Factors Impacting Long-Term Outcomes in Higher Education

Alex Axthelm & Adam Hilgenkamp

Abstract

The purpose of this research is to illuminate patterns in long term outcomes for students at Indiana colleges and universities. In addition to the general patterns in long term outcomes, we will explore the differences in the patterns for transfer students, changes in GPA and changes in major selection. We will use data provided by the Indiana Commission for Higher Education to study the patterns and to discuss how different events may affect the long term outcome for a student.

Introduction

Tracking student transfer is important for identifying the complex outcomes of students in the modern higher education landscape. 1 in 4 Indiana College students will transfer to another university within 6 years of starting (Indiana Commission for Higher Education, 2018). Transfer between colleges and universities is an important aspect of predicting student success, especially for minority, low-income, and non-traditional students (Grubb, 1991; Dougherty, 2006). For these students community college allows for lower barrier to entry to higher education, which can then act as a springboard to a four-year degree. In fact, students transferring a from a 2 year to a 4-year institution were twice as likely to graduate with a degree after transferring than those moving from a 4 to a 2-year program (Indiana Commission for Higher Education, 2018).

Preliminary analysis by the Commission suggests that there is an important connection between stopouts (when a student has an extended period without enrollment), and transfer (when a student changes institutions). This connection can be contextualized through the common factors that lead to both events, namely large extra-curricular life events, which lead a student to reconsider the value of their academic course.

In this paper we will begin by looking at the relationship between transfers and the long term completion rates for students at Indiana Institutions. We will then dig deeper by visualizing and studying the patterns in GPA changes. Finally, we will look at how students transfer between majors. Specifically, reviewing a student beginning their academic career as an “undecided” major.

Background

This work will build upon an analysis by Alex Axthelm for the I590, Network Science (Spring 2018) class, but with a different focus. Whereas that analysis centered on network connectivity, and distance/relation tradeoffs, this will focus on long term outcomes of students.

The key visualizations that will be developed will be a series of Sankey Diagrams, illustrating the overall flow of students in particular cohorts, with each column on the diagram representing an academic year. Each node within a column will represent a student cohort of interest.

In the data set there are some students that are enrolled at multiple institutions simultaneously. We will not be including these students in this study as we want to focus on transfers, GPA and program as factors impacting long term outcomes.

The dataset will allow us to parse out additional information about the students, by looking at breakdowns such as demographic characteristics, financial aid eligibility, and other student-level information. In the process of collecting the data and cleaning it for use in our visualizations we encountered some data quality issues that cause our number to be slightly different than the official numbers reported by the Indiana Commission for Higher Education. However, the numbers are generally represent the overall trends and breakdowns seen in the official results.

Objectives and Research Questions

The main research goals for this analysis are to consider:

- What are the patterns for students who transfer to a different institution in Indiana?
- Is a significant GPA change a factor that has a noticeable influence on long term completion?
- What are the patterns for major selection and changes in academic programs?
- What does the path of a student that begins as an "undecided" major look like?

Choosing a Visualization

With the previously stated research questions in mind we needed to determine how we wanted to display our results. To do this we needed to think critically about the data that we were looking at. The data and our analysis has three key components. The first is that we want to look at the flow between students between cohorts. Second, we want to be able to see these trends over a period of time. Lastly we want to highlight the behavior of a particular cohort. To accommodate these requirements there are several options, each with their own drawbacks and benefits. One option would be to use a waterfall chart to represent the changes from one point in time to another. This has the benefit of visualizing the students moving in and out of a particular cohort but the key drawback is that it does not scale well to visualize the larger system. This could prevent us from seeing the larger trends in the data. Another option would be to look at a one hundred percent stacked bar chart. This would do a good job of showing the differences in the cohorts between time periods. However, it would not show what is moving in and out of each time period bar. Thus we would not be able to see the flow of the students. What we decided to use was a Sankey diagram. Specifically the Sankey diagram available in the plotly R package. This chart does a great job of showing the flow across years and is clear about what is moving in and out. The main drawback of this option is that it is showing movement at a high level so it is difficult to track the path of any individual student. Still, it is accurate at showing the general flow for cohorts in the data.

