A Multivariate Approach to Disaggregating Data

Giovanni Sosa
Craig Hayward
Jared Lessard
Carolyn Holcroft

June 2017
Revised May 2018
www.rpgroup.org
Acknowledgements

We would like to express our gratitude to both the College of Marin and Santa Monica College for offering us clear examples of what it means to disaggregate data in a multivariate context. We would also like to thank Andrew Kretz, Ireri Valenzuela, and the RP Group editors for their insightful feedback and assistance developing this paper.

Finally, many thanks to the Institutional Effectiveness Partnership Initiative (IEPI) and the California Community Colleges Chancellor’s Office for the funding and support critical to making this work possible.
Table of Contents

Executive Summary ............................................................................................................... 4
Introduction .......................................................................................................................... 6
College of Marin's Student Equity Plan .............................................................................. 7
Santa Monica College's Student Equity Plan ..................................................................... 13
Conclusion ........................................................................................................................... 18
References ........................................................................................................................... 19
Research and Planning Group for California Community Colleges................................. 20
Executive Summary

Purpose

Many colleges examine disproportionate impact in the context of one demographic characteristic at a time. This entails, for example, first examining potential differences in student success across ethnic groups and then, in a separate analysis, examining such differences across age groups. What these comparisons would not reveal, however, is whether ethnicity and age jointly influence student success. Therefore, it is possible that there are no differences in student success across age groups, but that such differences are readily apparent when both age and ethnicity are jointly considered. This joint influence, or interaction, between two or more characteristics can only be explored by simultaneously disaggregating data across multiple variables—that is, via multivariate data disaggregation.

A multivariate approach to data disaggregation allows us to develop more targeted interventions designed to help bridge observed achievement gaps among highly specific groups of students (e.g., older veterans, male students of color). This paper offers the reader an introduction to multivariate disaggregation in the context of real-world examples; the aim is to both illustrate the data analytic procedures associated with multivariate data disaggregation and offer the reader concrete examples that highlight real-world applicability. What follows is a discussion of Student Equity Plans from two separate community colleges that showcase this approach.

Case Studies

Both the College of Marin and Santa Monica College offer examples that highlight the benefit of multivariate data disaggregation. In the case of the College of Marin, we see that examining the success rates on the basis of both foster youth status and gender offers us more insight about which specific student group is in need of support services (i.e., male foster youth) than would otherwise be the case had the college only examined their data by foster youth status. Similarly, Santa Monica’s approach to examining basic skills English rates by both ethnic group and gender demonstrated the pressing need for support that Black male students in particular, relative to all other examined groups.

Conclusion

It is critical that we examine the relationships that exist across two or more demographic characteristics. After all, students do not only belong to a single demographic group; students’ identities span many possible identifications or groups simultaneously—one student may, for instance, identify as a male, Hispanic, veteran, foster youth, first generation, etc. Thus, a multivariate approach more closely reflects the reality of the complex lives and intersecting identities of the students we serve.
Introduction

When one subgroup of students attains an outcome, such as a certificate or degree completion, at a rate that is significantly lower than an established threshold rate, that subgroup may be referred to as “disproportionately impacted”. According to the Chancellor’s Office, “disproportionate impact is a condition where some students’ access to key resources and supports and ultimately their academic success may be hampered by inequitable practices, policies and approaches to student support.” (California Community College Chancellor’s Office, 2013, p. 4). In a previous white paper (Sosa, 2017), we explored the various methods by which to examine disproportionate impact. However, as briefly discussed in that paper, many such analyses examine disproportionate impact in the context of one demographic characteristic at a time. That is, when examining potential disproportionate impact across three demographic characteristics—say ethnicity, gender, and age—we do so separately for each characteristic. We first compare ethnic groups, and then males and females, and then finally different age groups. We may find, for instance, that African American students demonstrate lower success rates than do other ethnic groups, or that males have lower completion rates than females, or that younger students have lower term-to-term persistence than older students. What these comparisons do not reveal, however, is whether, for instance, African American females have lower success rates than Hispanic males. What about younger (ages 18-20) Asian females—do they have higher or lower success rates than do older (ages 25+) white males? Exploring such complexities may clarify our understanding of what is driving disproportionate impact by revealing unexpected relationships and highlighting certain groups with particularly high levels of disproportionate impact. Uncovering such evidence allows us to develop more targeted interventions designed to help bridge observed achievement gaps among specific groups of students. If the evidence warrants it, one could place particular emphasis on improving completion rates of Hispanic males, rather than trying to address all male students, for example. Such insights require a multivariate disaggregation approach—one in which two or more demographic characteristics (e.g., ethnicity and gender) are examined simultaneously.

