DRAFT: Exploratory Analysis of Course Sequences in Math

Jeff Webb, SLCC Institutional Research

November 17, 2015

1 Introduction

What were the most typical math course-taking sequences of SLCC students? How did these differ by starting course? What were the most typical sequences for students who completed QL? Where were the largest loss points in math course sequences?

This paper engages these questions (and more) and is intended to serve to as a starting point for a discussion of math course taking pathways in relation to QL attainment at SLCC. Accordingly, we focused on descriptive statistics.

NOTE: This research is preliminary. The conclusions herein should be regarded as tentative and will likely be revised.

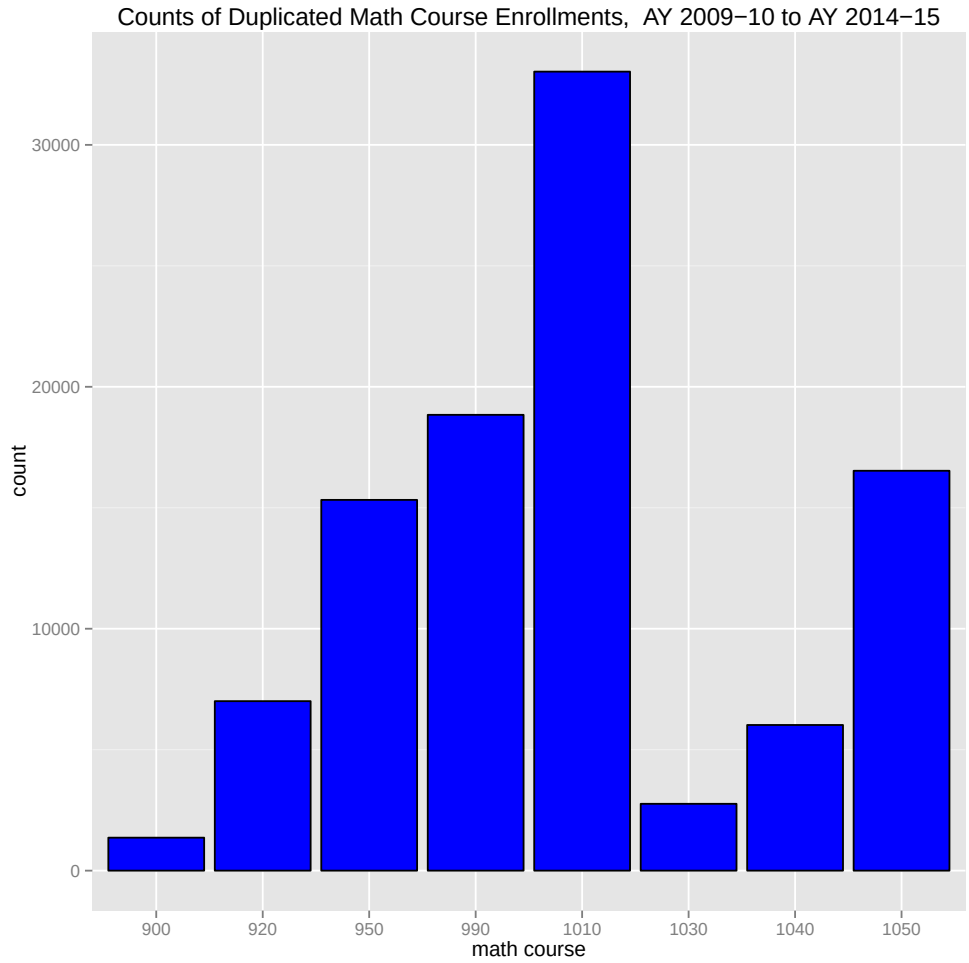
2 Summary of Results

3 Data

The data set included all students who took a math course from Fall 2009 through Summer 2015, excluding concurrent enrollment students. In order to focus on the core sequences to QL completion we have included in the dataset for this analysis only math 900, 920, 950, 990, 1010, 1030, 1040, and 1050. This means that students who took any courses other than those in their path towards QL have been removed.

4 Exploring Course Sequences

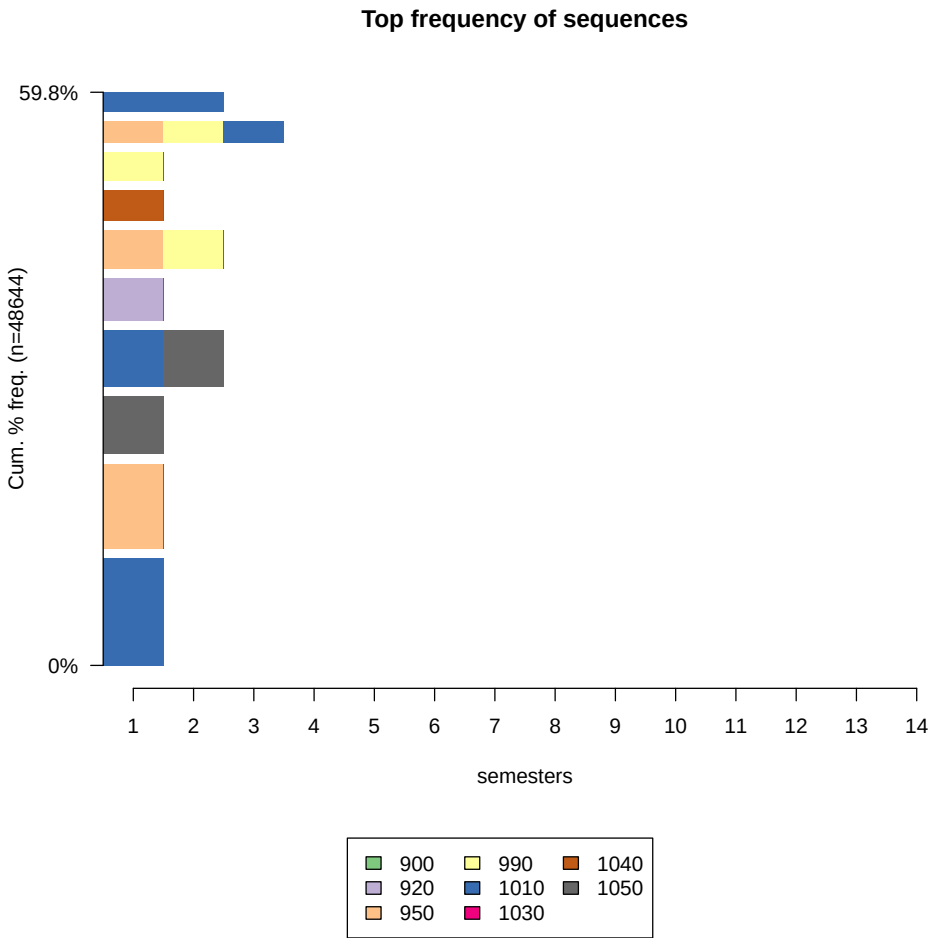
4.1 Counts of course enrollments



1010 is clearly the gateway course, with over 30,000 students in the past 6 years, feeding 1030, 1040 and 1050 in (perhaps unfortunately) unequal proportions. 900 is a small course; more students apparently tested into the higher Dev math courses, 920-990.

4.2 Counts of sequences

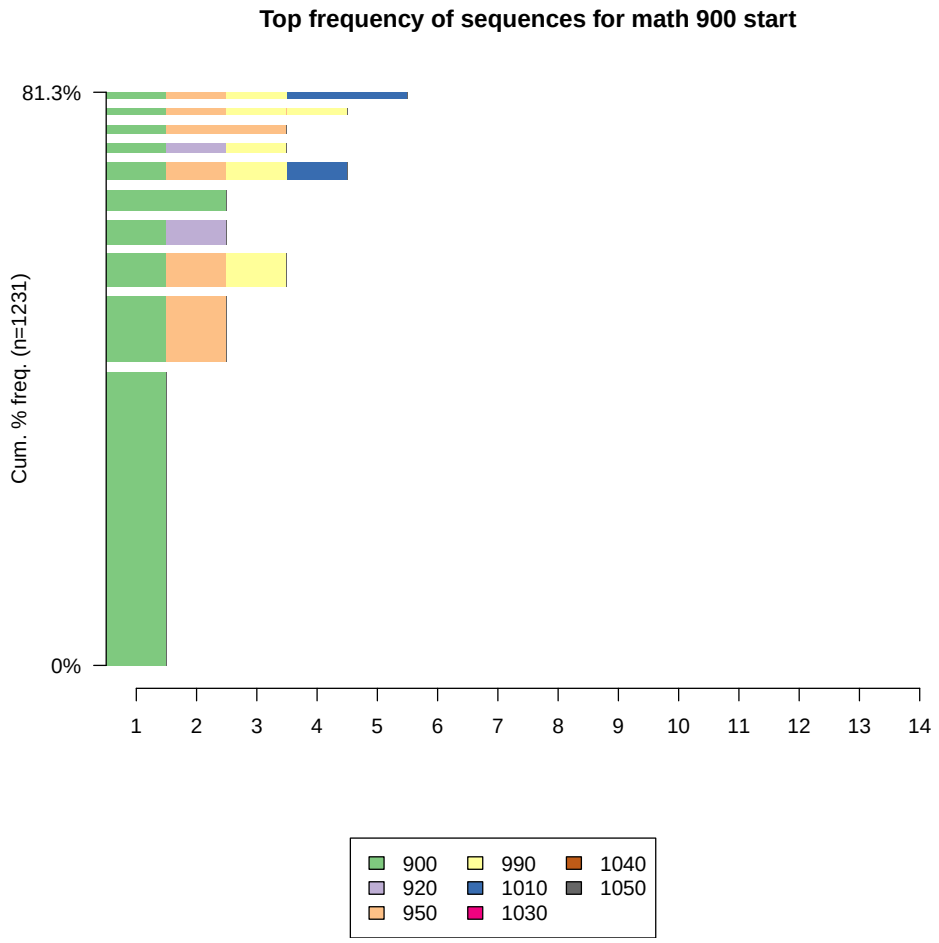
Here is a plot depicting the frequency of course-taking sequences, along with a supporting table containing the top 10 sequences.



