Correction to the "Student repeating courses" Section of "DRAFT: Variation in Math Student Pass Rates by Instructor and Section"

Jeff Webb, SLCC Institutional Research

October 19, 2015

1 Introduction

Earlier material presented in "DRAFT: Variation in Math Student Pass Rates by Instructor and Section" (October 2, 2015) on students repeating math classes was discovered to be in error upon review. This document presents the corrected numbers, which, it should be noted, differ pretty substantially from those reported earlier.

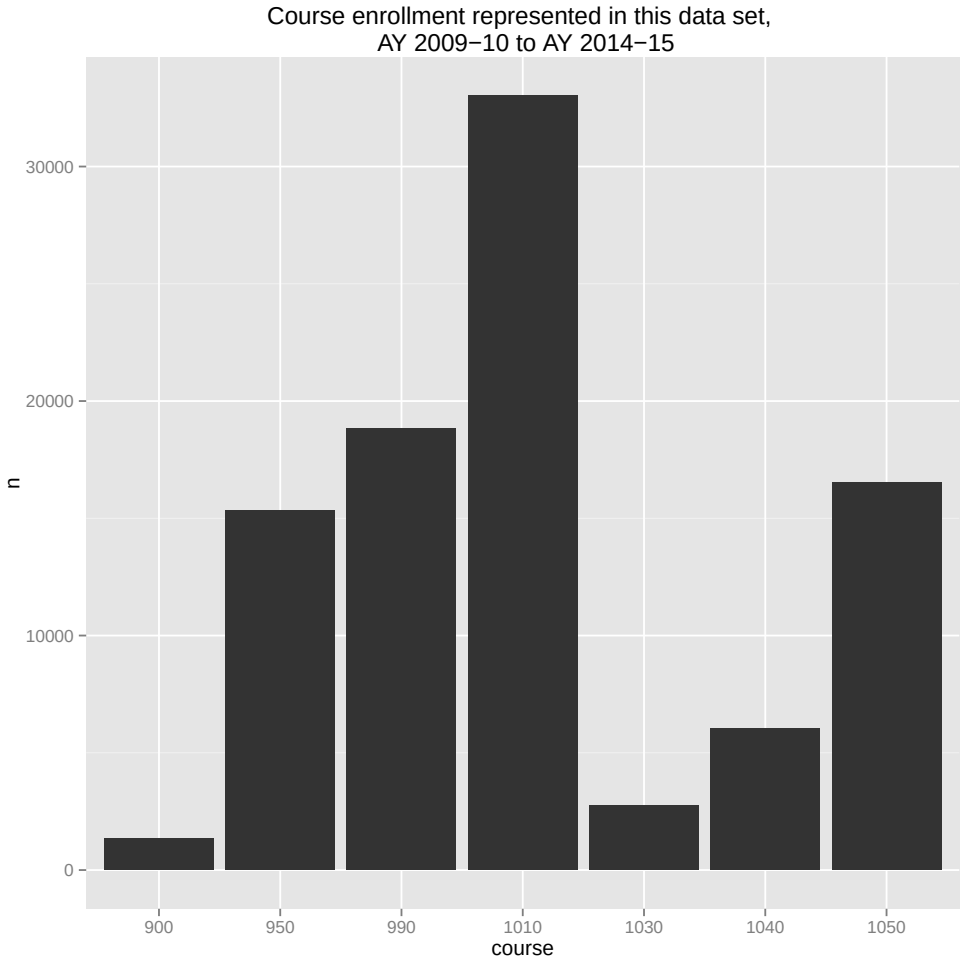
Regrettably, mistakes happen in data analysis. The best course of action is to to have a process of review in place, and through that process to identify mistakes as soon as possible and correct them.

2 Data

The dataset includes all students who took a math course from Fall 2009 through Summer 2015, excluding concurrent enrollment students. We made the following modelling choices for this analysis:

