iCourse Investigation

Will Morgan

21 June, 2018

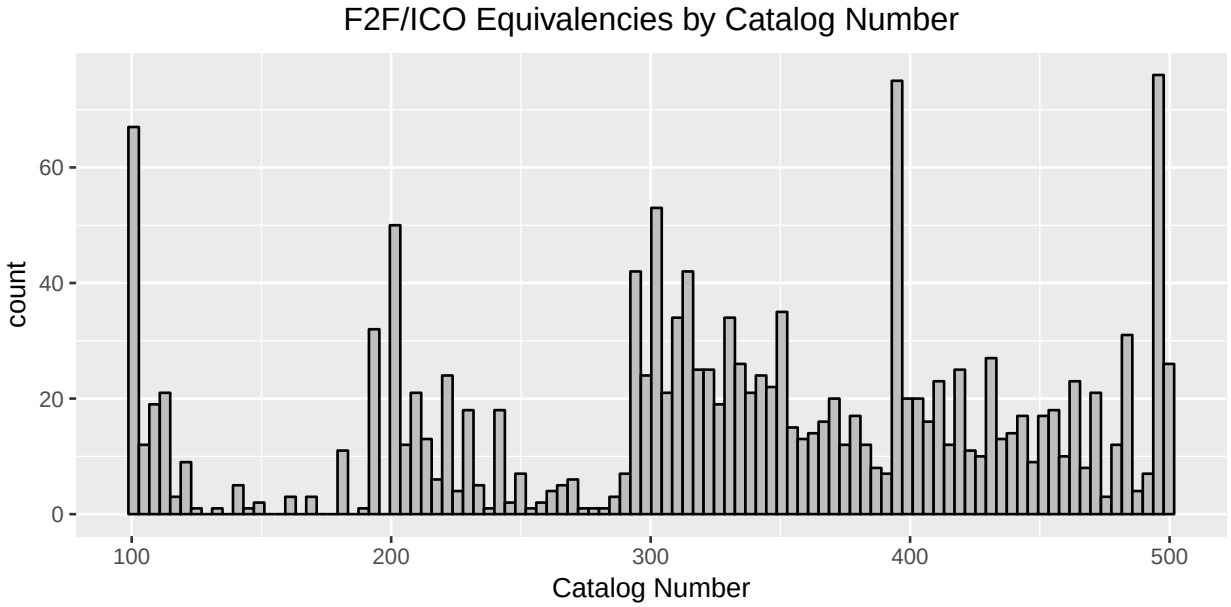
Purpose and Outline

This is an exploratory investigation into the characteristics of iCourses and is meant to be a log of the questions that arise as we delve further into the data. For now we focus on answering only a handful of questions we think will be informative down the line when we begin looking into student course selection. The current list of questions is as follows:

1. How many and what types of courses exist in both modalities?
2. What is the proportion of face-to-face (F2F) courses that have iCourse (ICO) equivalents?
    - How does this change when we disaggregate by college? program?
3. What proportion of iCourses are upper-division?
4. What proportion of on-campus students' courses are taken as iCourses?
    - Does this vary w.r.t. other variables of interest? (GPA, college, etc.)
5. Is there a relationship between iCourse enrollment and the amount of credits a student is taking?
6. What proportion of iCourses are being taken to fulfill core degree requirements?