---

1 Please note that, as of October 2017, the California Community College Chancellor’s Office recommends the use of the percentage point gap (PPG) method as the primary method of determining disproportionate impact (CCCCCO, 2017).
Case Study 1: College of Marin’s Student Equity Plan

Figure 1 illustrates the mean success rates of former foster youth relative to other students as described in the 2014-2015 Student Equity Plan submitted by College of Marin (College of Marin, 2014). Students identified as foster youth achieved an average course success rate of 43.2% as compared to 68.0% achieved by non-foster youth students. Using the percentage point gap method along with any other disproportionate impact methods described in Sosa (2017), this finding would indicate that students identified as foster youth are disproportionately impacted (i.e., they achieve course success at a significantly lower rate than non-foster youth).

Figure 1. Mean Course Success Rates by Foster Youth Status at College of Marin in Fall 2012 and Fall 2013

The example depicted in Figure 1 is an example of bivariate data disaggregation. It is referred to as bivariate because one examines two values of a single demographic characteristic at one time. Figure 1, for instance, reflects the examination of foster youth status (a demographic characteristic) with two values: Foster youth and non-foster youth. One could also create a separate figure to depict the success rates of males and females, reflecting an examination of gender as the demographic characteristic. In both cases, these serve as examples of bivariate data disaggregation because two statuses of a demographic characteristic (whether it be foster youth status or gender) is being examined at one time.

But there is more to the story that we discover when we disaggregate by two demographic characteristics at one time, also known as multivariate disaggregation. We can, as College of Marin did, examine the link between foster youth status and success rates by gender, as shown...
in Figure 2. The difference between what is depicted in Figure 1 as compared to Figure 2 is that the latter reflects the additional layer of disaggregation by gender. In looking at this figure, we can see that both female and male foster youth achieve lower success rates than their same-gender non-foster youth peers (i.e., female foster youth achieved lower success rates than non-foster youth females). However, the potential impact that being a foster youth has on success rates varies depending upon students’ gender: the discrepancy between female foster youth and non-foster youth is much smaller (12.8 percentage points) than the discrepancy between male foster youth and non-foster youth (25.8 percentage points). In other words, the gap between males and females is much larger among foster youth than among non-foster youth. This type of conjoint impact can be described as an interaction effect; gender and foster youth status interact. Specifically, knowing a student’s gender help us to understand the impact of foster youth status on course success rates.

**Figure 2. Mean Course Success Rates by Foster Youth Status and Gender at College of Marin in Fall 2012 and Fall 2013**

Therefore, bivariate data disaggregation allows one to identify if males have different success rates than do females, and separately, whether foster youth have different success rates than do non-foster youth. However, bivariate data analysis does not allow one to determine if both of the aforementioned demographic characteristics jointly impact a student outcome. Only multivariate data disaggregation therefore allows one to determine whether two demographic characteristics, such as gender and foster youth status, interact with one another to affect an outcome, such as success rates.