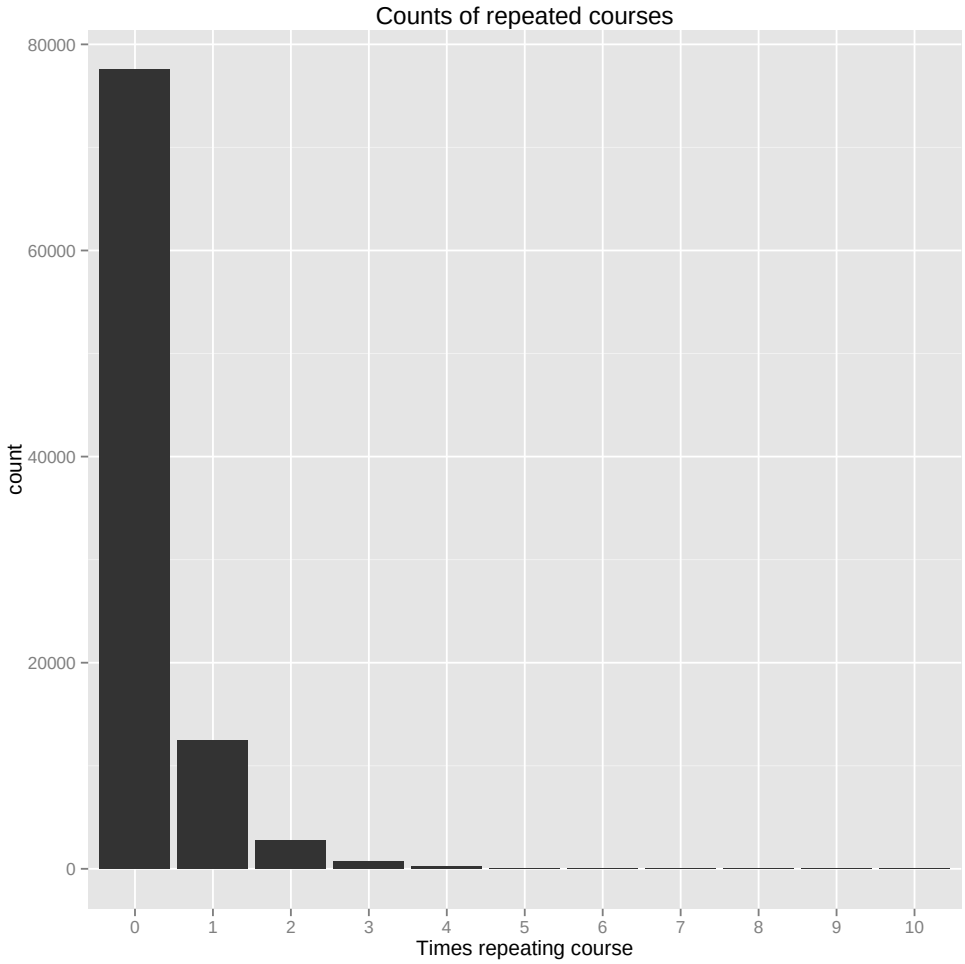
- Passing grade was coded as C or above. Non-passing grades included C-, D+, D, D-, E, I or W.
- The math courses in this analysis included Math 900, 950, 990, 1010, 1030, 1040, and 1050.
- A course was counted as repeated only if it was repeated in the data set. Because repetitions prior to Fall 2009 were therefore not counted, the number of repeated courses reported here probably slightly undercounts the actual number.
- Prior math grade was simply the last recorded math grade, even if it was a D, E, or W, *in the courses included in the data set*. If, for example, a student got a D in 1010, and later took 1015 and received an A, then for the purposes of this analysis the D was used as the prior grade. In cases where students took multiple math classes in a single semester, the grade in the most advanced math class was used as the prior grade.

There were 45820 unique students represented in the data set, with a total of 93868 course enrollments. The number of enrollments per course was as follows:



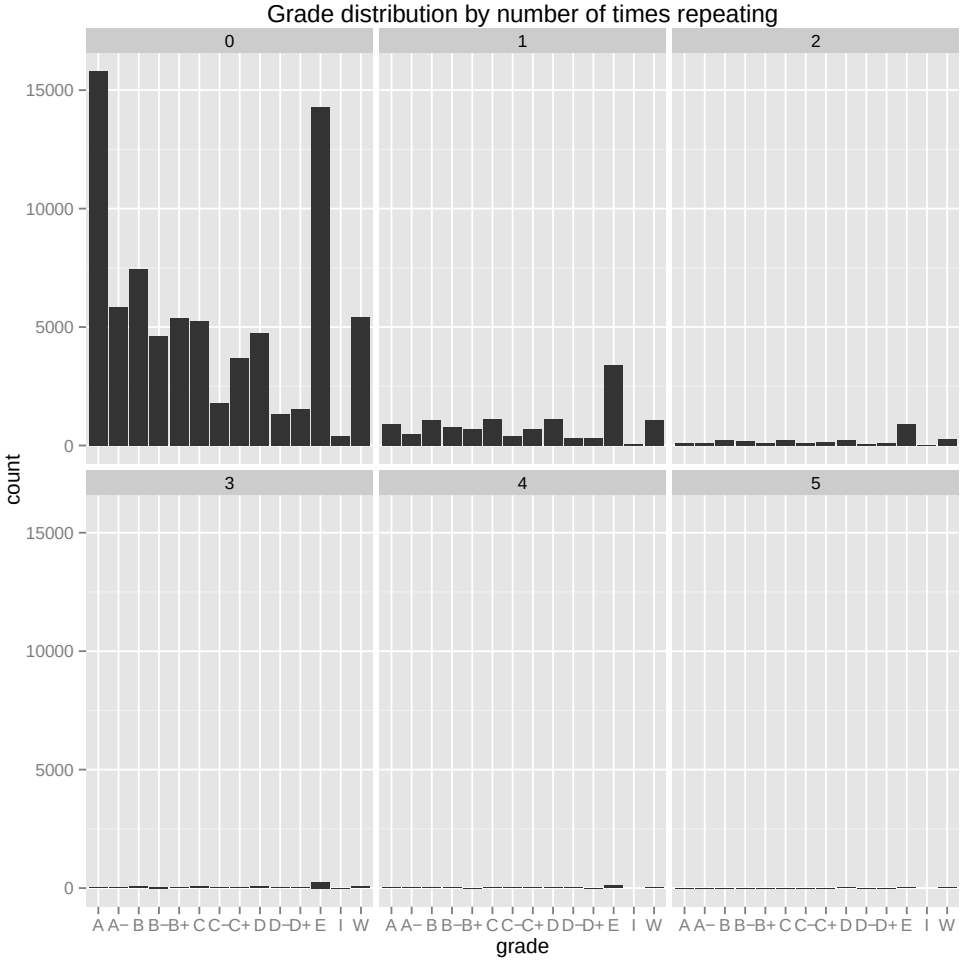
3 Results: Students Repeating Courses

Here are counts of repeated courses (0 means that the course was being taken for the first time):