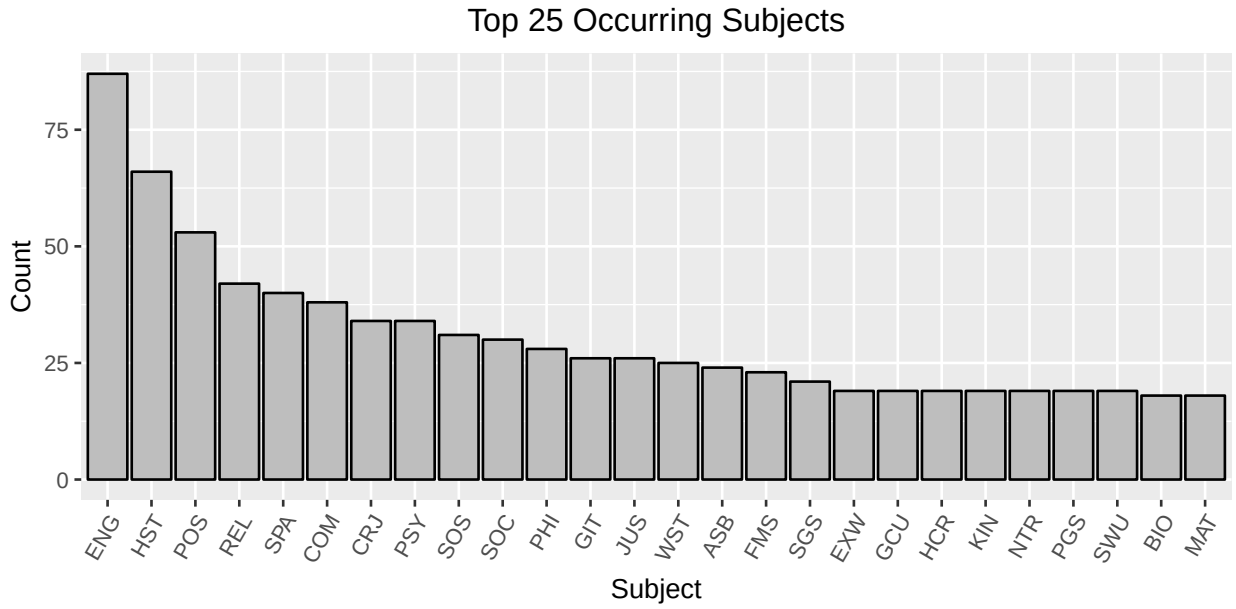
Question 1: Course Equivalencies

The first objective will simply be to investigate the number and types of courses that exist in both modalities. Before quantifying anything, we eliminate observations occurring before the Summer 2011 term as the `iCourse` modality was not fully in place before that time. Any observations occurring in the Winter 2014 term are also dropped because only Face-to-Face courses were offered in that term. Lastly, we drop anybody that is not an undergraduate student.



Takeaways:

- More upper division than lower division
- Lower division courses appear to be mostly 101, 102, 201, 202 courses
- Most commonly occurring upper division catalog numbers seems to be numbers that are usually reserved for topic courses (394, 398, 494, 498)



Takeaways:

- English, History, and Political Science have a significant number of unique courses that are offered

---

Question 2: Unique Course Counts

Here we want some basic stats on the amount of courses each modality offers and the amount of cross-over there is between each modality (how many courses exist in both, how many are unique to a given modality). Obviously we should expect there to be a relatively high number of courses that are only in F2F - what will be interesting is the amount of iCourses that do not have a F2F equivalent.

```
## Number of unique F2F courses: 6122
```

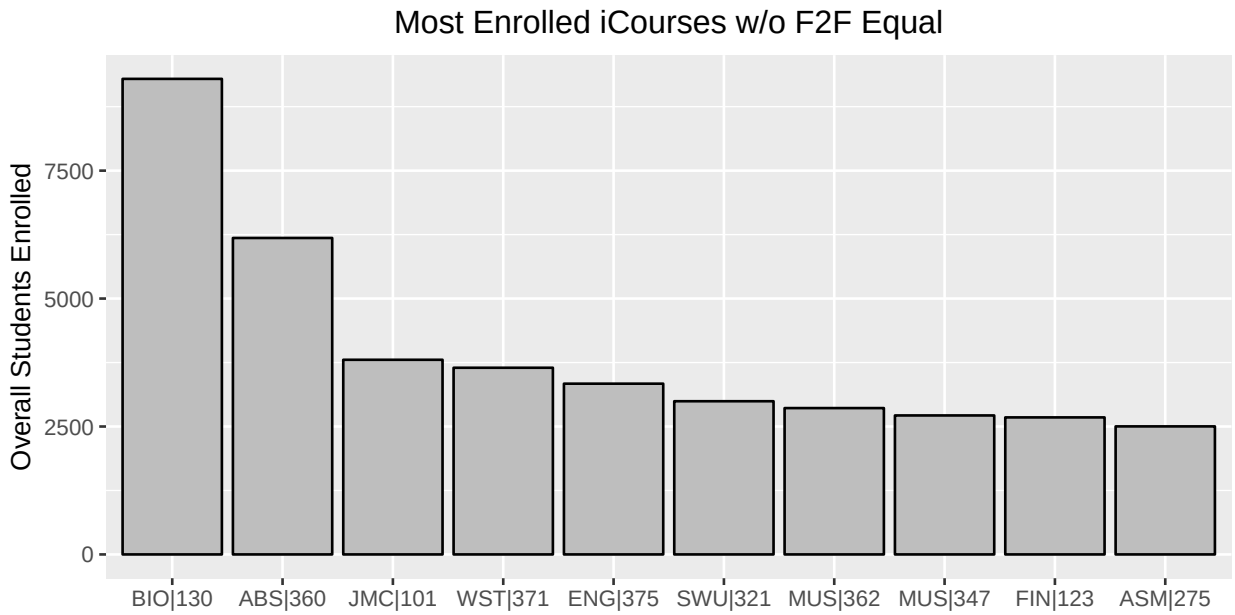
```
## Number of unique iCourses: 1934
```

```
## Number of courses that exist as ICO and F2F: 1532
```

```
## Proportion of F2F courses with ICO equivalents: 0.2502
```

```
## Proportion of iCourses with an F2F equivalent: 0.7921
```