	Freq	Percent
1010	6439	13.2
950	5049	10.4
1050	3442	7.1
1010-1050	3367	6.9
920	2559	5.3
950-990	2287	4.7
1040	1829	3.8
990	1696	3.5
950-990-1010	1272	2.6
1010-1010	1147	2.4

In this plot (and those that follow) the width of the bars on the y-axis corresponds to course-taking frequency. The x-axis represents semesters of taking math, with the maximum value of 14 set by the maximum in the data (the path with 14 semesters is very low frequency and is not represented in this plot). Most students started in 1010 or 950 and did not continue (the first two bars moving vertically up the y-axis). Of those who complete QL, most either start in 1040 or 1050, or start in 1010 and then take 1050.

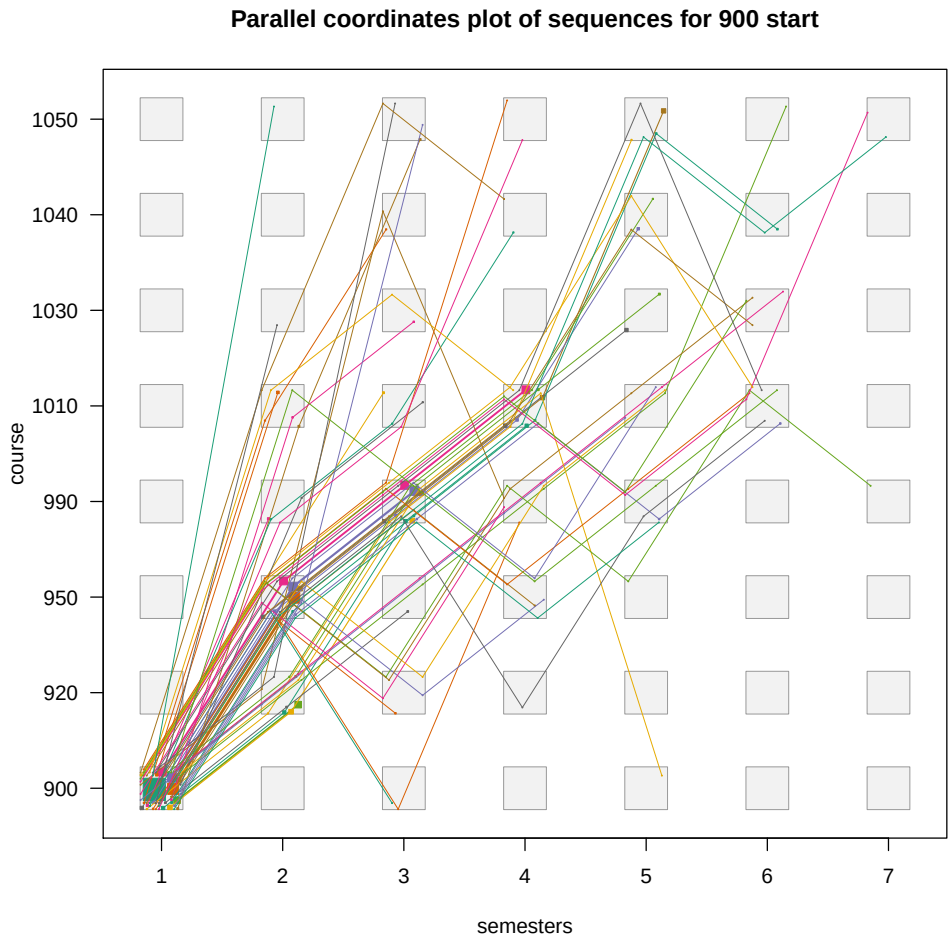
What were the most typical course sequences for the different starting courses—900, 920, 950, 990 and 1010?

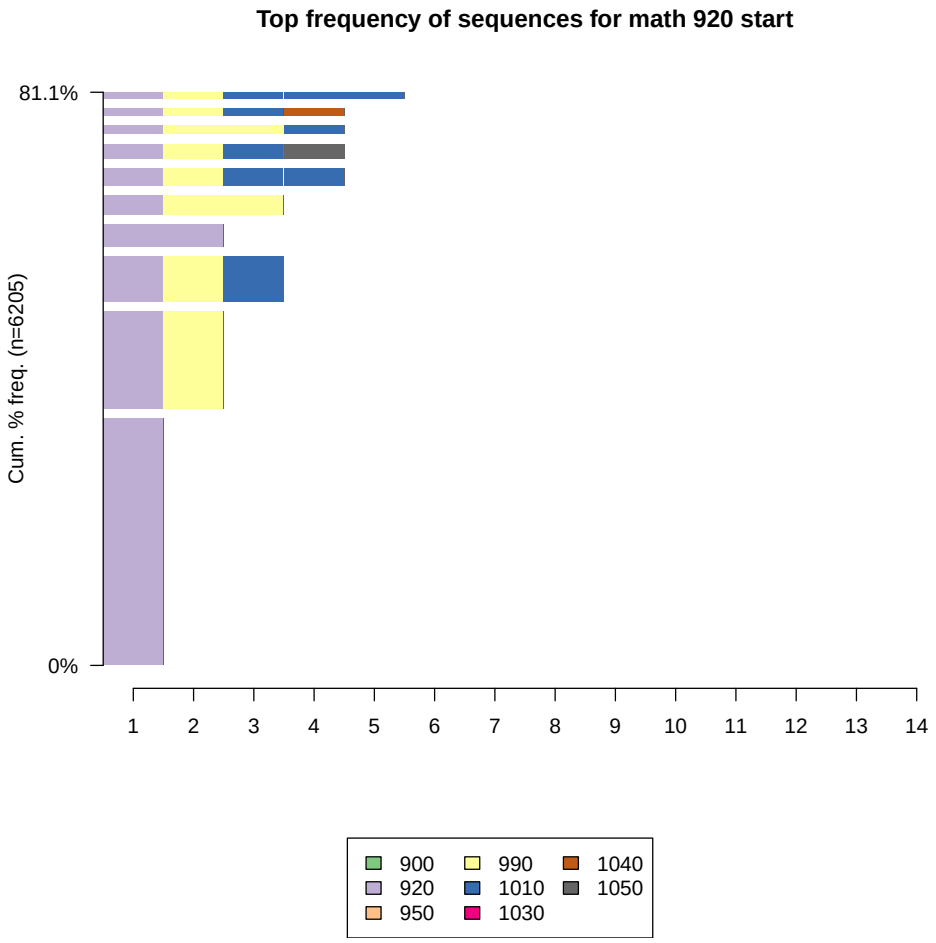


	Freq	Percent
900	605	49.1
900-950	135	11.0
900-950-990	68	5.5
900-920	50	4.1
900-900	42	3.4
900-950-990-1010	36	2.9
900-920-990	20	1.6
900-950-950	18	1.5
900-950-990-990	14	1.1
900-950-990-1010-1010	13	1.1

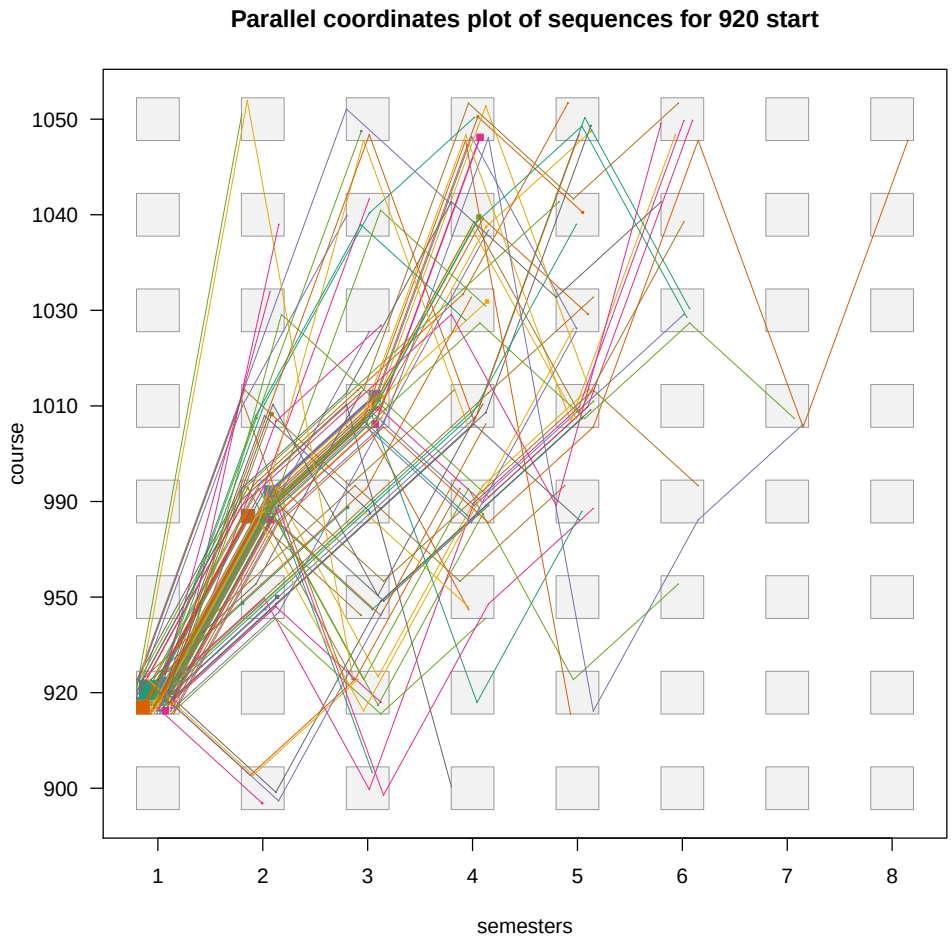
Astoundingly, almost half of the students who start in math 900 do not continue.