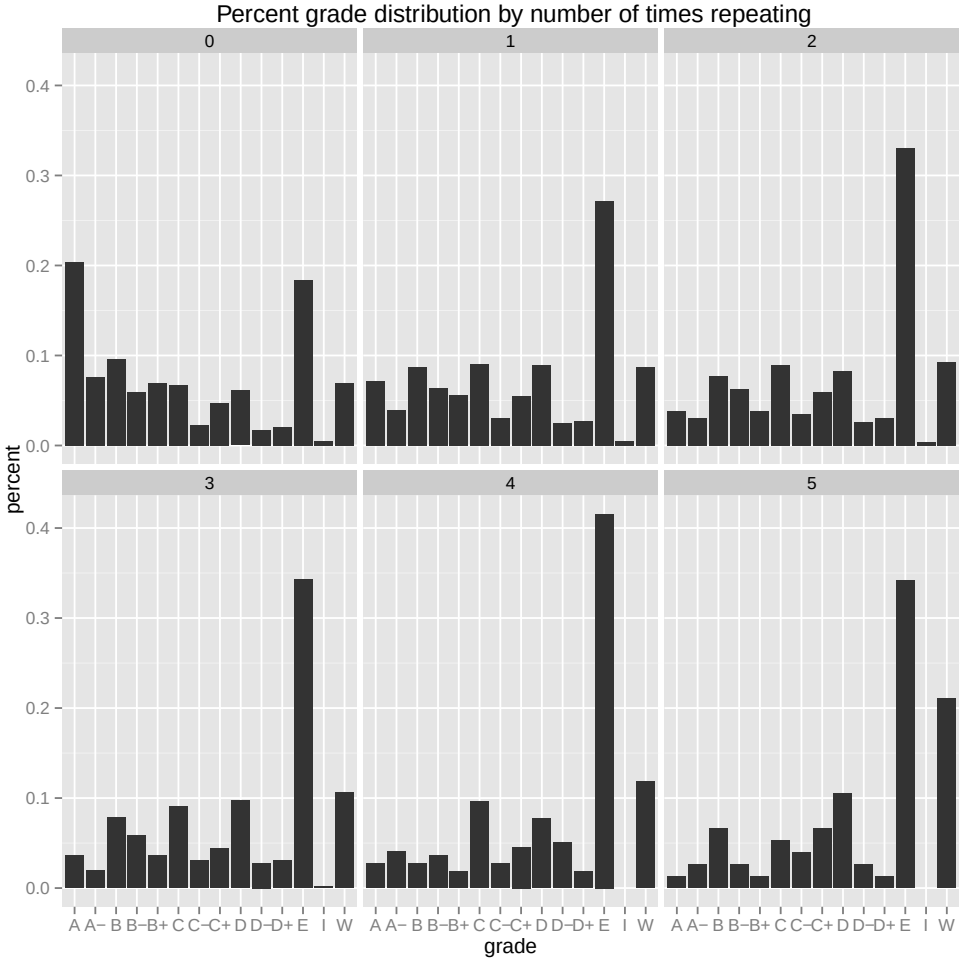
(A small number of students did indeed repeat courses up to 10 times in the 6 year period represented in this dataset.)

How were student grades impacted by course repetition?



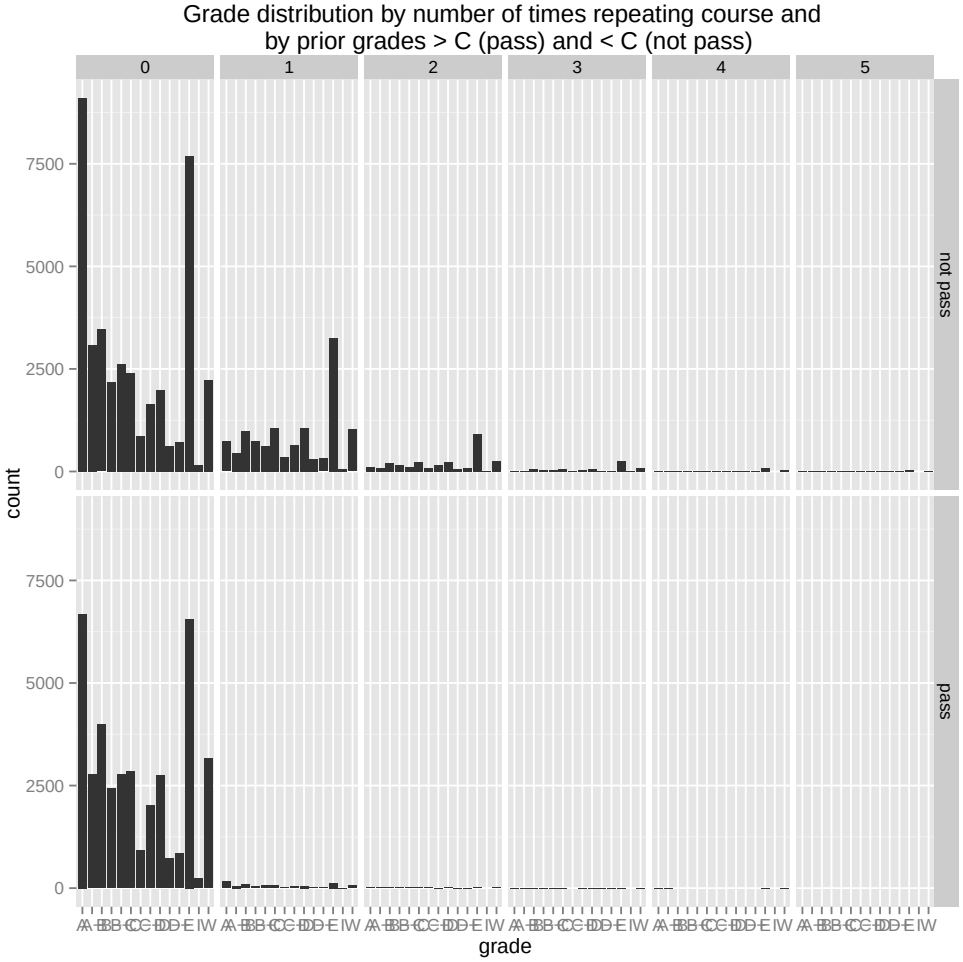
The shape of the grade distribution changed as students repeated courses, but not as dramatically as we reported previously.