Cleaning the Data

Once we selected a visualization technique we needed to clean the data in order to make sure we could provide the format needed by the chart functions and to ensure that our results were accurate. There were two main challenges that needed to be overcome. The first is that the data contains a lot of noise and is also spread across many tables. In order to derive the cohorts we wanted to look at we needed to join together several tables which we did using data frames in R. We then needed to exclude factors that could skew the results. We grouped the data down to only look at academic years in order to avoid the noise of

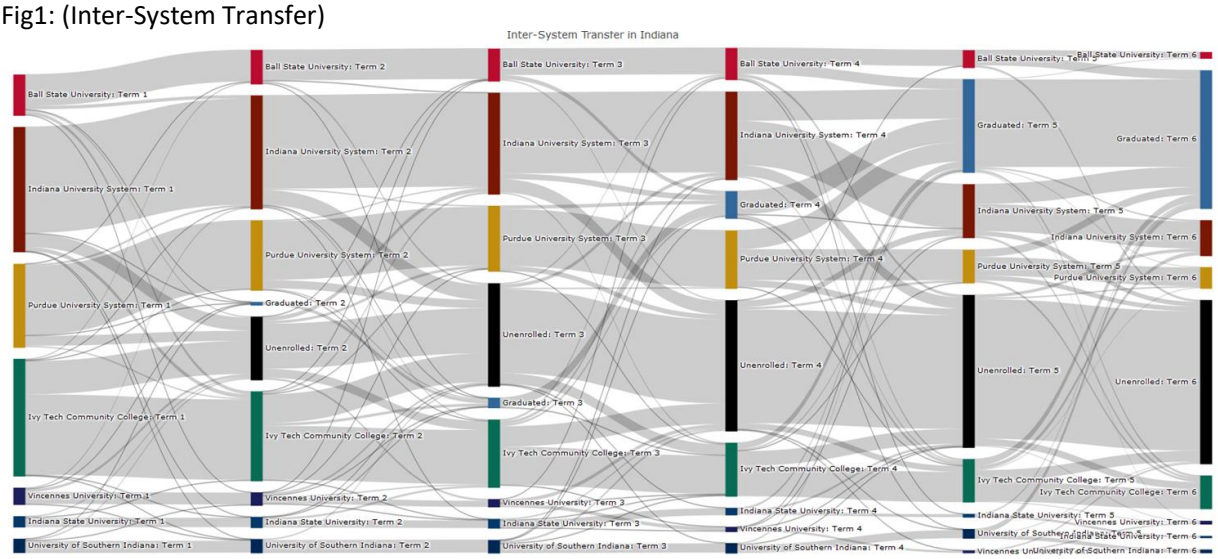
students taking breaks between semesters. The next challenge was getting the data into a format that would be accepted by the plotly Sankey function. We needed to create source and target columns in our data to tell plotly where a particular transition is taking place. Once we had current academic year variables and next year variables we could show our cohorts in the diagram. In addition to creating source and target columns ordering the data in a specific way when the plotly Sankey diagram renders is not very well supported and is challenging with the current version of the module. Once we were able to work through the data and plotting challenges we could then view our results.

Results

The data cleaning and manipulation consumed a majority of the time taken to complete this research but it payed off when we were able to take a look at the resulting visualizations.

The first question that we took a look at was inter-system transfer in Indiana. The aim of this visualization was to highlight the patterns in the flow of students from one university to another in Indiana. What we found is that Ivy Tech students are far more likely to become unenrolled in the next term without earning a degree. As seen in fig1 there are smaller lines showing the transfer between institutions but a large flow is coming from Ivy Tech and moving to the unenrolled cohort. Transfer was also studied within a university system. We looked at the transfer patterns of students in the Indiana University system. There were not as many noticeable trends at this level of detail as most of the movement was small between schools.

Fig1: (Inter-System Transfer)



The next question that we studied was the flow of students categorized in particular GPA cohorts. We were able to look at the movement within GPA cohorts as well as the long term completion rates of these groups. We found that a significant number of students who have a GPA of less than 2.0 much more likely to be unenrolled at the end of the first academic year as seen in fig2. This may confirm the general assumptions associated with GPA's but it is interesting to see the immediacy of that impact. Another interesting insight is the on-time graduation rates. Students who have a GPA of 3.0 or higher have a larger share of their cohort graduating on time.

