StudentPaths Visualization

March 21, 2023

Agenda

1. Student History and Trajectory
   a. Recap
   b. Sample Data
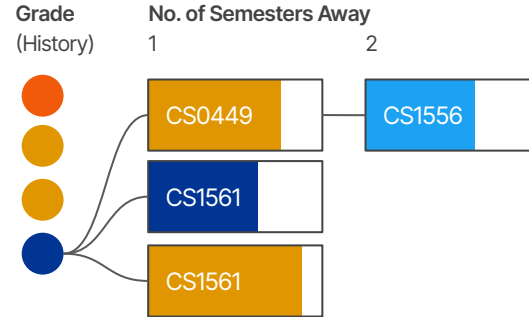   c. Adjustments
2. Clusters Visualization
   a. Challenges
   b. Revisions
   c. Observations

Student History and Trajectory

1. Historical GPA is displayed using colored circles, where 1 circle represents one course, and the color represents their final grade
2. Trajectory is visualized through a directed graph, where the color represents the probable grade, and the length represents the likelihood of that grade.
3. The different possible trajectories based on semester taken is then visualized using a line plot
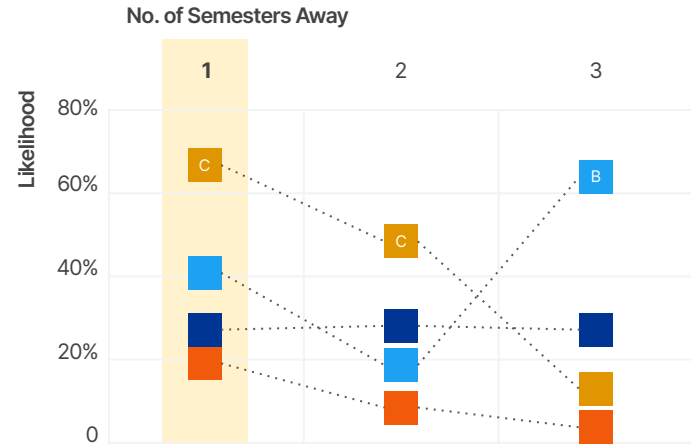
Student History and Trajectory - Sample Data

Observations

1. Accommodating empty semesters in between
2. Visualizing courses that were not taken or didn't have a letter grade

Student's Grades per Course, ordered by average GPA

Student History and Trajectory - Adjustments

1. Accommodating empty semesters

Student ID 2
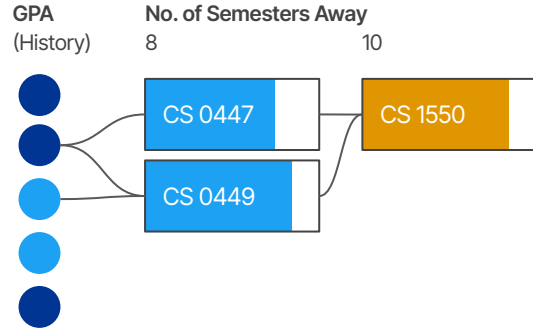
GPA	No. of Semesters Away
(History)	8 10

2. Visualizing courses that were not taken or didn't have a letter grade

Student ID 14

GPA	No. of Semesters Away
(History)	2 3
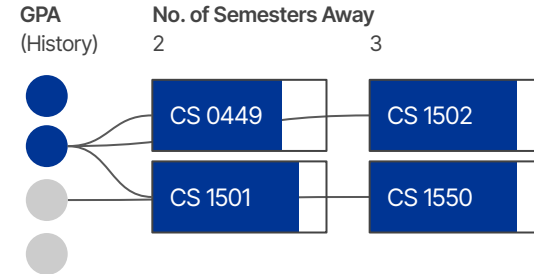
Assumptions made: for each student, I made assumptions about which courses they've taken or are yet to take

Cluster Visualization

Challenges:

1. Visualizing multiple categories on three axes will typically lead to an information-dense visualization that is challenging to interpret at first glance
2. The additional processing steps behind each axis adds to the cognitive load needed to extract insights from the visual
   a. Numbers need to be mapped to course codes
   b. A reader would have to consider that the grade and semester axis is affected by a stand-in value (e.g. -100, 2.5, or another number) that represents the student not having taken or not having a letter grade for that course