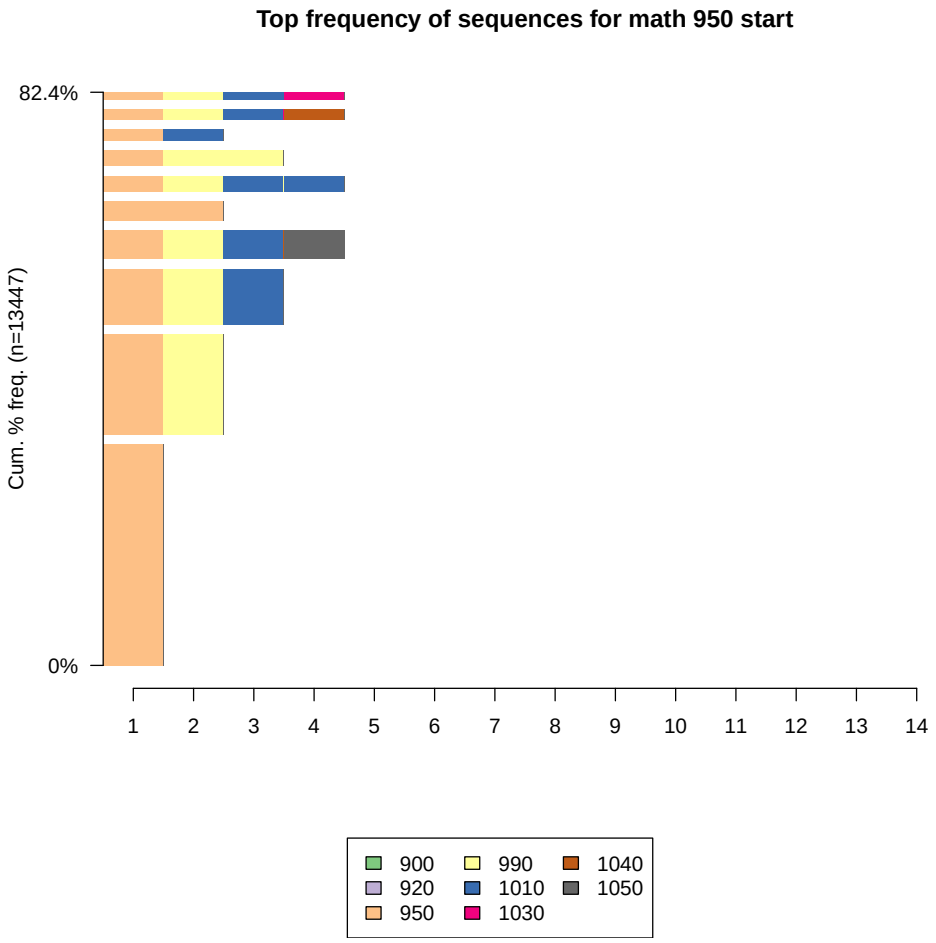
Here is a parallel coordinates plot of these pathways. This type of plot provides an overall view of the different course-taking paths; each line is an individual path, which would have been traveled by—usually—multiple students. The thickness of the lines are frequency-weighted, corresponding to the proportion of students who traveled that path. In the majority cases, as we've seen, students did not continue on from 900. That non-path is represented by the green square in the lower left starting square. The plot should also show the 900-900 sequence, though for some reason I've been having trouble getting that to display.


Parallel coordinates plot of sequences for 900 start
course
1050
1040
1030
1010
990
950
920
900
1
2
3
4
5
6
7
semesters

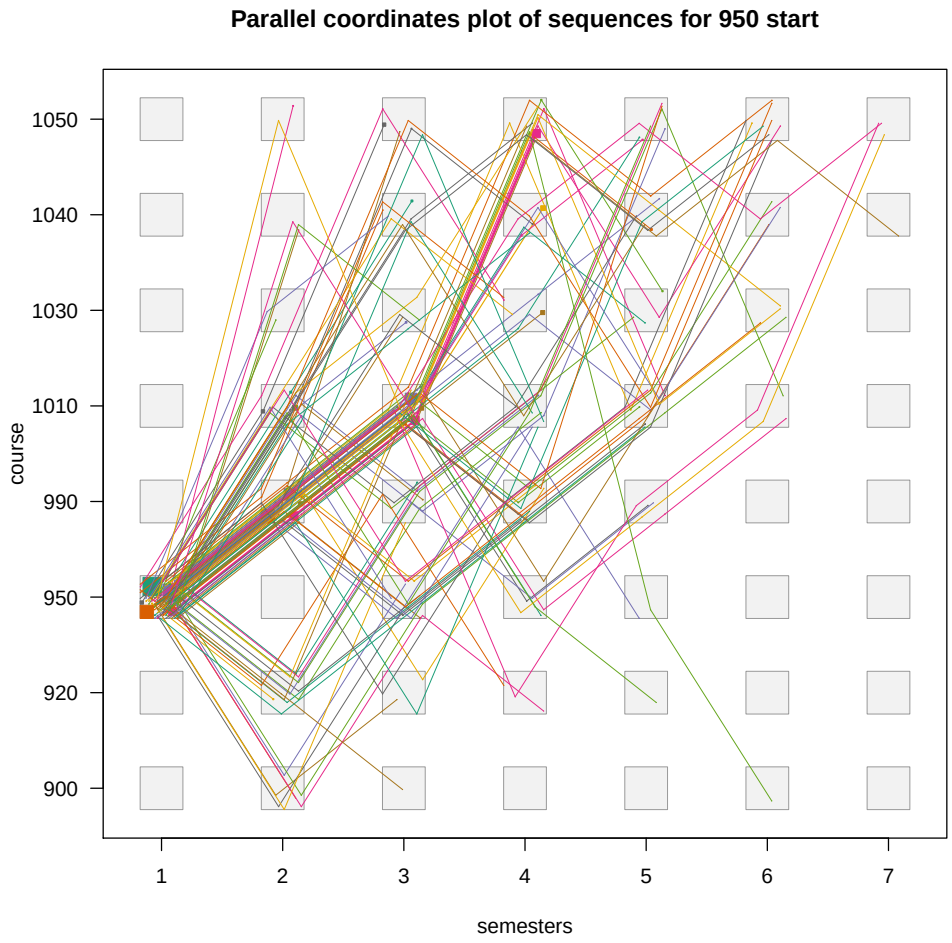


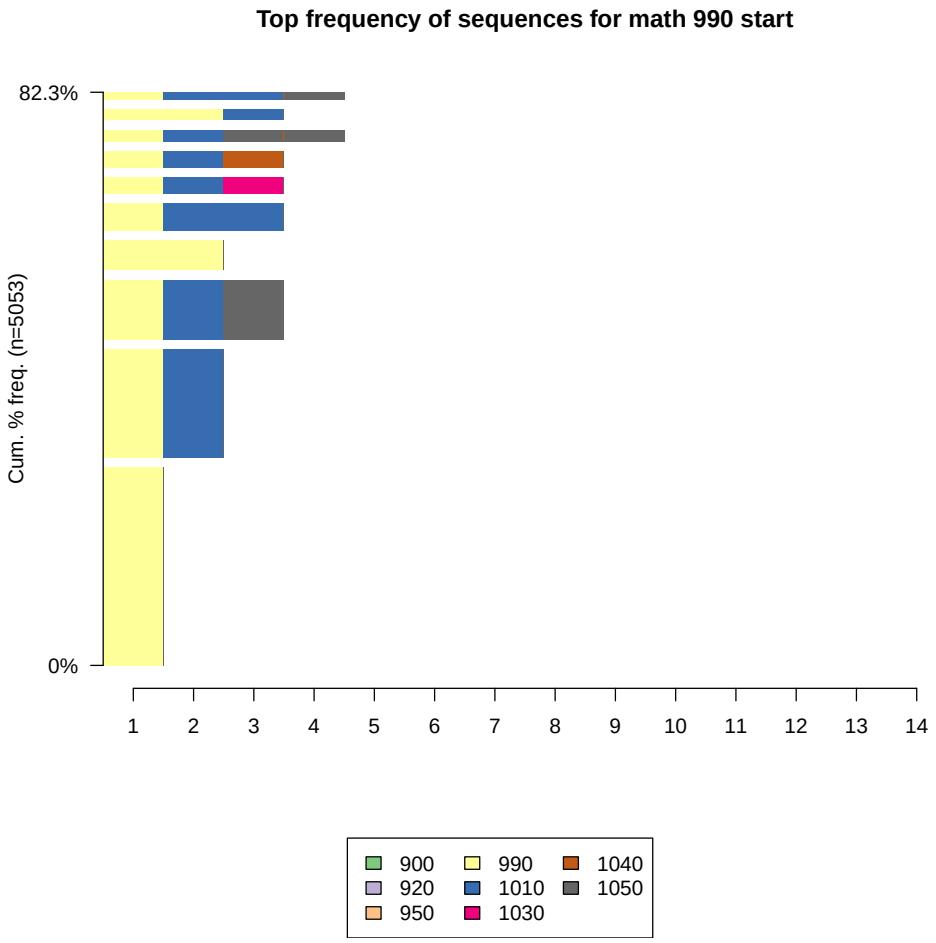
	Freq	Percent
920	2559	41.2
920-990	1009	16.3
920-990-1010	467	7.5
920-920	233	3.8
920-990-990	199	3.2
920-990-1010-1010	182	2.9
920-990-1010-1050	154	2.5
920-990-990-1010	88	1.4
920-990-1010-1040	76	1.2
920-990-1010-1010-1010	66	1.1