Here are grade distributions represented as a percentage for each time repeated:

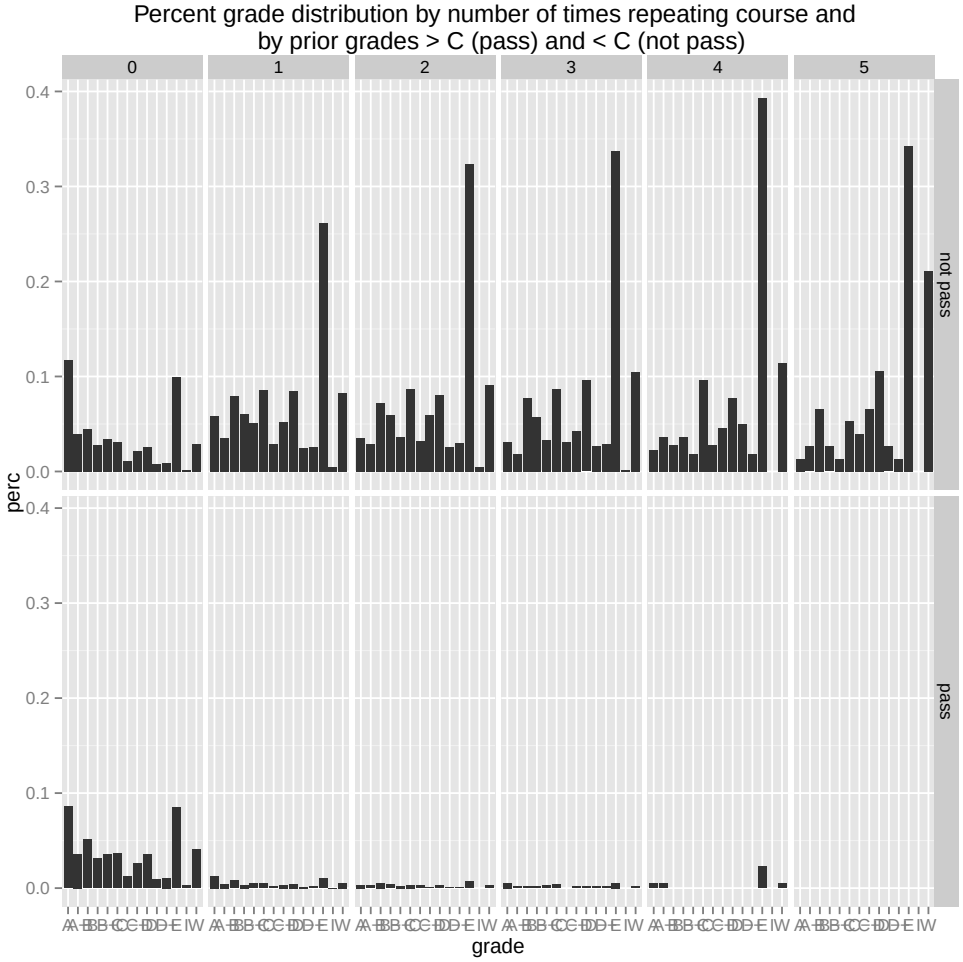


In the above graph we can see that percentage of students getting an A dropped substantially with course repetition, the percentage receiving an E went up, but the percentage getting a B or C remained somewhat constant through at least repetition = 3.

How was the grade distribution for each time repeating a course affected by prior grade performance in that course? The following plot summarizes prior grade performance as pass/not pass.