Cluster Visualization

Challenges:

1. Visualizing multiple categories on three axes will typically lead to an information-dense visualization that is challenging to interpret at first glance
2. The additional processing steps behind each axis adds to the cognitive load needed to extract insights from the visual
   a. Numbers need to be mapped to course codes
   b. A reader would have to consider that the grade and semester axis is affected by a stand-in value (e.g. -100, 2.5, or another number) that represents the student not having taken or not having a letter grade for that course
3. Trajectories are difficult to compare because the of 2b

Cluster Visualization

Visualization tweaks:

Minimize processing steps for the reader by simplifying the values

1. Instead of showing the semester as-is, assign a rank to the semester values for each cluster
2. Instead of showing the grade as-is, assign colors that differentiate below zero, above zero

Cluster	1	2	3	4	5	6	7	8
1	2 -22.1, -22.5	4 -15.0, -17.5	8 4.8, 3.1	5 -14.3, -17.5	7 -7.1, -12.8	6 -7.2, -13.3	1 -100.0, -100.0	3 -20.6, -23.2
2	2 -24.4, -25.5	1 -28.8, -31.1	3 -11.3, -14.6	4 -8.5, -13.4	5 5.4, -1.2	7 11.1, 2.8	6 10.9, 2.8	8 12.6, 2.6
3	2 -43.5, -42.9	5 -31.4, -31.0	6 -28.6, -28.8	4 -32.3, -33.8	3 -37.1, -39.8	7 -24.0, -27.4	8 4.6, 1.9	1 -100.0, -100.0
4	7 -20.0, -18.6	6 -25.9, -25.9	8 3.3, 2.3	5 -42.8, -44.3	3 -82.1, -82.8	4 -59.0, -60.5	1 -100.0, -100.0	1 -100.0, -100.0
5	1 -100.0, -100.0	2 -99.1, -99.1	3 -72.3, -71.5	4 -49.7, -48.4	5 -35.1, -34.3	6 -18.7, -17.4	7 2.2, 2.8	8 3.2, 2.8
6	1 -17.3, -15.3	3 1.7, 3.2	2 -0.9, 0.0	4 3.1, 3.2	5 4.3, 3.1	6 5.1, 2.9	7 5.6, 2.9	8 6.9, 2.7

Cluster Visualization

Cluster	1	2	3	4	5	6	7	8
1	2	4	8	5	7	6	1	3
2	2	1	3	4	5	7	6	8
3	2	5	6	4	3	7	8	1
4	7	6	8	5	3	4	1	1
5	1	2	3	4	5	6	7	8
6	1	3	2	4	5	6	7	8

Course (column header spanning columns 1–8)

- Color: GPA
  - Darker orange: more -100s
  - Lighter blue: less -100s
- Text: Semester rank
  - 1: relatively lower semester value
  - 8: relatively higher semester value

Observations

- Cluster 6 `has the least -100 values
- Cluster 5 semester rank follow the course number exactly, and has increasingly less -100 values
- Cluster 2 is somewhat similar to cluster 5 in terms of having increasingly less -100 values, and sem rank is just slightly varied from course number
- Course 4 is consistently taken in the 4th or 5th ranked semester

Cluster Visualization

Cluster	1	2	3	4	5	6	7	8
1	2	4	8	5	7	6	1	3
2	2	1	3	4	5	7	6	8
3	2	5	6	4	3	7	8	1
4	7	6	8	5	3	4	1	1
5	1	2	3	4	5	6	7	8
6	1	3	2	4	5	6	7	8

Course (column header across top)

- Color: GPA
  - Darker orange: more -100s
  - Lighter blue: less -100s
- Text: Semester rank
  - 1: relatively lower semester value
  - 8: relatively higher semester value

Weaknesses

- Not interpretable when later courses have earlier semester ranks (e.g. clusters 4 and 1)
  - A semester rank of 1 likely represents a high value of -100s
- Unable to visualize what the exact semester and grades are

Further Improvement

- We may consider using *null* instead of -100, and separating the amount and percentage of records with these values as another aggregation

End