Parallel coordinates plot of sequences for 920 start
course
1050
1040
1030
1010
990
950
920
900
1
2
3
4
5
6
7
8
semesters

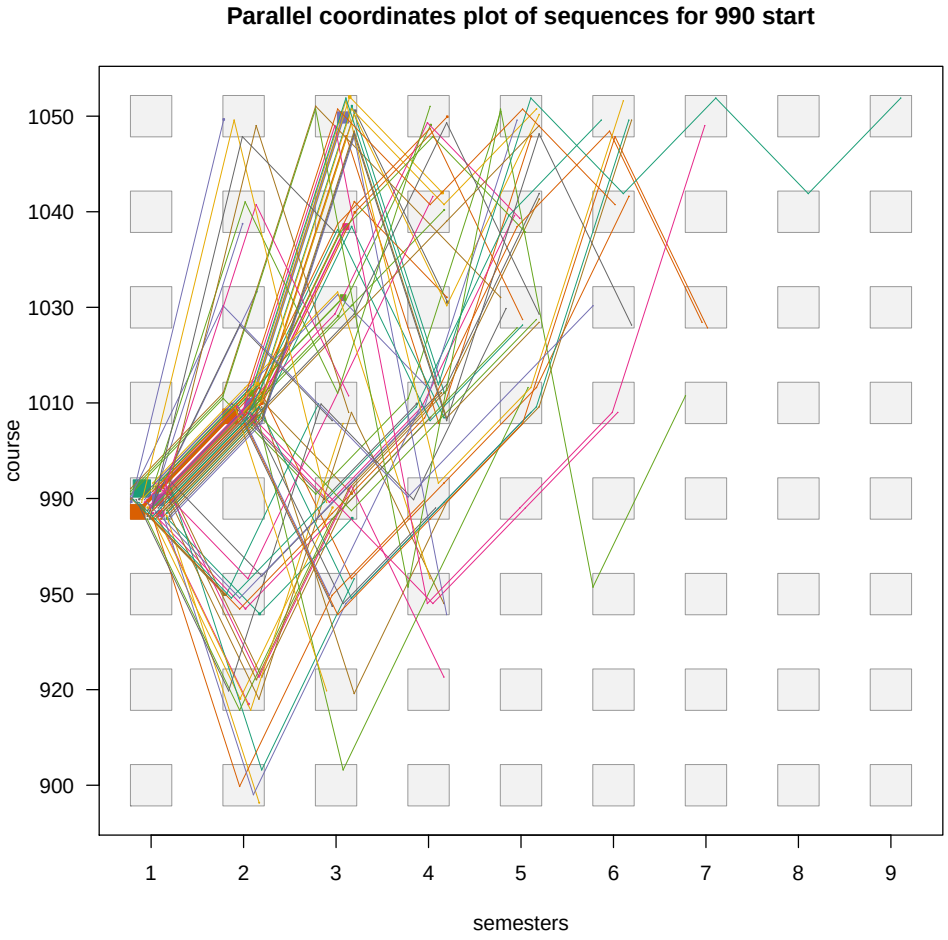


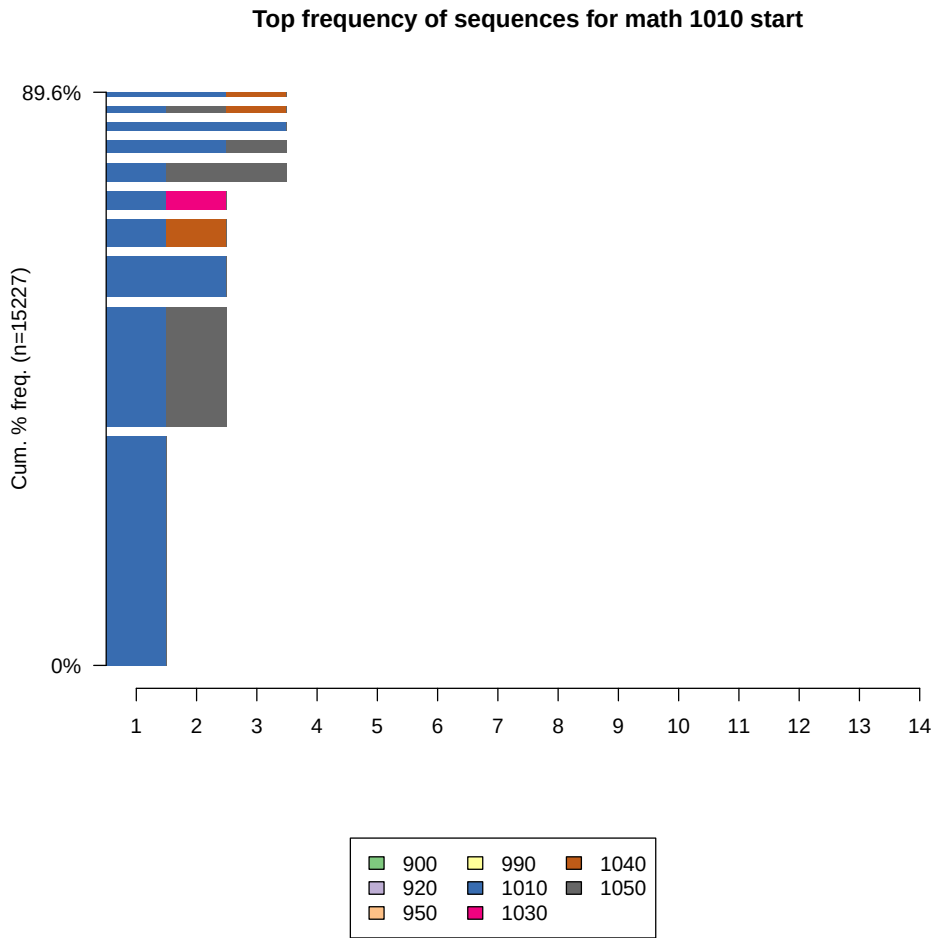
	Freq	Percent
950	5049	37.5
950-990	2287	17.0
950-990-1010	1272	9.5
950-990-1010-1050	654	4.9
950-950	440	3.3
950-990-1010-1010	364	2.7
950-990-990	362	2.7
950-1010	259	1.9
950-990-1010-1040	240	1.8
950-990-1010-1030	160	1.2


Parallel coordinates plot of sequences for 950 start
course
1050
1040
1030
1010
990
950
920
900
1
2
3
4
5
6
7
semesters