Approximately a quarter of F2F courses are also offerred as iCourses, and about one-fifth of iCourses do not have a F2F equivalent. This is higher than what I was expecting, so let's quickly investigate how populated these courses are:



This isn't terribly exciting, but it is interesting to see that these courses are mostly from different departments, save for MUS|362 and MUS|347. For now we won't do anything with this info but it is good to have available.

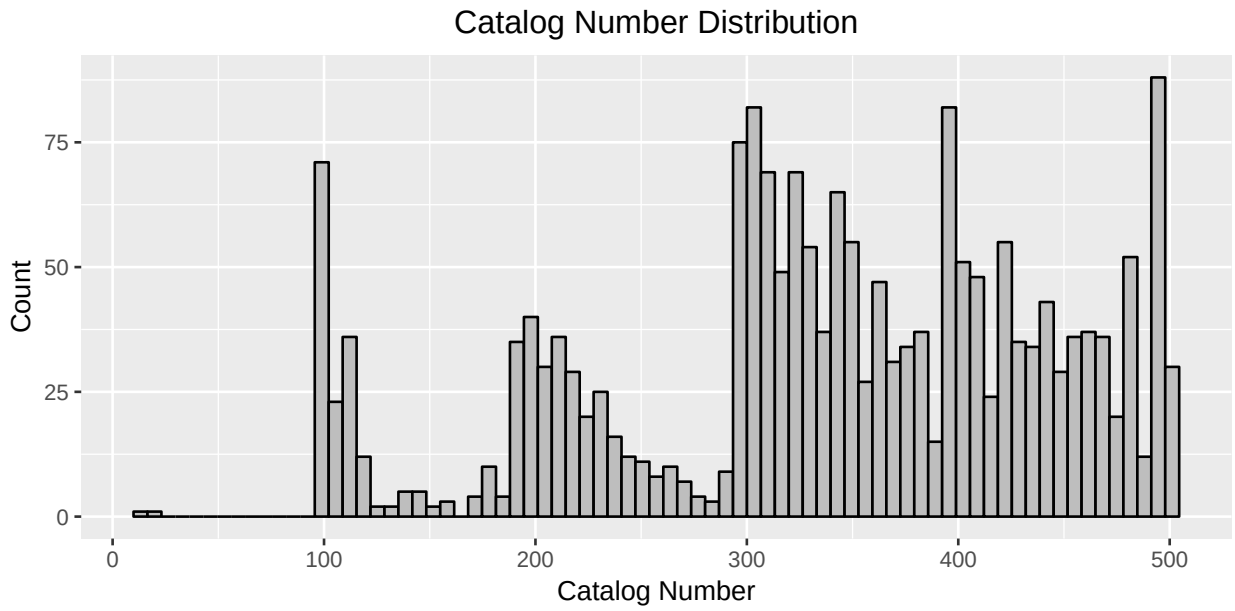
Getting back to the question at hand, let's now investigate the proportion of F2F-ICO equivalents by college. We aren't necessarily looking for any particular bit of information from this but at least we can get a sense of who is making the most use of iCourses.