The last question we wanted to look at was to better understand the movement between academic programs. We first looked at the changes for undergraduate students seeking bachelor's degrees. When looking at the flow for all students there was a lot of movement in the early years and movement seemed to settle down in year 3. To take a closer look at academic program movement we then looked at a

particular institution. When filtering our data to show Indiana University – Bloomington we can begin to see the most popular majors, on-time graduation rates and the movement between programs. Since we used a small sample the data may not exactly reflect the student population but it gives a general representation of the movement. The most interesting zoom that we did was to look at students who begin their academic career as an "undecided" major. As seen in fig3 there is still a significant percentage of students who remain undecided after the first year. The concerning trends appear when looking at on-time completion rates. Undecided students seem to have a smaller group that graduate in four years and a large portion are unenrolled after four years. In the following section we will briefly discuss the further research that we would like to do on this particular trend and the other trends that we observed.

Fig2: (GPA Flow)

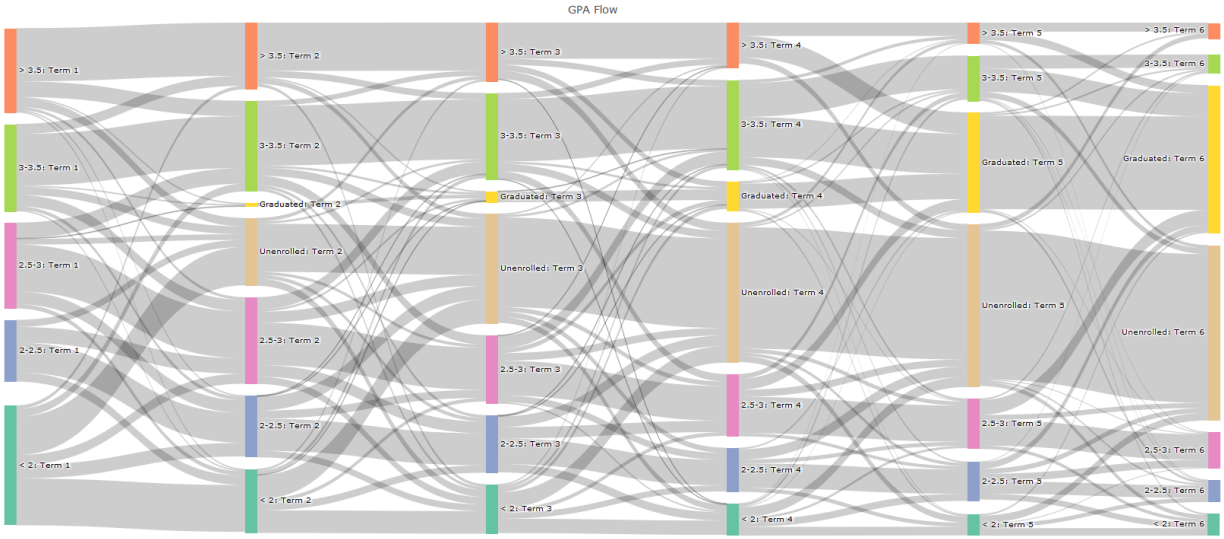
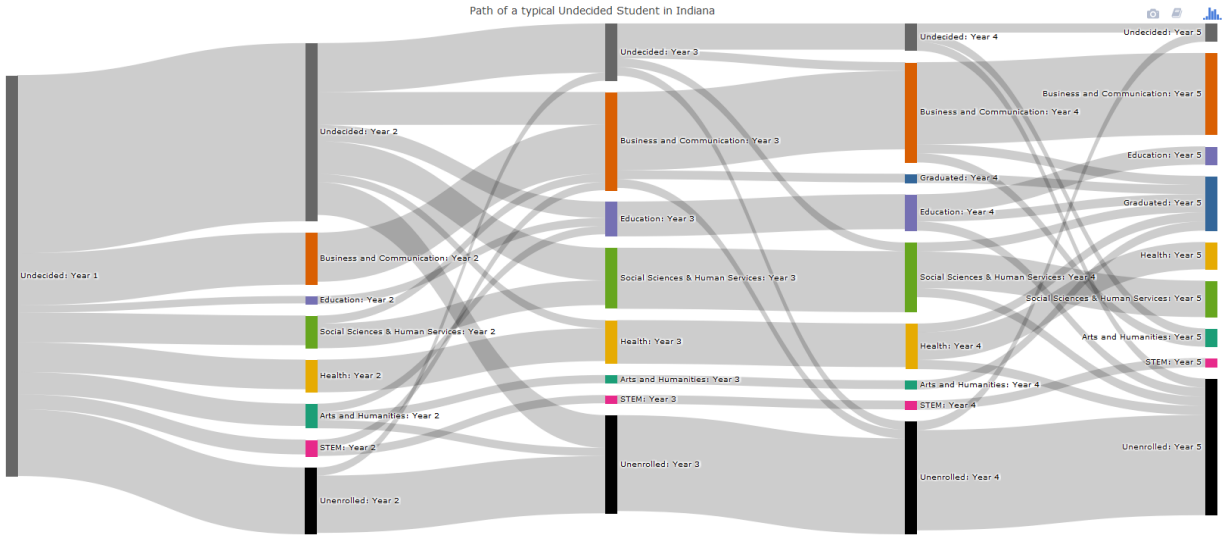