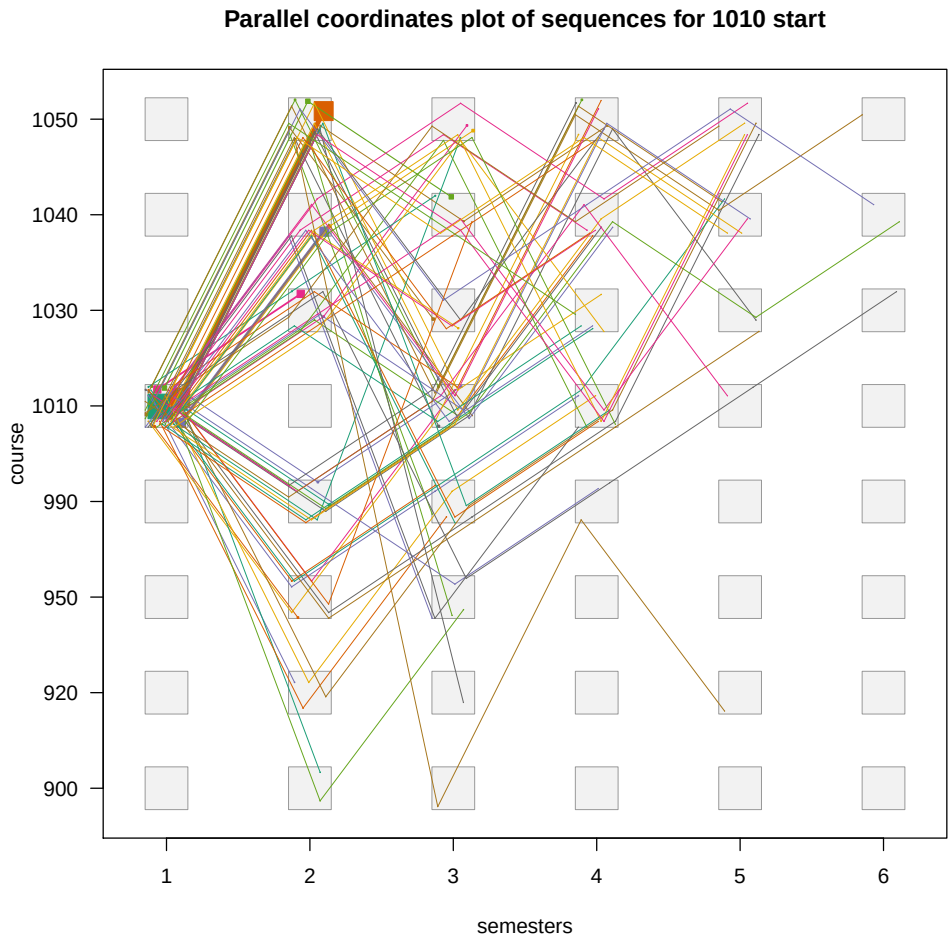


	Freq	Percent
990	1696	33.6
990-1010	926	18.3
990-1010-1050	513	10.2
990-990	251	5.0
990-1010-1010	237	4.7
990-1010-1030	141	2.8
990-1010-1040	137	2.7
990-1010-1050-1050	101	2.0
990-990-1010	87	1.7
990-1010-1010-1050	68	1.3


Parallel coordinates plot of sequences for 990 start
course
1050
1040
1030
1010
990
950
920
900
1
2
3
4
5
6
7
8
9
semesters



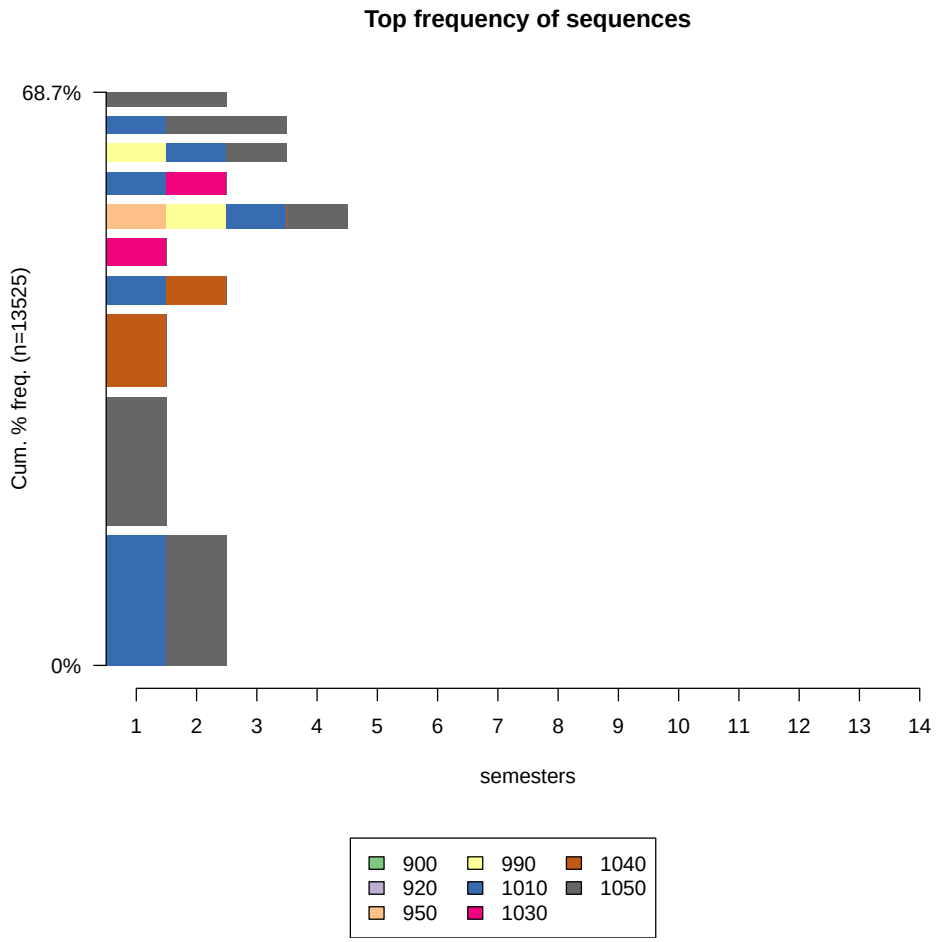
	Freq	Percent
1010	6439	42.29
1010-1050	3367	22.11
1010-1010	1147	7.53
1010-1040	765	5.02
1010-1030	523	3.43
1010-1050-1050	514	3.38
1010-1010-1050	362	2.38
1010-1010-1010	231	1.52
1010-1050-1040	181	1.19
1010-1010-1040	118	0.77



For students starting in math at 1010 or below there was a non-continuation rate, *in each course*, of between 30% and 50%. This massive attrition after one course almost makes the analysis of sequences irrelevant.

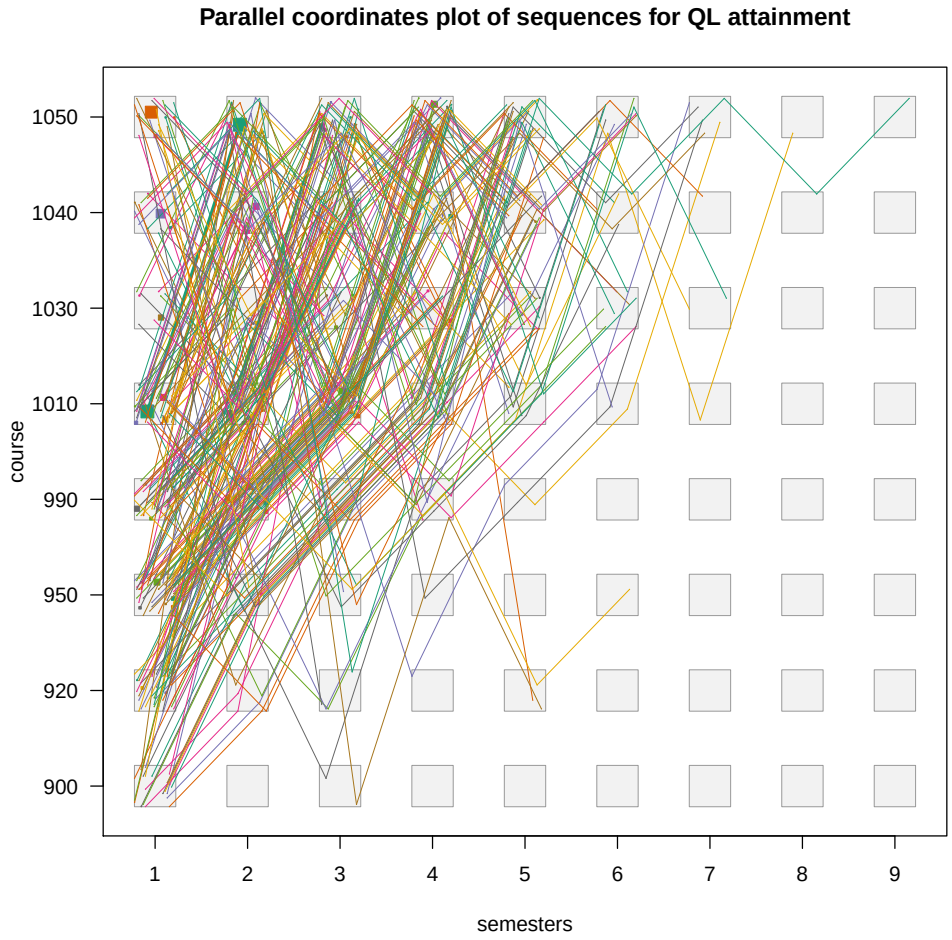
4.3 Counts of sequences that lead to QL

The following plots look at sequences a little differently. Rather than starting with a particular course, we look at QL completions—course sequences that conclude with passing grades in 1030, 1040 or 1050.