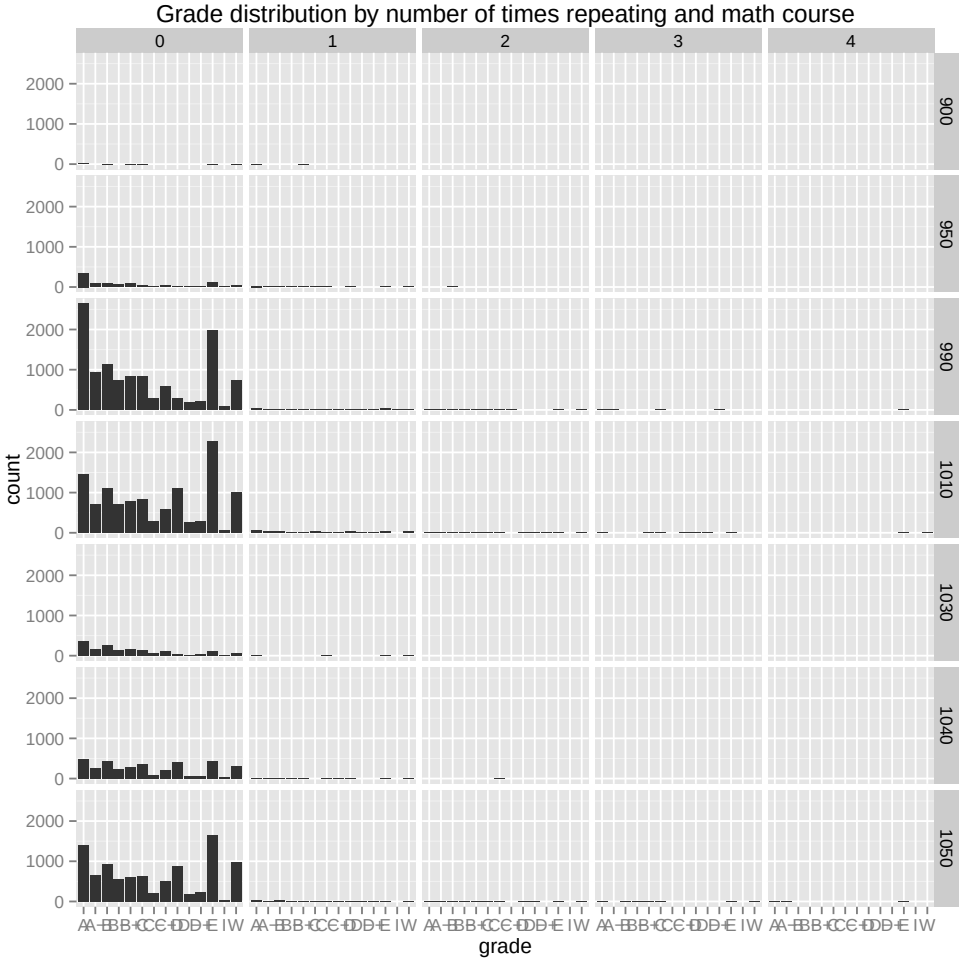
We can see that students sometimes repeated a course without being obligated to. Here is the same plot, but with grade distribution represented as a percentage for each time repeating a course.



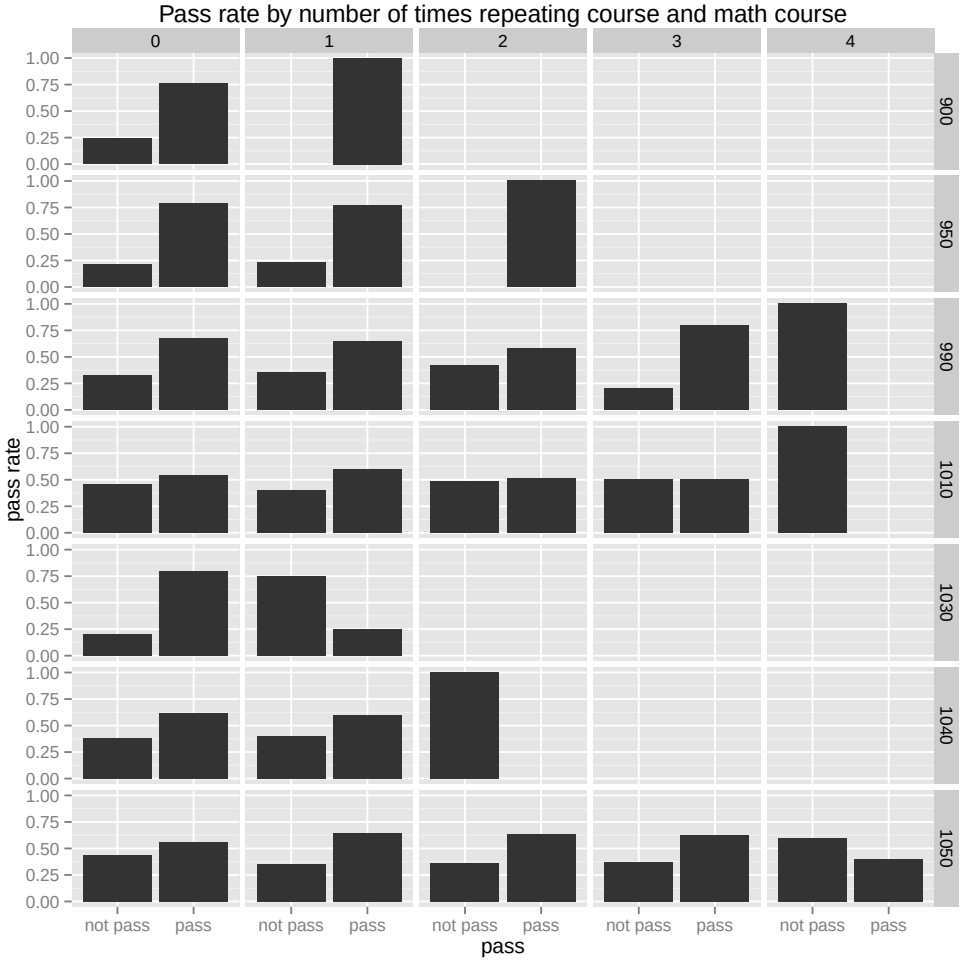
(For subsequent plots we will exclude students who were repeating courses they passed —the bottom panel in the above plot—so as to focus on the performance of students who were required to repeat the course.)

