Government Department Curriculum Mapping Prototype[1]
Spring/Summer 2017

Liz Davis (G-3)
Olivia Goldberg (UG-17)
Max Gopelrud (G-3)
David Ifkovits (G-4)
Gabriel Katsh (Lecturer)
Caroline Marshall (UG 17)
Dustin Tingley (Director of Graduate Studies, Government Department)

Overview

This specific project was initiated by Harvard's Department of Government to start a process and conversation around our curriculum. It is an example of a broader exploration into curriculum mapping supported by the Vice Provost for Advances in Learning (VPAL) via its Harvard Initiative for Learning and Teaching (HILT) and the VPAL-Research Group organizations.

The particular deliverables discussed below should be read as an 'experiment': what data could a collaboration of students (graduate and undergraduate), faculty, staff, and institutional actors amass, analyze, and render explorable?

We hired two undergraduate research assistants (RAs) and one Lecturer to map out the learning objectives in the Government Department undergraduate curriculum, using the following approach:

- *They compiled a comprehensive list of courses offered for Government Department concentration credit in recent years.*
- *They then searched Harvard websites to obtain syllabi for the aforementioned courses.*
- *An RA read each syllabus to determine the learning objectives for the course. For syllabi lacking explicit learning objectives, an RA attempted to ascertain the implied learning objectives from what was stated on the syllabus and then contacted the instructors to discuss their goals when designing the course.*

[1] We thank the Social Science Dean's innovation fund for providing resources, Jennifer Hochschild, Danielle Allen, and Cheryl Welch for support, Karen Kaletka and Diana Wojcik for administrative support. We are especially grateful to Karen Pearce and her FAS Institutional Research team for providing enrollment data. We thank IQSS for supporting the R Shiny deployment.

- *The RAs then drafted sample plans of study linking the course learning objectives with topics of interest to undergraduate concentrators. These drafts are available at https://scholar.harvard.edu/files/dtingley/files/govcurricmemo.pdf.*

Both the database of learning objectives and the repository of syllabi are available for future use as appropriate.

We also hired three graduate students to visualize the data collected from the FAS and the undergraduate RA's. Along with Tingley, they produced the code necessary for the interactive website at https://hmdc.shinyapps.io/govmap/

- They transformed the data provided by the FAS Institutional Research that had the Government Department course enrollment history for every student since 2002 that took a Government Department course.
- Using R, they wrote scripts to produce a variety of graphical depictions of the FAS data including analyses of enrollment similarities and trends over time.
- Using R, they wrote scripts to conduct Natural Language Processing analyses on the learning objectives (topic modelling, hiearchial clustering)
- Using R, they stood up an R Shiny website that enables all of these in an interactive manner

Summary Cost of Project

Total Cost: ~$4250 and ~30 hours of Tingley's time.

Course Learning Objective Data

Our goal in this project was to analyze the courses available to Government Department undergraduates to determine the types of skills and subject-matter knowledge that concentrators were gaining from their required courses to ascertain whether the courses fit together into a coherent plan of study. To that end, our two undergraduate research assistants (RAs) carefully went through all course offerings in the Government Department, as well as cross-listed courses at the College and in other Harvard schools (e.g., the Harvard Kennedy School), and mapped out the learning objectives of each course that had a syllabus accessible through my.harvard.edu or the Harvard Kennedy School website.

The RAs first looked at the syllabus of each course to determine if the learning objectives were explicitly listed and evident to prospective students. If not, the RAs attempted to ascertain the implied learning objectives from the syllabus, and then contacted the course instructors to see if their judgment of the learning objectives matched what the course instructors had in mind when

designing the courses. The RAs completed this curriculum mapping for 187 courses (132 in the Government Department, 12 in other FAS departments, and 43 at HKS), and they communicated with about half the instructors to discuss the pedagogical goals in mind when the syllabi were written.[2]

While completing this curriculum mapping, the RAs compiled a database of syllabi for the courses being analyzed, which is now available for further research by members of the Government Department.

Finally, the RAs each wrote a **memorandum** summarizing their work on this project and discussing their own experience designing a coherent plan of study for themselves while students in the Government Department. The memos also provide sample plans of study for future students based on their academic interests and future career goals (e.g., an interest in international relations, future plan to attend law school, etc.). These memos can be used to guide future concentrators in selecting courses, and may be helpful for faculty in ensuring that the Government Department offers a sufficient range of courses to allow students with a variety of different academic interests and career plans to benefit from concentrating in Government.

Course Enrollment Data

To obtain data on Government Department enrollments, Tingley contacted Karen Pearce in FAS institutional research with a proposed data format. We were able to easily get de-identified data for every student who took a Gov class from 2002 to the current time period. This includes 11,931 students. The average number of courses taken by students in this period was 3.46; the median was 2.