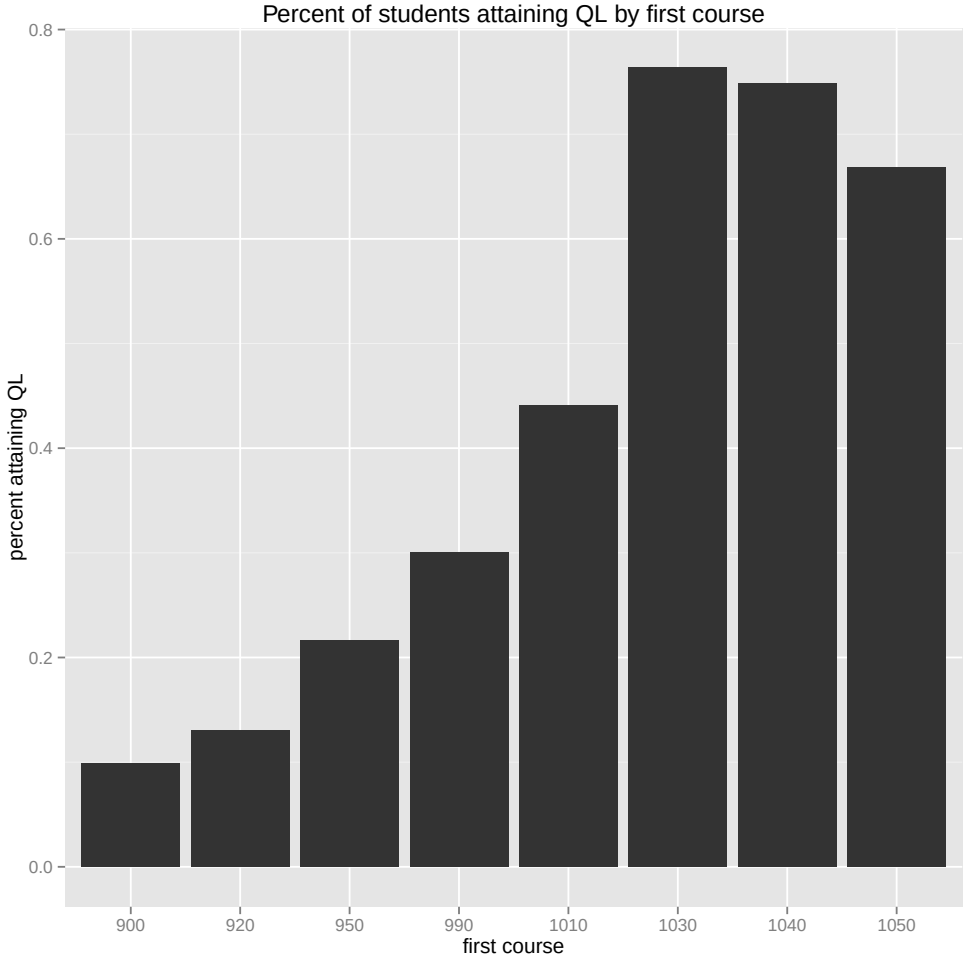


	Freq	Percent
1010-1050	2492	18.43
1050	2458	18.17
1040	1391	10.28
1010-1040	551	4.07
1030	526	3.89
950-990-1010-1050	470	3.48
1010-1030	434	3.21
990-1010-1050	357	2.64
1010-1050-1050	337	2.49
1050-1050	269	1.99
1010-1010-1050	221	1.63
950-990-1010-1040	177	1.31
1010-1050-1040	157	1.16
1050-1040	155	1.15
950-990-1010-1030	138	1.02
990-1010-1030	128	0.95
1040-1040	105	0.78
990-1010-1040	95	0.70
920-990-1010-1050	93	0.69
950-990-1010-1050-1050	91	0.67
1010-1010-1030	85	0.63

950-1010-1050	81	0.60
1040-1050	77	0.57
990-1010-1050-1050	65	0.48
1010-1010-1040	63	0.47
1010-1040-1040	62	0.46
1010-1050-1050-1050	62	0.46
1050-1050-1050	60	0.44
1010-1010-1050-1050	57	0.42
950-990-1010-1010-1050	51	0.38

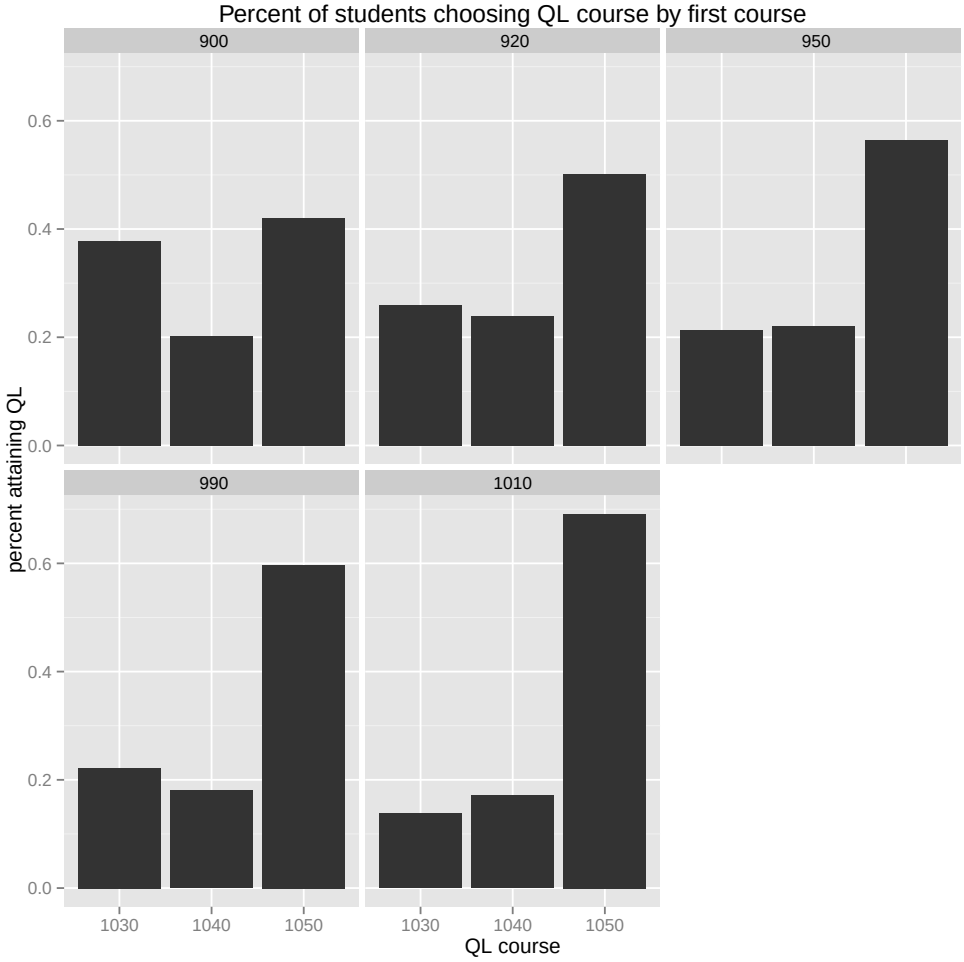


The strongest, or most populated, pathways are clearly the short ones: starting in 1040 or 1050, or 1010-1050. But what are the highest probability pathways? Let's look at percentage QL completion for a given first course.



Clearly, the shorter the path, the higher the probability of QL attainment. In the case of the shortest paths—i.e., starting and ending with 1030, 1040, or 1050—1050 is the clear loser, which is unfortunate given that, as we see above, the majority of students choose that course to satisfy QL.

What are the paths of those who end up completing in 1030 vs. 1040 vs 1050?



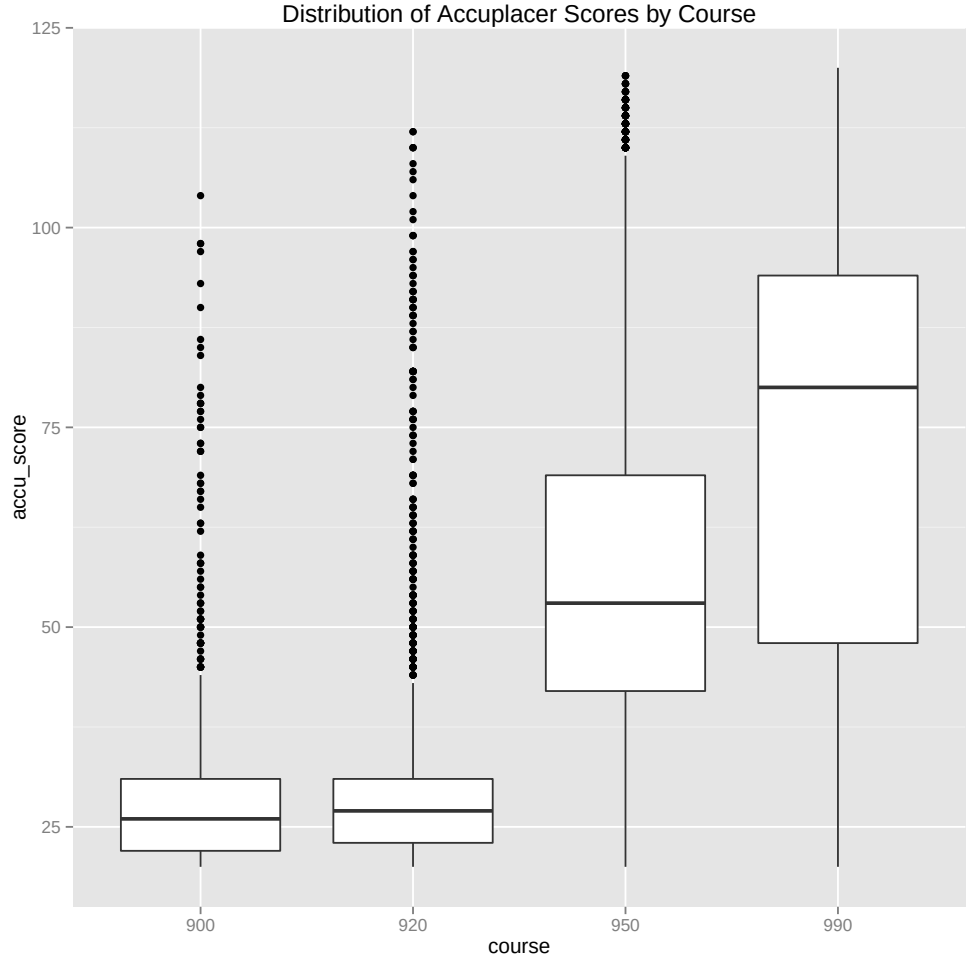
This plot suggests that students who start higher in math choose 1050 more frequently as their QL course.