Note that with each time repeating a course through repetition = 4, the percentage of students failing the course rose

How did grade distributions vary by math course?

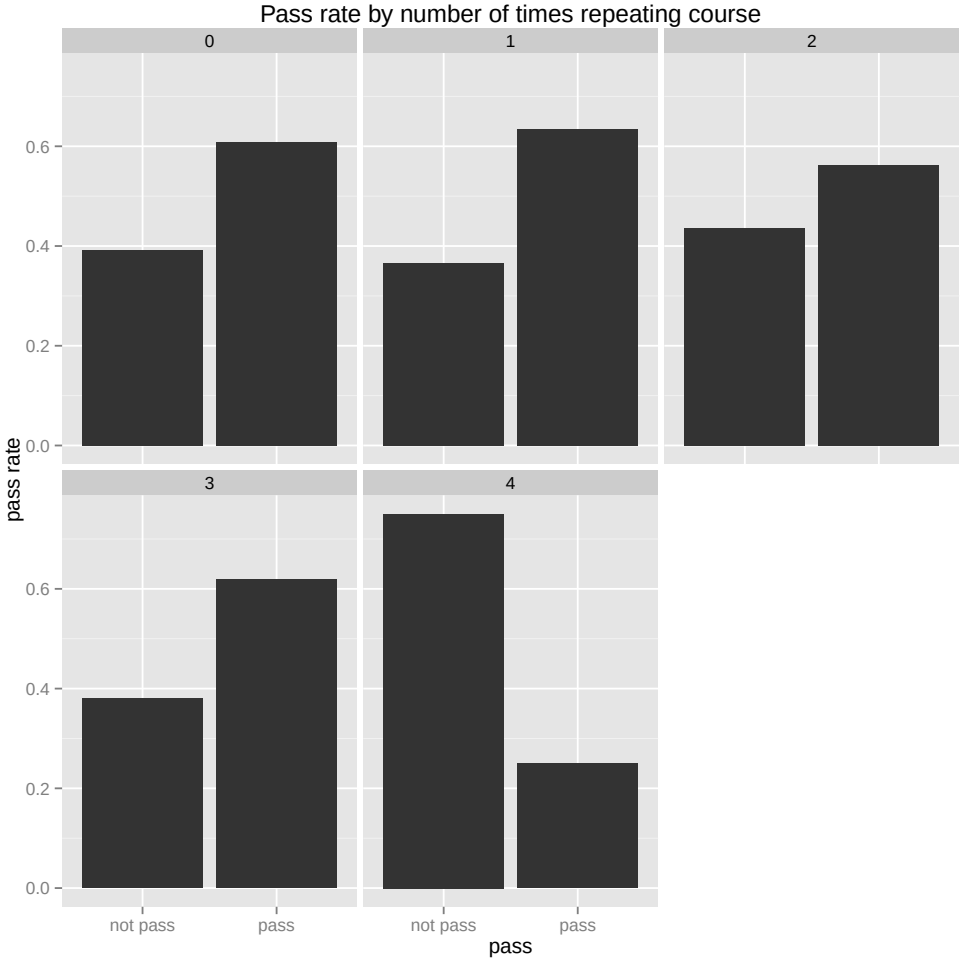


We can simplify the above plot by summarizing the grade distribution as pass rates.