Also noteworthy is that multivariate data disaggregation still offers one all the benefits of bivariate data disaggregation, but in addition, also allows one to explore for potential interactions. For example, the findings depicted in Figure 2 point to lower success rates among
foster youth as compared to non-foster youth, which is the same bivariate finding illustrated in Figure 1. However, because Figure 2 simultaneously depicts the success rates by gender and foster youth status, the findings also indicate that male foster youth achieve the lowest course success rates of the four groups depicted in Figure 2 (i.e., male foster youth, male non-foster youth, female foster youth, and female non-foster youth). Using established disproportionate impact methods (including the CCCCO recommended percentage point gap method), male foster youth are disproportionately impacted even as compared to female foster youth. In this way, a multivariate approach to disaggregating data uncovers relationships that would otherwise remain obscured by examining each demographic characteristic in isolation. So while the bivariate findings suggest that all foster youth are at risk for achieving lower success rates and that strategies designed to help this group of students should be developed, the multivariate findings also point to a specific group—male foster youth—that is at particularly high risk.

Table 1 offers us a tabular version of the data depicted in Figure 2—that is, it illustrates each of the four course success rates by foster youth status and gender. The key to identifying an interaction is to determine whether the impact that one demographic characteristic (e.g., gender) has on success rates is influenced by a second demographic characteristic (e.g., foster youth status). In our case, if the observed impact that gender has on success rates was not influenced by foster youth status, then the observed gender differences for each foster youth status would be very similar (or even identical). In other words, if we were to find that the female foster youth success rate was 10 percentage points higher than the male foster youth success rate, then we would expect to find that the female non-foster youth success rate was also 10 percentage points higher than the male non-foster youth student success rate—this would indicate that female students earn higher success rates regardless of foster youth status. However, as shown in Table 1, the gender difference is much larger among foster youth (Gender Difference = 20.5) than among non-foster youth (Gender Difference = 7.5%). Similarly, the difference by foster youth status differs by gender. Specifically, the difference between male foster youth and male non-foster youth (Foster Youth Status Difference = 25.8%) is larger than the difference between female foster youth and female non-foster youth (Foster Youth Status Difference = 12.8%). These findings indicate that the observed gender difference in success rates is dependent upon students’ foster youth status, and by extension, that the observed foster youth difference in success rates is dependent upon students’ gender. Thus, the findings presented in Table 1 (along with Figure 2) are illustrative of the interaction between gender and foster youth status.2

2 Interactions are formally identified via inferential statistical techniques—in this case, those techniques would focus on establishing whether the aforementioned gender difference for foster youth status and non-foster youth
Table 1. Course Success Rates by Foster Youth Status and Gender

<table>
<thead>
<tr>
<th>Foster Youth Status</th>
<th>Male</th>
<th>Female</th>
<th>Gender Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foster Youth</td>
<td>42.3%</td>
<td>62.8%</td>
<td>20.5%</td>
</tr>
<tr>
<td>Non-Foster Youth</td>
<td>68.0%</td>
<td>75.5%</td>
<td>7.5%</td>
</tr>
<tr>
<td>Foster Youth Status Difference</td>
<td>25.8%</td>
<td>12.8%</td>
<td></td>
</tr>
</tbody>
</table>

The next step in the analytical process is to determine which specific subgroup(s) are disproportionately impacted. While we may know that the impact that foster youth status has on success rates depends upon students’ gender, we need to formally determine which subgroup(s) are disproportionately impacted. In other words, are female or male foster youth at especially high risk of earning non-successful grades? On the basis of both Figure 2 and Table 1, it appears that male foster youth are at especially high-risk (42.3% success rate). However, we can formally determine this by employing disproportionate impact methods, such as the percentage point gap (PPG; Sosa, 2017).

Table 2 illustrates course success rates for each of the aforementioned subgroups (i.e., gender and foster youth status). Additionally, the table illustrates findings stemming from the use of the PPG. As discussed in Sosa (2017), the PPG method measures the difference in percentage points between a given subgroup’s outcome and the overall average (or mean) for those outcomes across all subgroups. The difference may be positive (as when a subgroup achieves better than average mark) or negative (as when a subgroup achieves a lower than average mark). It is therefore defined as follows:

\[
\text{Percentage point gap} = \frac{\% \text{ of outcome for students in subgroup}}{\% \text{ of outcome for all students}}
\]

Given that the overall success rates across all four subgroups is 72.0%, calculating PPG values entails taking the differences between each subgroup’s success rate and the overall average of 72.0%. PPG values resulting from the use of this approach are shown in Table 2.

status was statistically significant. Please see http://www.skidmore.edu/~hfoley/StatLabs/S1.Interaction.Lab.pdf for a more detailed statistical discussion of interactions.
Table 2. Course Success Rates by Foster Youth Status and Gender—Illustration of Percentage Point Gap Index

<table>
<thead>
<tr>
<th>Ethnic Group</th>
<th>Cohort Count</th>
<th>Outcome Count</th>
<th>Success Rate (Per Group)</th>
<th>MOE Threshold</th>
<th>Point Gap Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male Foster Youth</td>
<td>265</td>
<td>112</td>
<td>42.3%</td>
<td>-6%</td>
<td>-29.7</td>
</tr>
<tr>
<td>Female Foster Youth</td>
<td>408</td>
<td>256</td>
<td>62.8%</td>
<td>-5%</td>
<td>-9.2</td>
</tr>
<tr>
<td>Male Non-Foster Youth</td>
<td>14,221</td>
<td>9,672</td>
<td>68.0%</td>
<td>-3%</td>
<td>-4.0</td>
</tr>
<tr>
<td>Female Non-Foster Youth</td>
<td>19,330</td>
<td>14,594</td>
<td>75.5%</td>
<td>-3%</td>
<td>+2.7</td>
</tr>
<tr>
<td>Total</td>
<td>32,224</td>
<td>24,633</td>
<td>72.0%</td>
<td>-3%</td>
<td>+2.7</td>
</tr>
</tbody>
</table>

Note. Percentage point gaps values were computed by taking the difference between each subgroup’s success rate and the overall success rate (i.e., 42.3% - 72.0% = -29.7 percentage point gap). MOE thresholds are based upon Appendix A of the chancellor’s office document on working with the percentage point gap (CCCCO, 2017). For a review of the procedures involved in determining a given MOE threshold, please see Sosa (2017).

Source: College of Marin’s Student Equity Plan (2014-2015)

Identifying an instance of disproportionate impact depends upon a corresponding cohort’s margin of error (MOE) threshold. That is, a given PPG value must be lower (that is, more negative) than its corresponding MOE threshold to be considered an instance of disproportionate impact. Since the obtained PPG value for male foster youth (-29.7), female foster youth (-9.2), and male non-foster youth (-4.0) were all lower than their corresponding margin of error, each is identified as an instance of disproportionate impact. Figure 3 offers readers a visual depiction of the PPG values relative to corresponding MOE values. PPG values exceeding corresponding MOE values are shown in red. We thus conclude that the success rates of foster youth students depended upon their gender. That is, while the success rates were lower among foster youth than non-foster youth, they were especially lower among male foster youth.

---

3 For a review of the procedures involved in determining MOE thresholds, please see Sosa (2017).
4 The authors also employed two additional disproportionate impact methods, the 80% method and the proportionality index. Both approaches identified male foster youth as disproportionately impacted; this was the only group identified as disproportionately impacted using either method. The authors recommend employing all three methods of disproportionate impact. See Sosa (2017) for more information on all three of the methods.
To reiterate, the advantage of the multivariate approach over the bivariate approach is that, in addition to being able to provide the same information as a univariate approach, it also addresses a third question: Are success rates affected jointly by both gender and foster youth status? Indeed, a multivariate approach may more accurately reflect real-world students in that students are not just male or female; instead they simultaneously identify with a host of demographics groups (e.g., Hispanic, veteran, foster youth) that collectively help to shape students’ academic and personal experience. A multivariate approach therefore helps to shed light on how these various characteristics interact to bring about student outcomes. The one disadvantage of the multivariate approach is the additional complexity in interpreting the disproportionate impact associated with two separate demographic characteristics. Identifying the disproportionate impact among males is relatively straight-forward; identifying the disproportionate impact among males that are also foster youth can present a greater conceptual and data-analytic challenge.
Case Study 2: Santa Monica College’s Student Equity Plan