4.4 Accuplacer scores?

The attrition in the developmental courses especially makes one wonder about placement. Are students prepared? Here are box plots of the distribution of accuplacer scores (the box represents the interquartile range, IQR, and the black dots represent outliers, observations that are 1.5 times or greater than the upper or lower limit of the IQR). The courses in the plot below were a student's first course, and the accuplacer scores were the arithmetic scores recorded prior and closest to that first course.

	id	last	course	term	accu_score	accu_date	act_score	act_date
1	12589	Spackman	950	201030	22	2010-04-14	NA	<NA>
2	32020	White	950	201020	24	2009-08-10	NA	<NA>
3	95532	Segura	950	201130	33	2010-07-12	NA	<NA>
4	97512	Cordova	950	200940	24	2009-08-17	NA	<NA>
5	106675	Willey	950	201340	34	2012-06-21	NA	<NA>
6	114848	Dorrell	950	201220	32	2011-12-06	NA	<NA>
7	124854	Anderson	950	201020	30	2009-04-22	NA	<NA>
8	125923	Dredge	950	201120	25	2010-09-02	16	2006-12-01
9	126347	Anderson	950	201120	29	2010-12-21	NA	<NA>
10	128500	Nish	950	201330	30	2012-11-09	NA	<NA>
11	129473	Nelson	950	201230	21	2011-12-05	NA	<NA>
12	130169	Walles	950	200940	30	2009-04-24	NA	<NA>

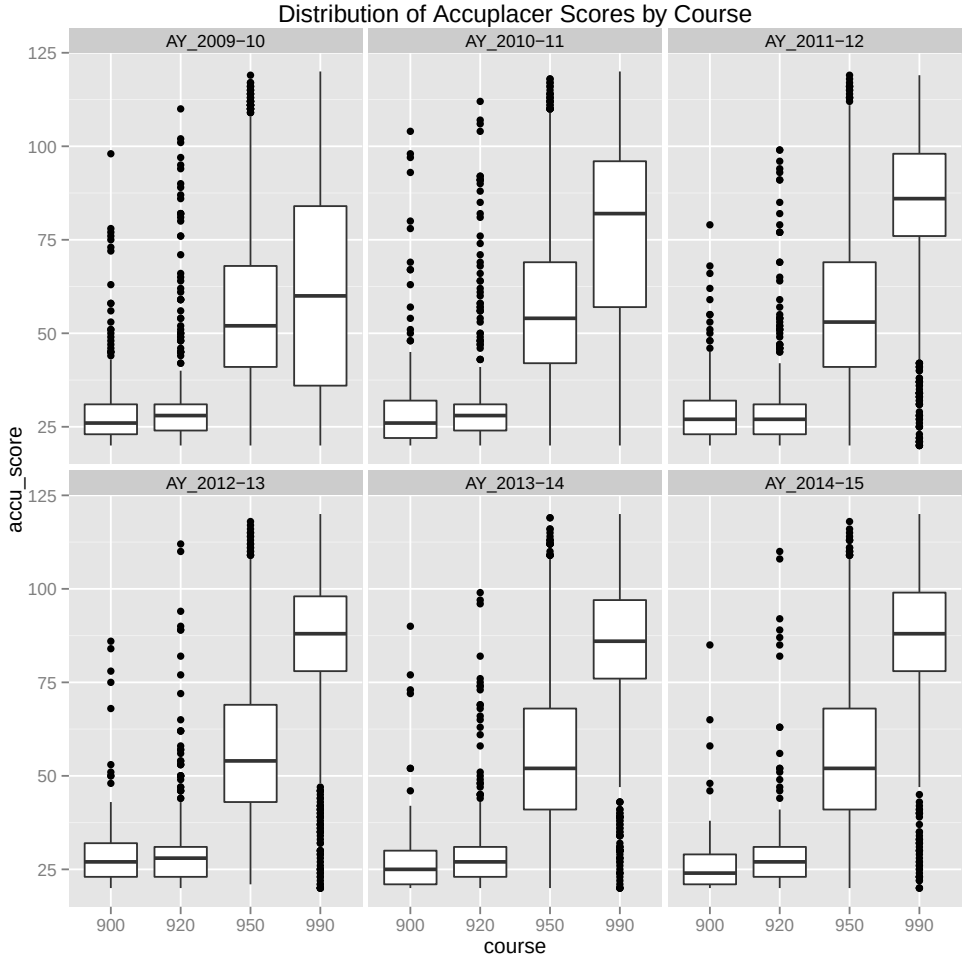
```
13 135019 Hermansen    950 201440         34 2014-08-07        NA       <NA>
14 137631    Daniel    950 201020         27 2010-01-04        NA       <NA>
15 145653   Hewlett    950 201040         24 2010-07-22        NA       <NA>
16 149278     Bills    950 201140         31 2007-07-10        NA       <NA>
17 152015     Lopez    950 201520         29 2014-12-17        NA       <NA>
18 157191     Stott    950 201020         30 2008-12-02        NA       <NA>
19 164138   Nielson    950 201420         28 2013-08-14        14 2001-10-01
20 173905   Roberts    950 201120         26 2004-08-18        NA       <NA>

      id      last course   term accu_score  accu_date act_score act_date
1   1644      Tate    990 200940         50 2009-03-11        NA     <NA>
2   5788      Gold    990 201040         56 2002-03-08        NA     <NA>
3   7398       Lee    990 200940         42 2008-10-10        NA     <NA>
4   8031     Deuel    990 200940         60 2008-01-02        NA     <NA>
5  11749    Gundry    990 201020         43 2000-08-22        NA     <NA>
6  11785     Salls    990 201030         67 2009-01-14        NA     <NA>
7  12226     Prada    990 201420         39 2013-11-20        NA     <NA>
8  12702 Orlandini    990 200940         58 2008-11-05        NA     <NA>
9  15922     Rabow    990 201120         56 2010-12-20        NA     <NA>
10 17533    Anasco    990 200940         53 2009-05-12        NA     <NA>
11 23904    Jacobs    990 201330         32 2005-12-15        NA     <NA>
12 27232  Rogonjic    990 201230         37 2012-01-03        NA     <NA>
13 28583     Scott    990 201220         54 2011-01-07        NA     <NA>
14 30539  Hamilton    990 201040         41 2007-08-14        NA     <NA>
15 30940   Raphael    990 201020         45 2007-07-16        NA     <NA>
16 31531     Smith    990 201040         57 2010-08-17        NA     <NA>
17 33717 Arredondo    990 200940         66 2008-12-01        NA     <NA>
18 37519 Rodriguez    990 201040         31 2004-04-01        NA     <NA>
19 42662     Noble    990 200940         63 2009-03-13        NA     <NA>
20 46827   Hausman    990 200940         45 2008-11-25        NA     <NA>
```



A couple of observations about the above plot. According to the Accuplacer placement guide, students with scores of 34 and below go into 900/20, those with scores between 35-74 go into 950, and those with scores of 75 and above (and elementary algebra scores of 40-54) go into 990. Higher scores in each course might not be a concern. Essentially everyone should be able to take 900/20, regardless of their score—there's no minimum. (More on that in a moment.) But in both 950 and 990 we see scores that are lower than should be possible. The first quartile is just about at the cutoff in 950 and is well below the cutoff in 990.

Still, the 950 and 990 distributions show the effects of the cutoff, with rising IQR. In contrast, 900 and 920 have no cutoff and it shows. Should they have a cutoff? Do these courses truly presume no arithmetic knowledge? There are community education options for math education, which are cheaper than SLCC. Should low scoring students be advised into those programs first?

Perhaps this distribution differs by year—if, for example, minimum accuplacer scores have been enforced more consistently in recent years.



It looks like the situation has improved in recent years—the boxes in 950 and 990 in the bottom row are more consistently separate than they are in the top—but in every year the minimum accuplacer score of 20 is represented in each course.