Fig3: (Path of an Undecided Student)



Conclusion & Further Research

For next steps the first thing that comes to mind is improving the data quality. When performing this analysis we had to work through some issues and further validation of the data is required before moving forward. As a proof of concept we were able to get some interesting results with the data that we have but we would like to also expand this to look at a larger sample size. Some interesting trends were uncovered as a part of this research. We would like to further look at GPA trends and the trends associated with students who begin their academic career as undecided majors. Once we have been able to further research these trends we think it would be interesting to explore the impacts of financial aid as well as looking at the correlation between dimensions in this dataset. In conclusion, visualizing the Indiana higher education student dataset through Sankey diagrams highlighted some interesting trends. Over time we were able to look at the long term impacts of transferring to a different institution, changes in a student's GPA and the impact of changing majors. These preliminary insights will be helpful in identifying places to dig deeper to understand what students need to be successful and how the state may be able to enable them to achieve the education that they set out to get.

Other Research and Sources

Allen, J., Robbins, S. B., Casillas, A., & Oh, I. (2008). Third-year College Retention and Transfer: Effects of Academic Performance, Motivation, and Social Connectedness. Research in Higher Education, 49(7), 647-664. doi:10.1007/s11162-008-9098-3

Bradburn, E. M., G, D., & Peng, S. (2001). Community College Transfer Rates to 4-Year Institutions Using Alternative Definitions of Transfer. Research and Development Report.
Publisher: ERIC

Diaz, P. E. (1992). Effects Of Transfer On Academic Performance Of Community College Students At The Four-Year Institution. Community Junior College Research Quarterly of Research and Practice, 16(3), 279-291. doi:10.1080/0361697920160307

Dougherty, K. J., & Kienzl, G. S. (2006). Its Not Enough to Get Through the Open Door: Inequalities by Social Background in Transfer from Community Colleges to Four-Year Colleges. Teachers College Record, 108(3), 452-487. doi:10.1111/j.1467-9620.2006.00658.x

Flaga, C. T. (2006). The process of transition for community college transfer students. Community College Journal of Research and Practice, 30.

Grubb, W. N. (1991). The Decline of Community College Transfer Rates. The Journal of Higher Education, 62(2), 194-222. doi:10.1080/00221546.1991.11774115

Hills, J. R. (1965). Transfer Shock. The Journal of Experimental Education, 33(3), 201-215. doi:10.1080/00220973.1965.11010875

Indiana Commission for Higher Education. (2018, February). College Transfer & Student Success. Retrieved from https://www.in.gov/che/files/2018_Transfer_Report_2-12-18 FINAL.PDF

Jenkins, D., & Fink, J. (2016). Tracking Transfer: New Measures of Institutional and State Effectiveness in Helping Community College Students Attain Bachelor's Degrees. Community College Research Center, Teachers College, Columbia University.

Jr, J. C., & Harrington, A. R. (2002). Academic Performance Of Community College Transfer Students And "native" Students At A Large State University. Community College Journal of Research and Practice, 26(5), 415-430. doi:10.1080/02776770290041774

Kuh, G. D., Cruce, T. M., Shoup, R., Kinzie, J., & Gonyea, R. M. (2008). Unmasking the Effects of Student Engagement on First-Year College Grades and Persistence. The Journal of Higher Education, 79(5), 540-563. doi:10.1353/jhe.0.0019

Laanan, F. S. (1996). Making the Transition: Understanding the Adjustment Process of Community College Transfer Students. Community College Review, 23(4), 69-84. doi:10.1177/009155219602300407

Townsend, B. K. (2001). Redefining the Community College Transfer Mission. Community College Review, 29(2), 29-42. doi:10.1177/009155210102900203

Wang, X. (2009). Baccalaureate Attainment and College Persistence of Community College Transfer Students at Four-Year Institutions. Research in Higher Education, 50(6), 570-588. doi:10.1007/s11162-009-9133-z