Course Enrollment and Learning Objective Visualization

We then transformed this data to make a variety of interactive visualizations deployed in an R Shiny app at https://hmdc.shinyapps.io/govmap/. The left side of the app breaks apart different ways of viewing the data into 'tabs'.

Course Data Tab

There are four separate graph options within the "Course Data" tab. The "Course Popularity (1)" graph allows the user to subset students on the basis of how many courses they took and which years they were enrolling in courses, and it displays the most commonly taken courses for this student group. The "Course Popularity (2)" graph also allows the user to subset

[2] The automated analyses of learning objectives in the GovMap dashboard focus for now on Government Department courses.

to students active within a certain time-frame, and it displays, for a single, user-selected course, the probability that a student takes it, given they have taken *x* Gov courses. The first two graphs can illustrate the difference in course selection between concentrators and non-concentrators, and they can show how course selection has changed for these groups across time.

The third graph, “Course Enrollment” allows the user to select a course, and it shows the total enrollment in that course across terms, for all students.

The fourth tab, “Compare Courses,” allows the user to examine how courses are related to each other over time. The user selects two courses to examine and can subset students by the total number of courses they took. The user can select “Took Both” to see the probability that a student took both courses; “Took Both Given One” to see the probability that a student took both if they’re in the set that took at least one of the specified courses; “Took C1 Cond. on Taking C2” to see the probability that a student took the first course specified, given they took the second; and vice versa for “Took C2 Cond. on Taking C1.” Each probability is calculated for the set of students in a four-year window, starting at the year displayed (so the value for 2013 is calculated using data for students active from 2013 to 2016). This set of graphs can highlight how the relations between courses changes over time, especially when paired with each other. For instance, plotting each of these probabilities for Gov 1295 and Gov 20 reveals that students taking Gov 1295 are increasingly likely to also take Gov 20, while the reverse does not hold true. Gov 10 and Gov 30 also have an interesting pattern: fewer students took both courses in 2007 to 2012.

Network Data

The Network Data module visualizes how students move throughout the GOV curriculum and which courses they tend to take together. There are two ways to visualize these relationships, the “Course Network” option and the “Course Flow” option. Users can choose a class from the dropdown menu and the plots are then rendered for all students who took this class. The “cutoff” function is crucial to make the plots more interpretable. It subsets the data to classes that were taken by at least x students, where x is the cutoff. Picking a low cutoff value vastly increases the number of nodes and edges, making the plot potentially hard to interpret.

“Course Network” creates a so called forced network plot, depicting not only which courses are taken together, but also how *central* courses are to the curriculum. The closer a dot is to the center, the more it acts as a “hub”, i.e. a course that is included in many different pairings of courses. In other words: Courses at the edge of the graphic are not taken in unison with many other GOV courses, while more central courses are often requirements or quasi-requirements, forming are a shared experience for many GOV concentrators. Note that the course that is selected by the user is always in the center by default since it is taken by all students in the selected subsample. Lastly, the line width of the edges also has a meaning. It is thicker if more students take a certain pair of classes together.

“Course Flow” is a so called Sankey plot. Unlike in the network plot, the position of a class in the graph does not have a conceptual meaning. Rather, classes are ordered by a combination of name and number of students in pairings. The advantage of this plot is that it is easier to see differences in magnitude across flows. Wider bands represent a greater number of

students taking a certain pair of classes together. Important note: There is no temporal order in this plot. A band from left to right includes all students that took a certain pair of courses, it does not mean that students first took the course on the left, then the course on the right.

Learning objectives

Faculty and undergraduate students were tasked with describing the learning objectives of GOV courses, and this module uses text analysis to explore patterns in these documents. The raw data (texts) can be seen in the first tab. The second tab is a "word cloud" tool that represents the overall frequency of words in the texts.

The "dendogram" and "topic cluster" render two different text analysis algorithms to visualize how identify courses with similar learning objectives. The basic idea is that courses that share a large number of words in their learning objective texts can be considered as similar. The dendogram represents a plot with all courses and connects the more similar courses with lines. This plots the results of simple hiearchial clustering. Alternatively, we also implemented the Latent Dirichelet Allocation topic model. The user selects the number of topics to estimate, with a higher number being more granular. Then using the "Which topic to show" slider, the user can then look at each topic and see what words are most common for a given topic. These analyses are largely exploratory.

Conclusion

With collaboration from multiple University units (e.g., FAS Institutional Research) the Department of Government conducted a pilot curriculum mapping experiment. With relatively little expense we collected data and analyzed it quantitatively and qualitatively. The code base for the interactive app would require slight reworking in order to support a similar effort in another Department (~3 hours). Given the experimental nature of the project, there are of course more things that could be done in the future.