Maybe these students with low accuplacer scores are using their ACT scores to qualify instead? Here is a snapshot of rows for 950:

	id	last	course	term	accu_score	accu_date	act_score	act_date
1	12589	Spackman	950	201030	22	2010-04-14	NA	<NA>
2	32020	White	950	201020	24	2009-08-10	NA	<NA>
3	95532	Segura	950	201130	33	2010-07-12	NA	<NA>
4	97512	Cordova	950	200940	24	2009-08-17	NA	<NA>
5	106675	Willey	950	201340	34	2012-06-21	NA	<NA>
6	114848	Dorrell	950	201220	32	2011-12-06	NA	<NA>
7	124854	Anderson	950	201020	30	2009-04-22	NA	<NA>
8	125923	Dredge	950	201120	25	2010-09-02	16	2006-12-01
9	126347	Anderson	950	201120	29	2010-12-21	NA	<NA>
10	128500	Nish	950	201330	30	2012-11-09	NA	<NA>
11	129473	Nelson	950	201230	21	2011-12-05	NA	<NA>
12	130169	Walles	950	200940	30	2009-04-24	NA	<NA>
13	135019	Hermansen	950	201440	34	2014-08-07	NA	<NA>
14	137631	Daniel	950	201020	27	2010-01-04	NA	<NA>
15	145653	Hewlett	950	201040	24	2010-07-22	NA	<NA>
16	149278	Bills	950	201140	31	2007-07-10	NA	<NA>

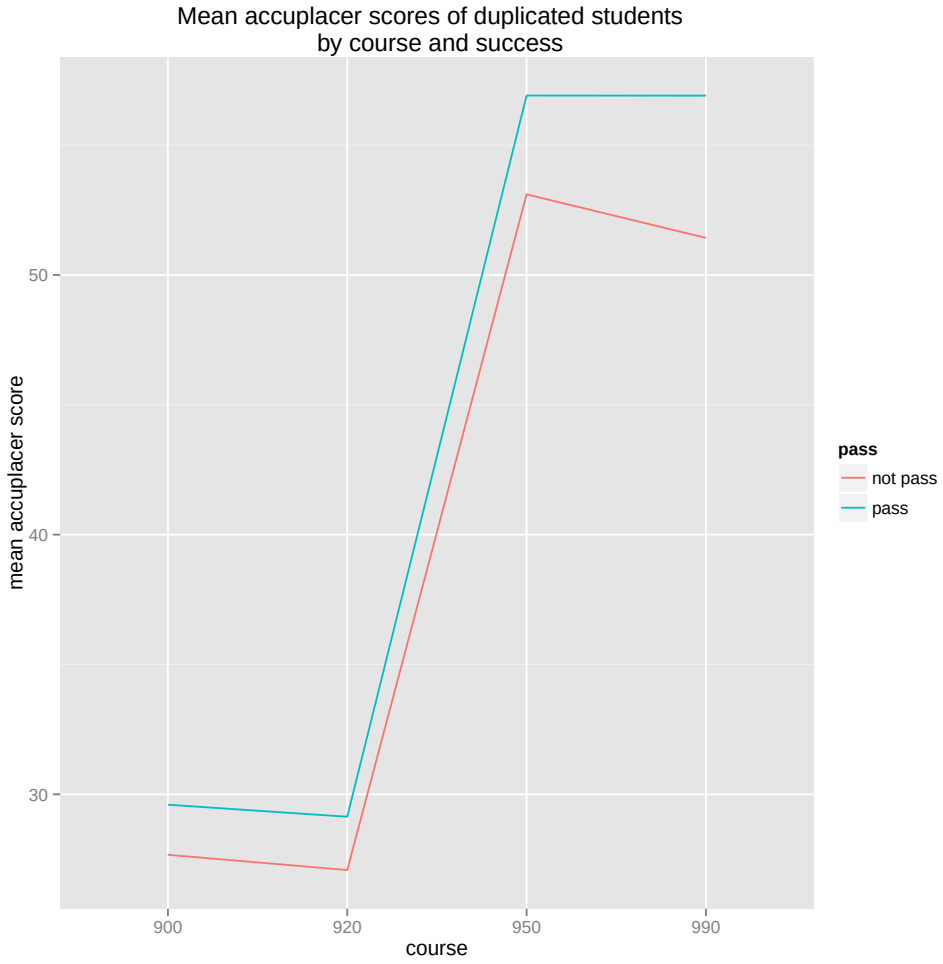
```
17 152015     Lopez    950 201520         29 2014-12-17        NA       <NA>
18 157191     Stott    950 201020         30 2008-12-02        NA       <NA>
19 164138   Nielson    950 201420         28 2013-08-14        14 2001-10-01
20 173905   Roberts    950 201120         26 2004-08-18        NA       <NA>
```

And here is a snapshot of rows for 950:

```
       id        last course   term accu_score  accu_date act_score   act_date
1   23904      Jacobs    990 201330         32 2005-12-15        NA       <NA>
2   37519   Rodriguez    990 201040         31 2004-04-01        NA       <NA>
3   47213     Zecevic    990 201030         29 2007-08-13        NA       <NA>
4   73228    Meinhart    990 201020         23 2008-04-25        NA       <NA>
5   73425      Toledo    990 201020         29 2008-08-13        NA       <NA>
6   73471      Toledo    990 200940         33 2008-07-25        NA       <NA>
7   80518     Nemelka    990 201020         27 2008-07-16        NA       <NA>
8   83996  Kranendonk    990 201030         23 2008-04-16        NA       <NA>
9   84900     Herrera    990 200940         27 2008-12-09        NA       <NA>
10  85258    Forsgren    990 200940         31 2008-12-08        16 2007-04-01
11  90707       Olsen    990 200940         34 2008-08-26        NA       <NA>
12  96258     Gilbert    990 201020         25 2008-06-09        NA       <NA>
13  96870    Gatherum    990 200940         27 2006-01-06        NA       <NA>
14  99102      Pluhar    990 201020         30 2008-10-13        NA       <NA>
15  99639      Gurule    990 200940         27 2008-12-16        NA       <NA>
16 100338         Lee    990 200940         32 2008-11-26        NA       <NA>
17 101529     Collier    990 200940         28 2008-12-02        NA       <NA>
18 103432       Wyatt    990 200940         33 2009-04-27        NA       <NA>
19 111979       Moore    990 201040         20 2006-03-20        NA       <NA>
20 115443 Christensen    990 200940         27 2008-11-18        NA       <NA>
```

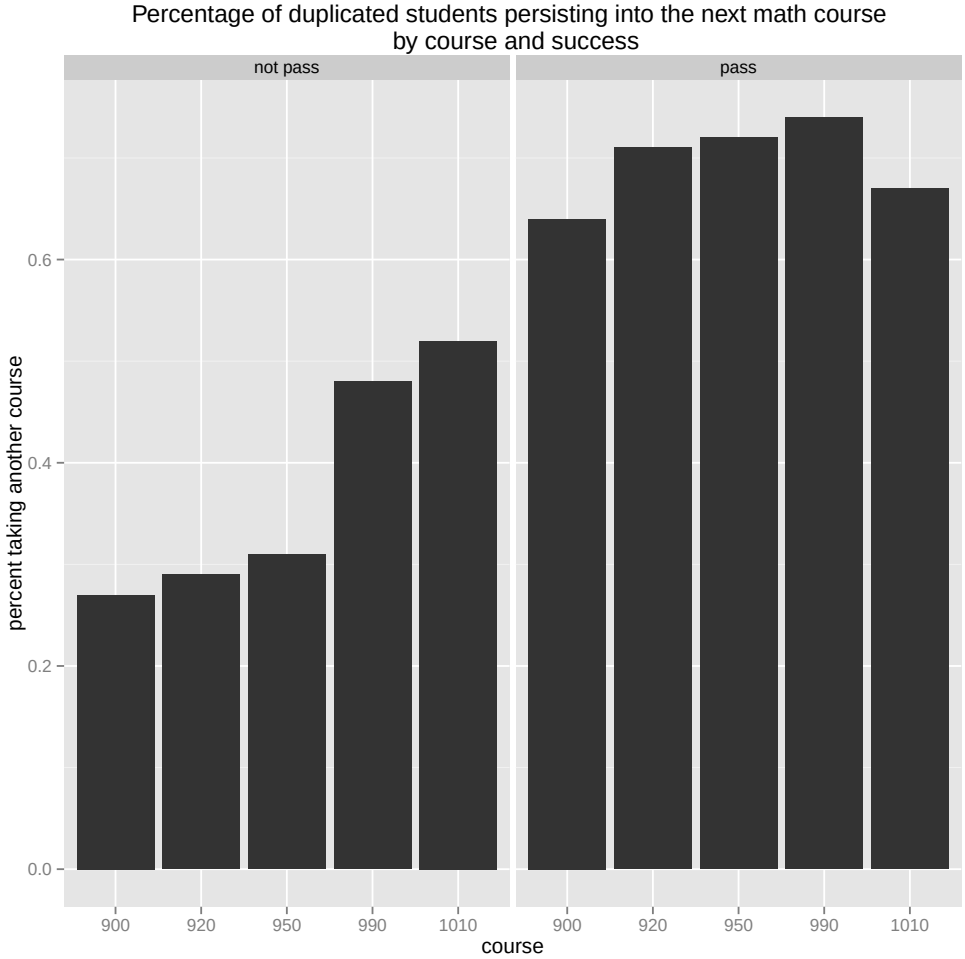
While ACT scores may in some cases explain low accuplacer scores in these courses, but that was clearly not usually the case, as it happened not once in the rows included above.

Was arithmetic accuplacer score associated with math course success? Yes, for each math class passing was associated with higher mean accuplacer score.



4.5 Biggest loss points

The following plot does not look at sequences but at continuation or persistence rates for students who took that course (which could be their first or one in a series).



In the above plot we see that students who were not successful in 990 or 1010 displayed greater persistence than those in 900/920/950 in persevering and retaking the course (around 50%). For those who were successful, a couple of questions: 1. Why the difference between 900 and 920? How might the differences between these courses explain the difference in student persistence? 920 is a 6 credit hour course. The difference in persistence is probably explained by differences in the students—those choosing the 6 credit hour course are more directed and motivated. 2. Why the low continuation rate in 1010? This despite the impressive tenacity of 1010 students who did not pass the course. The students who did not pass 1010 retake it at rates of about 50% whereas those who did pass retake it at rates of only 65%?