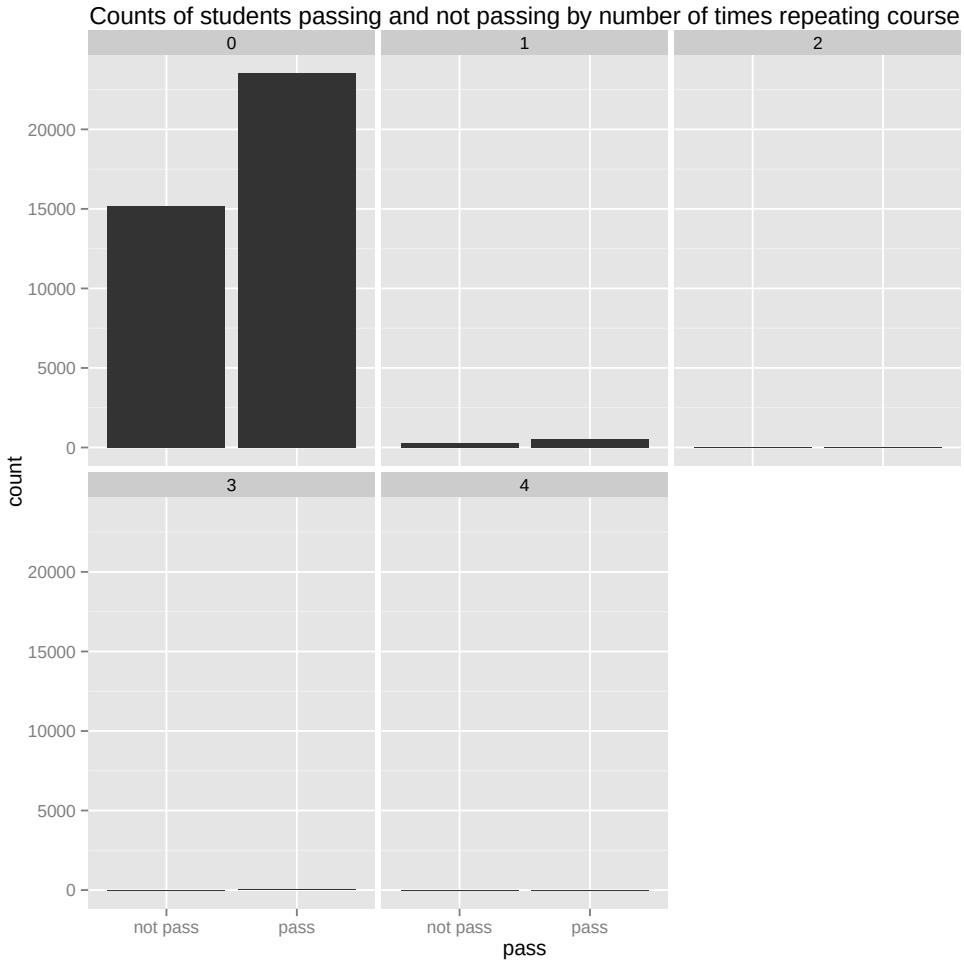


The above plot makes it clear that students actually do reasonably well when repeating classes (except in the case of 1030, where passing initially appears to be fairly easy). The percentages of those passing who were repeating the course for the first time were greater than or equal to the percentage of those passing when taking the course for the first time.

Let's simplify this plot even further by aggregating over courses. This allows us to see larger patterns.

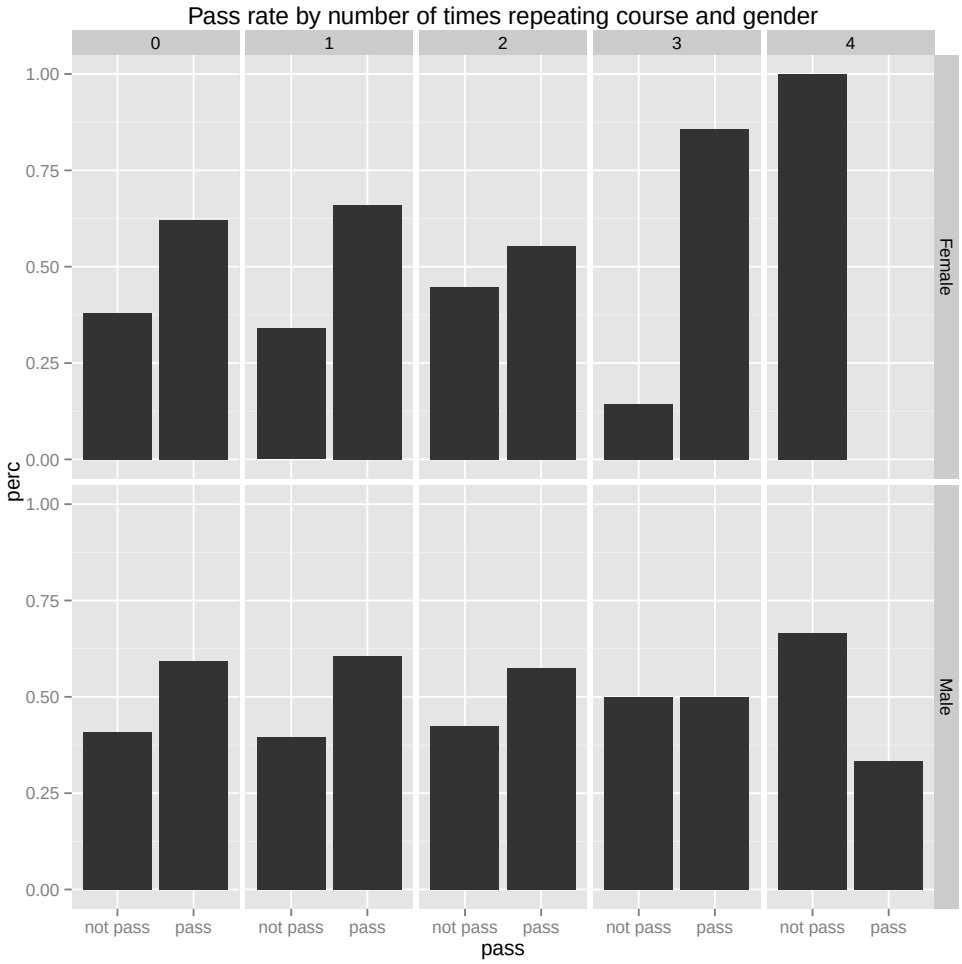


It is important to keep in mind the scale of course repeating—quite small relative to the number of students taking courses for the first time.