As part of its 2014-2015 Student Equity Plan, Santa Monica employed a multivariate approach to data disaggregation (Santa Monica College, 2014). Across several outcomes, the college examined its data by simultaneously examining two demographic characteristics. One of the outcomes for which the campus employed a multivariate approach was with respect to its basic skills progress rates. Instead of only separately examining ethnic and gender differences in basic skills progress rate, they examined both demographic characteristics simultaneously. Although Santa Monica included all ethnic groups in its multivariate analyses, for simplicity’s sake, we have limited our presentation of their findings to only two ethnic groups.

Table 4 illustrates the progression rates for all four possible groups resulting from examining both ethnicity and gender together: (a) Black males, (b) Black females, (c) White males, and (d) White females. Based upon the overall progression rates shown by ethnicity and gender, there appear to be two findings. First, there appears to be a relationship between ethnicity and basic skills progression. That is, Black students achieved a lower progression rate than did White students. To specifically examine this issue, one need only compare the combined Black student progress rate across both males and females (30.2%) to the combined White student progress rate across both males and females (50.3%). A second finding is that female students achieved a higher progress rate than did male students, regardless of students’ ethnicity. Specifically, female students achieved a progress rate of 43.5% (combined for Black and White female students) while male students achieved a progress rate of 36.8% (combined for Black and White male students). While the calculations are not shown here, the Black students were found to be disproportionately impacted as were male students.

5 The Basic Skills Progression Rate describes the ratio of the number of students completing a college-level course compared to the number students in the same cohort who began their sequence of courses in basic skills. Santa Monica employed a six-year track out period; in this case, it meant they tracked students enrolling in basic skills English in 2007-2008 and identified the percentage of those same students successfully completing college-level English by 2012-2013.

6 Based upon all three methods for determining disproportionate impact, including the percentage point gap method.
Table 4. Basic Skills Progression Rates by Ethnic Group and Gender

<table>
<thead>
<tr>
<th>Ethnic Group</th>
<th>Male</th>
<th>Female</th>
<th>Ethnic Group Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black Students</td>
<td>24.6%</td>
<td>34.7%</td>
<td>30.2%</td>
</tr>
<tr>
<td>White Students</td>
<td>47.6%</td>
<td>52.9%</td>
<td>50.3%</td>
</tr>
<tr>
<td>Gender Overall</td>
<td>36.8%</td>
<td>43.5%</td>
<td></td>
</tr>
</tbody>
</table>

*Source. Santa Monica College's Student Equity Plan (2014-2015)*

In statistical terms, finding disproportionate impact for each individual demographic characteristic is referred to as uncovering a *main effect*. We compared groups within each demographic characteristic; that is, we first examined ethnicity by comparing Black students to White students, and then we examined gender by comparing males to females. They are considered examples of a bivariate approach because each set of findings reflects a comparison between groups within each demographic characteristic separately; in other words, ethnicity was examined separately from gender.

We can also simultaneously examine the joint influence of both ethnicity and gender on progression rates, as Santa Monica college did. Doing so would reflect a multivariate approach to disaggregation. When data are disaggregated in this fashion, it is to determine whether an interaction exists between the two demographic characteristics under investigation. In this case, an interaction would mean that the difference in progress rates between Black and White students depends upon students’ gender—this is an interaction because the impact that ethnicity may have on progress rates is dependent upon students’ gender.

To determine whether an interaction exists, we need to consult the information presented in Table 5. To determine whether an interaction is present, we ask ourselves this question: “Does the difference in progress rates that exists between male and female students differ between Black and White students?” In other words, is the gap between males and females larger for Black students than it is for White students? As shown in Table 4, the gender gap among Black students (34.7% – 24.6% = 10.1%) is almost twice the size of the gender gap among White students (52.9% – 47.6% = 5.3%). Similarly, the gap in success rates between male Black students and male White students is larger (47.6% – 24.6% = 23.0%) than the gap observed between female Black and female White students (52.9% – 34.7% = 18.2%). Thus, these
findings indicate that both demographic characteristics jointly influence progression rate—that is, they interact\(^7\). See Figure 4 for a graphical depiction of the progression rates by subgroup.

Table 5. Basic Skills Progression Rates by Ethnic Group and Gender

<table>
<thead>
<tr>
<th>Ethnic Group</th>
<th>Male</th>
<th>Female</th>
<th>Gender Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black Students</td>
<td>24.6%</td>
<td>34.7%</td>
<td>10.1%</td>
</tr>
<tr>
<td>White Students</td>
<td>47.6%</td>
<td>52.9%</td>
<td>5.3%</td>
</tr>
<tr>
<td>Ethnic Group Difference</td>
<td>23.0%</td>
<td>18.2%</td>
<td></td>
</tr>
</tbody>
</table>

Figure 4. The 2012-13 Basic Skills English Progression Rate Among Black and White Students by Gender at Santa Monica College (2007-2008 Cohort)

\(^7\) As noted earlier, Interactions are formally identified via inferential statistical techniques. Please see [http://www.skidmore.edu/~hfoley/StatLabs/S1.Interaction_Lab.pdf](http://www.skidmore.edu/~hfoley/StatLabs/S1.Interaction_Lab.pdf) for a more detailed statistical discussion of interactions.
As described earlier, the next step after establishing an interaction is to identify which of the examined subgroup(s) may be disproportionately impacted. To that end, Table 6 illustrates the progression rates for each of the aforementioned subgroups (i.e., ethnicity and gender) along with findings stemming from the use of the PPG (see earlier discussion of College of Marin’s data for all the calculation procedures). The PPG findings indicate that male Black students (-15.7) and female Black students (-5.6) were disproportionately impacted.8 What does this all mean in practical terms? The evidence suggests that both male and female Black students were disproportionately impacted; however, the specific subgroup of students that is in most need of support services when considering English progress rates are Black males.

Table 6. Basic Skills Progression Rates Ethnicity and Gender—Illustration of the Percentage Point Gap Index

<table>
<thead>
<tr>
<th>Ethnic Group</th>
<th>Cohort Count</th>
<th>Outcome Count</th>
<th>Success Rate (Per Group)</th>
<th>MOE Threshold</th>
<th>Point Gap Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male Black Students</td>
<td>264</td>
<td>65</td>
<td>24.6%</td>
<td>-6%</td>
<td>-15.7</td>
</tr>
<tr>
<td>Female Black Students</td>
<td>326</td>
<td>113</td>
<td>34.7%</td>
<td>-5%</td>
<td>-5.6</td>
</tr>
<tr>
<td>Male White Students</td>
<td>296</td>
<td>141</td>
<td>47.6%</td>
<td>-6%</td>
<td>+7.3</td>
</tr>
<tr>
<td>Female White Students</td>
<td>306</td>
<td>162</td>
<td>52.9%</td>
<td>-6%</td>
<td>+12.6</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1,192</strong></td>
<td><strong>481</strong></td>
<td><strong>40.3%</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* Percentage point gaps values were computed by taking the difference between each subgroup’s success rate and the overall success rate (i.e., 24.6% - 40.3% = -15.7 percentage point gap). MOE thresholds are based upon Appendix A of the chancellor’s office document on working with the percentage point gap (CCCCO, 2017). For a review of the procedures involved in determining a given MOE threshold, please see Sosa (2017).

*Source: Santa Monica’s Student Equity Plan (2014-2015)*

---

8 The authors also employed two additional disproportionate impact methods, the 80% method and the proportionality index. Both approaches identified male Black students as disproportionately impacted and the 80% method (but not the proportionality index) identified female Black students as disproportionately impacted. The authors recommend employing all three methods of disproportionate impact. See Sosa (2017) for more information on all three of the methods.
Figure 5 offers readers a visual depiction of the PPG values relative to corresponding MOE values. PPG values exceeding corresponding MOE values are shown in red. We thus conclude that the success rates of Black students were meaningfully lower than were the rates among White students (i.e., both male and female Black students are disproportionately impacted). In addition, the multivariate findings indicate that the gap in success rates between Black and White students depended upon their gender: While the success rates of Black students (regardless of gender) were lower than that of White students, they were especially lower among male Black students.

Figure 5. Percentage Point Gap Values and Margin of Errors by Ethnic Group and Gender
Conclusion

The first step in identifying equity gaps is to closely examine disaggregated data. However, it is important to go beyond common practice and consider those multivariate relationships that may describe groups who are experiencing disproportionate impact. After all, students’ identities are complex; they do not belong exclusively to a single demographic group. Students simultaneously belong to a host of groups—they may, for example, identify themselves as male, Hispanic, veteran, foster youth, first generation, etc. Thus, a multivariate approach more closely reflects the actual population of students we serve because it takes into account these various identities at the same time.

A multivariate approach entails at least two levels of analysis. First, we gather evidence concerning how demographic characteristics are individually related to the outcome of interest. For instance, we may first uncover ethnic group differences, and then separately, uncover gender differences. Any differences observed at this level are referred to as main effects. The next level of disaggregation occurs when we examine the interaction between both demographic characteristics—a situation in which the impact that one demographic variable has on an outcome depends upon the level of a second demographic variable (e.g., when the impact of ethnicity on completion is only present for members of one gender). This is how we can discover, for instance, that male Black students are at risk (rather than just Black students or just male students).

While conducting DI analyses, it is important to include the PPG approach since it is required by the chancellor’s office (CCCO, 2017). However, we also recommend that practitioners employ more than one disproportionate impact method when determining instances of disproportionate impact. This is because employing a separate method allows one to determine whether there exists corroborating evidence for findings obtained via a single disproportionate impact method. Finding that all three methods point to a meaningful gap between the groups under comparison strongly suggests that an equity gap is powerfully present.

Finally, our hope is that, through this paper (and all the others that makeup the Data Disaggregation ASK) the reader may see that data disaggregation is not simply a matter for researchers to contend with. Subgroup differences in success and completion that may exist at our colleges are critical issues that pertain to all education practitioners. To the extent that data is used to inform an array of decision-making that takes place at any college, all stakeholders—including faculty, staff, and administrators—should be active participants in the analysis and interpretation of data-based findings, particularly those that work directly with students. The rich experience that practitioners have with students will likely help in interpreting the meaning of a given multivariate relationship uncovered through data disaggregation; and it is that experience with students that will likely help institutions proactively respond to identified equity gaps. Thus, the rich dialogue and support strategies that results from the application of the methods described in this paper are central to fully addressing the challenges that our students face.
References


College of Marin (2014). *Marin Community College student equity plan*.


Research and Planning Group for California Community Colleges

The RP Group strengthens the ability of California community colleges to discover and undertake high-quality research, planning, and assessments that improve evidence-based decision-making, institutional effectiveness, and success for all students.

Data Disaggregation Applied Solution Kits Team

Craig Hayward
Carolyn Holcroft
Andrew Kretz
Jared Lessard
Giovanni Sosa
Ireri Valenzuela

www.rpgroup.org