES