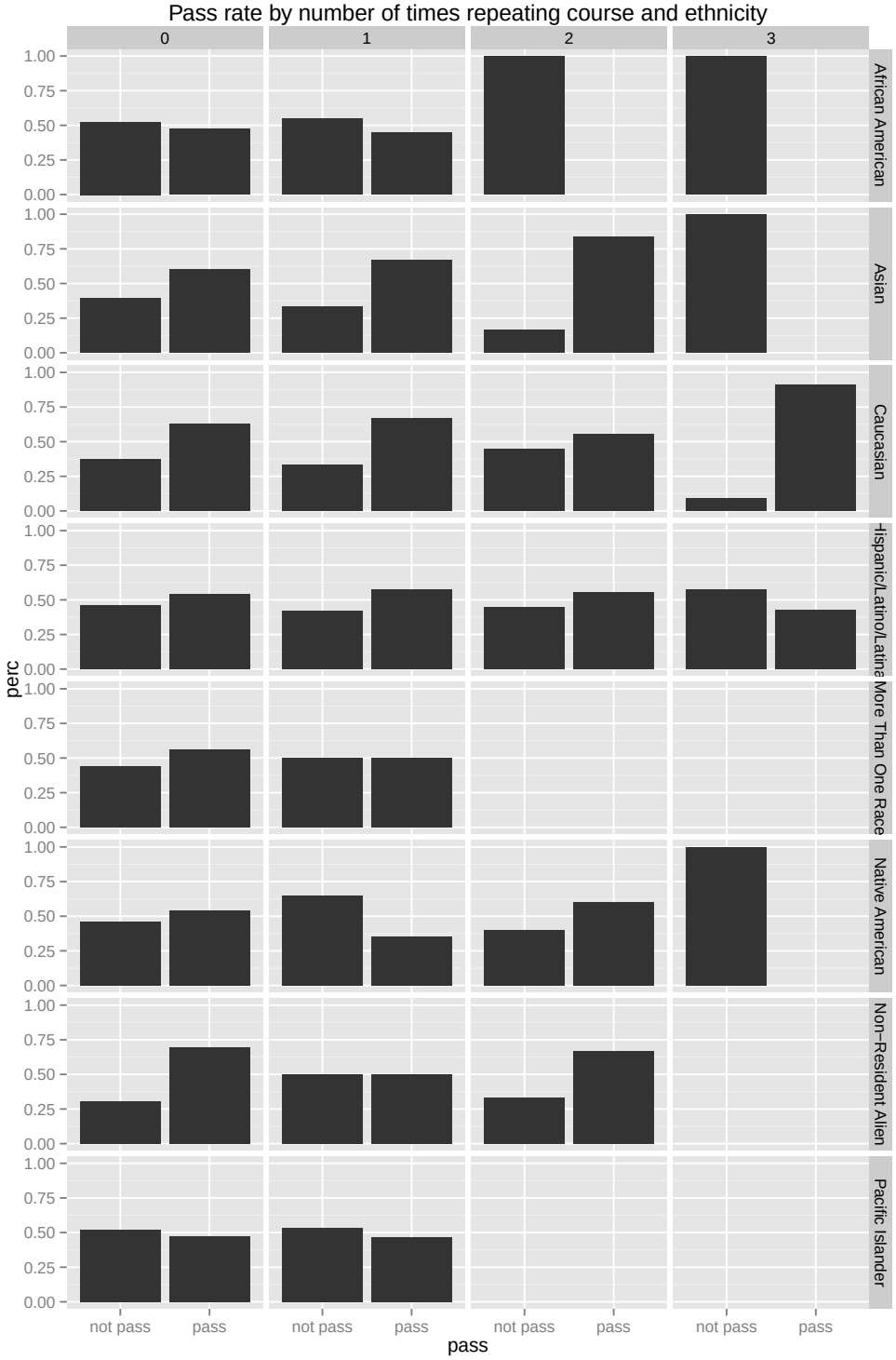


The above two plots indicate that (a) the numbers of students repeating courses was quite small relative to the overall population and (b) the percentage of students passing at repetition 1 - 3 was similar to the percentage passing when taking the course for the first time. However, this relationship changes dramatically in the case of repeating the course for the fifth time (repetition = 4).

How do pass rates vary by times taking the course and student characteristics? Age was not a predictor of course repetition but gender and ethnicity were



(The fluctuations at repetitions = 3 and 4 were due to small n.) Women did especially well the first time repeating a course.



(The fluctuations at repetitions = 2 and 3 were due to small n.) Here we see obvious disparities in pass rates with African American and Pacific Islander students having lower than 50% pass rates when taking math courses for the first time. Asian, Caucasian and Hispanic students fared best when repeating a course.

4 Recommendations

Two points are worth emphasizing in light of the above material:

1. Students who repeat courses have decent success percentage-wise (the actual numbers involved were, of course, quite small). At each course repetition prior to repetition = 4 students were able to convert about 60% of the failures from the previous try into successes. That means that for the students who repeated, .4 x .6 = .24 of the original failures (at repetition = 0) were converted into successes at repetition 1, .24 x .6 = .144 were converted into successes at repetition 2, and .144 x .6 = .086 were converted into successes at repetition 3. So, by the time students repeated a course 3 times, almost half (.24 + .14 + .09 = .47) of the initial failures for those students were converted into successes. This is not to say that their time and success would not have been served better by a different approach, but in this case persistence does seem to pay off eventually. There is perhaps a not-negligible cost, however, for those students who were never successful and experienced repeated failure.

2. The relationship between success and failure switches dramatically at repetition = 4. The change in success rates suggests that repeating a course 4 times is futile. If we are looking for a clear signal from the data about maximum number of course repetitions, this would be it. A math department policy might be to require an advisor meeting after 1 unsuccessful course repeat (prior to repeating a course for the second time) to assess what is going wrong and how it might be addressed, and then to prohibit course repeats after the fourth unsuccessful attempt to pass (before repetition = 4). At that point, the likelihood of not passing is far less than the likelihood of passing.