Important note: The original data we have been working with so far does not contain the correct course ownership information, so we need to bring another data source. From this data we'll use the `acad_org` (College) and `acad_group` (School) variables to do the disaggregation

Second Important Note: At the time of writing this there is no clear way to assign ownership of courses to colleges (apparently this is much harder to suss out than I had expected). Since this is not a huge interest and it can still be approached in alternative ways (via `subject`), we'll leave this alone for now.

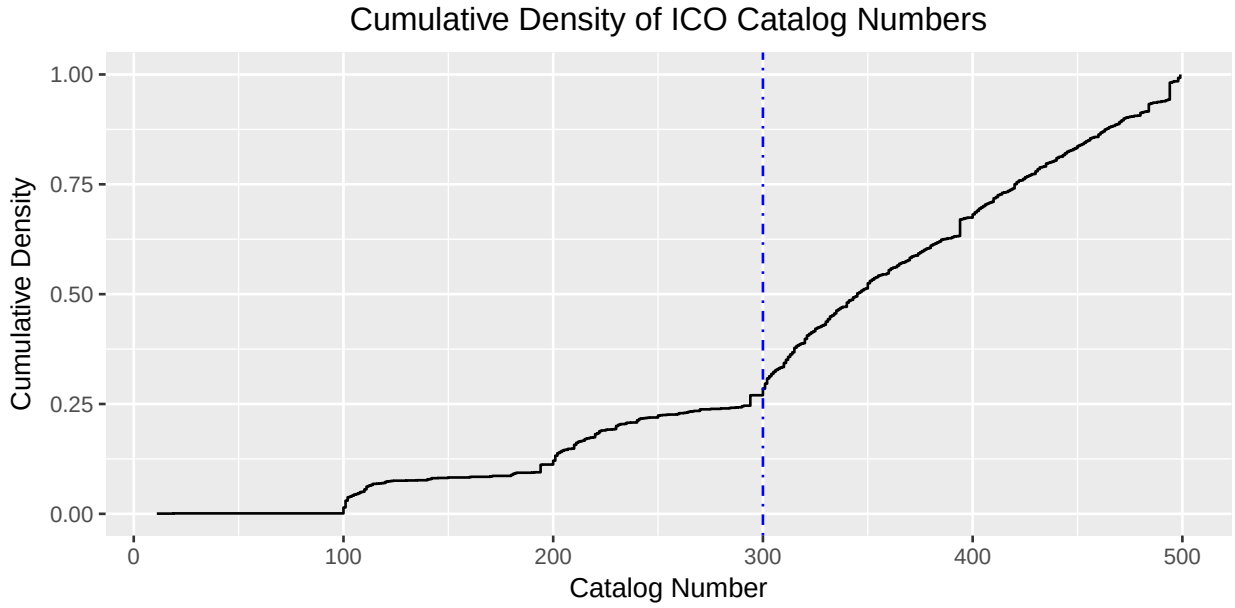
Question 3: Upper Division vs. Lower Division iCourses

This will be similar to the first question we answered, except now we will be looking at the ~1900 iCourses and no F2F courses. First up is a quick histogram of the catalog numbers for the 1900 iCourses in our sample.



We're not necessarily interested in specific numbers but rather just a vague sense of where most of the data lies. Clearly, most of the available iCourses are upper division courses.

Now onto the empirical CDF:

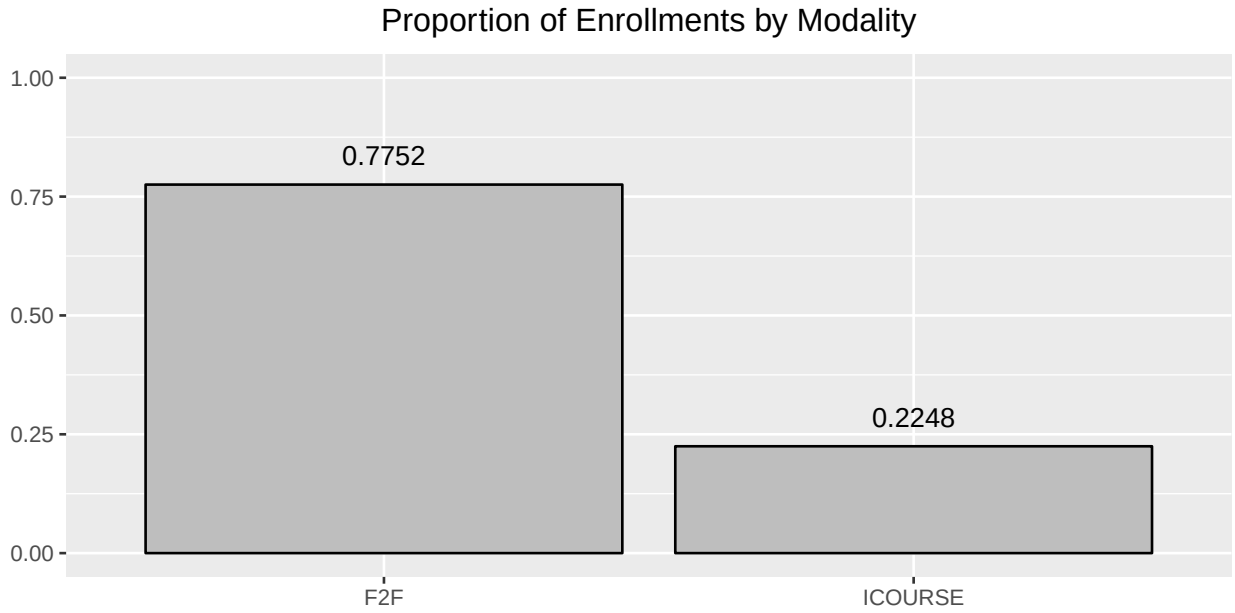


The ECDF displays this idea a bit more succinctly - the dashed line separates our sample into lower-division (to the left) and upper-division (right). Only a quarter of all iCourses are lower-division, so we have a decently left-skewed distribution.

A small note on this - there appears to be two courses with catalog numbers less than 100. These are likely two week-long "Welcome to ASU" sort of courses or some other form of irregular course type. We'll get rid of them now in case they become a problem later.

Question 4: Proportion of Courses taken as iCourses

This will be a very general estimate of how often on-ground students are taking iCourses. Once we have that down, we can disaggregate by college (`acad_prog`) to see which colleges' students are taking the most iCourses. If this doesn't reveal much, we can easily take it a step further and split this up by degree program. However, since the number of degree programs is so great it will probably be best to just look at the extremes.



Not terribly surprising - three-quarters of the enrollments are taken in-person. In fact, its a bit lower than I was expecting. Hopefully the disaggregation by college will give us more information.

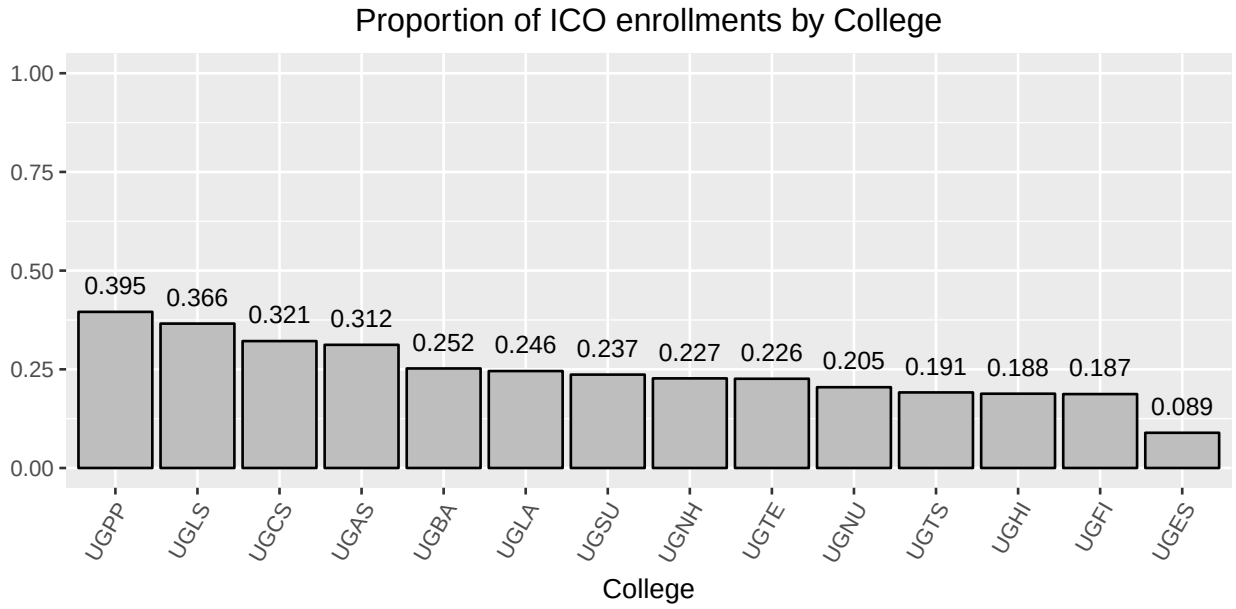
Just to be clear - we will be using the college associated with the students' degree to make these calculations, not the college to which a course belongs. In other words, a correct interpretation of this disaggregation might be:

"12% of all enrollments for *students* in the College of Public Service & Community Solutions were iCourses"

An incorrect interpretation would be:

"12% of all enrollments for *courses* in the College of Public Service & Community Solutions were iCourses"

Finally, when we do the disaggregation we eliminate college codes that don't correspond to a single college. This includes codes for non-degree students, students from the Provost's Office, and students at Thunderbird School of Global Management.



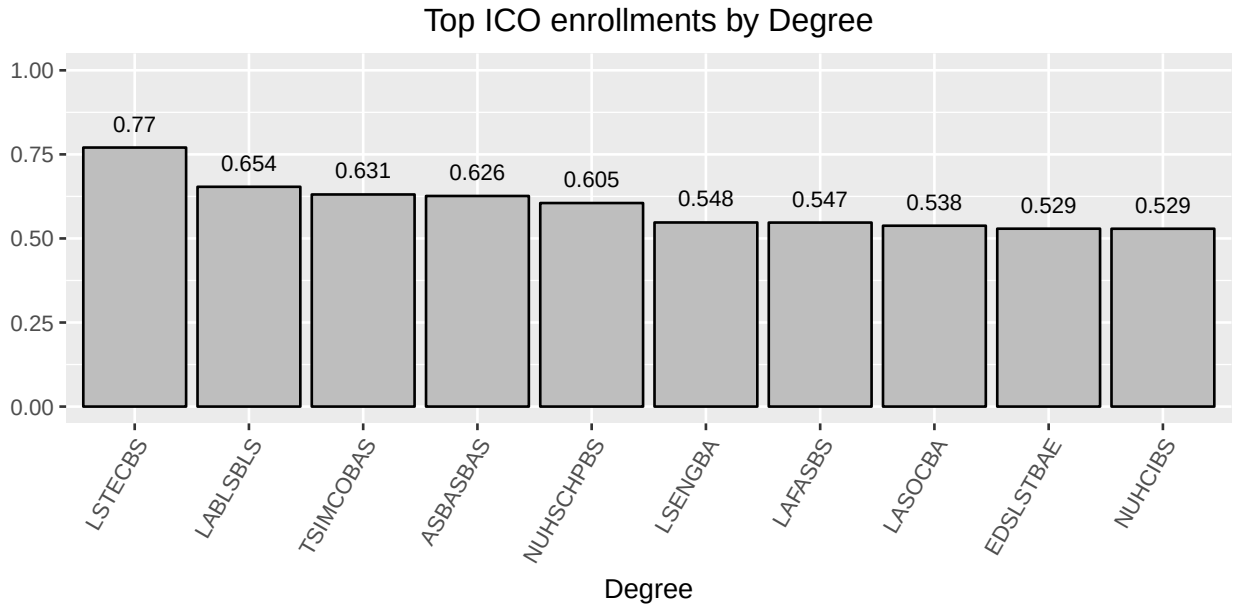
Top 3:

- College of Public Service & Community Solutions (UGPP), 39.5%
- College of Integrative Science & Arts (UGLS), 36.6%
- Cronkite School of Journalism & Mass Communication (UGCS), 32.1%

Bottom 3:

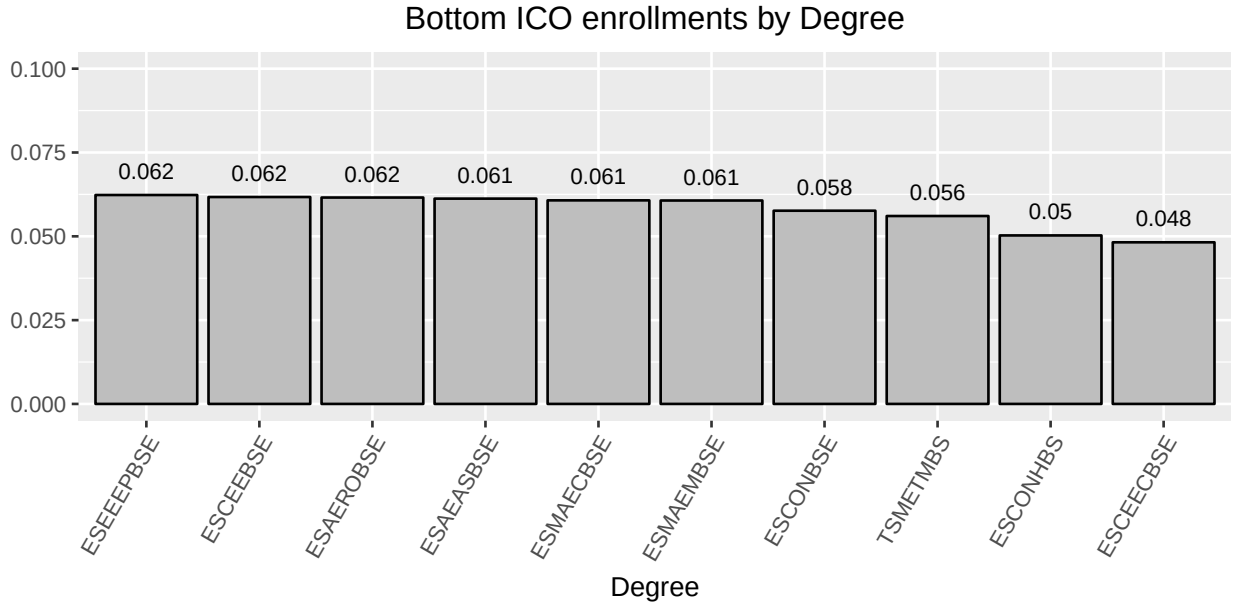
- Ira A. Fulton School of Engineering (UGES), 8.9%
- School for the Future of Innovation in Society (UGFI), 18.7%
- Herberger Institute (UGHI), 18.8%

We now go one step further and look at the actual degree programs that have the highest and lowest proportions of iCourse enrollments



Top 5 by degree:

- Technical Communications BS (LSTECBS), 77%
- Liberal Studies BA (LABLSBLS), 65.4%
- Applied Science/Operations Management (TSIMCOBAS), 63.1%
- Applied Science BAS (ASBASBAS), 62.3%
- Health Sciences/Health Policy (NUHSCHPBS), 60.5%



Bottom 5:

- Civil Engineering (ESCEECBSE), 4.8%

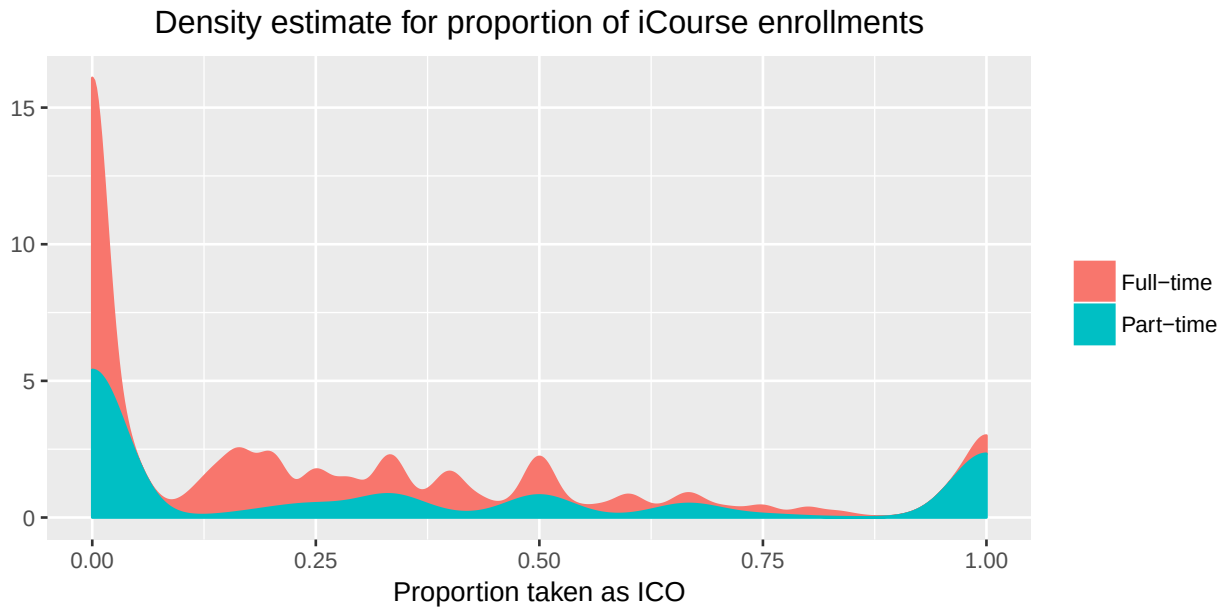
- Construction Management/Heavy Construction (ESCONHBS), 5%
- Manufacturing Engineering Technology (TSMETMBS), 5.6%
- Construction Engineering (ESCONBSE), 5.8%
- Mechanical Engineering (ESMAEMBSE), 6.1%

Question 5: iCourse enrollment and Course Load

The final bit we will investigate is how iCourse enrollment proportions vary by course load. In particular, we want to see if there is any visual difference between the number of full-time students taking iCourses versus part-time students.

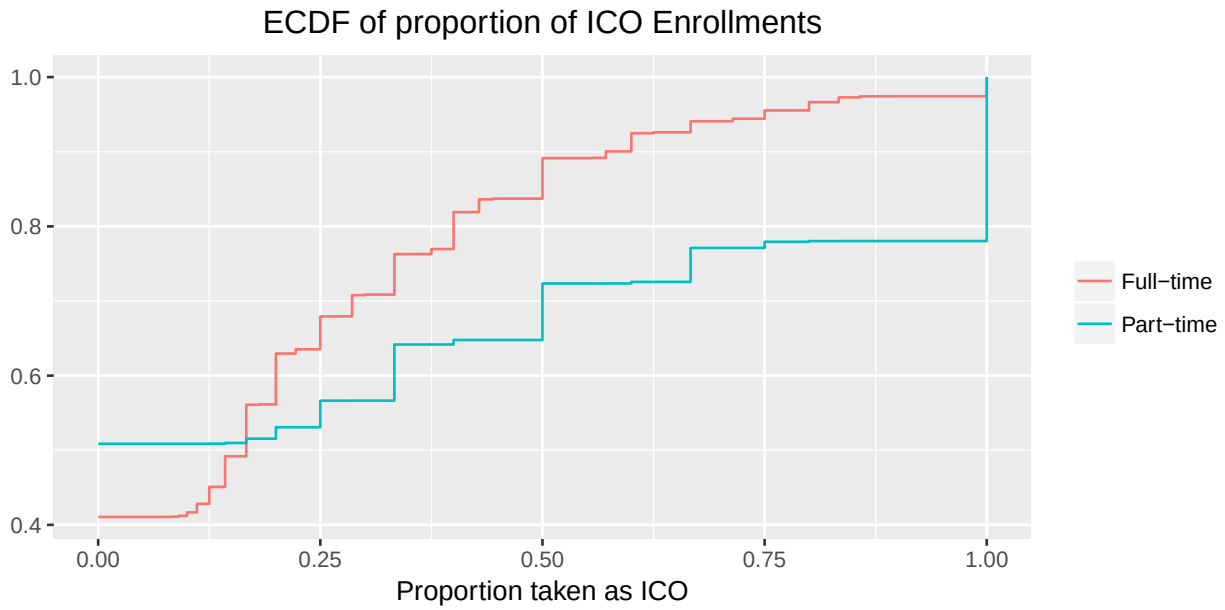
To be extremely clear, we calculate two data points for each student in a given term: the amount of units they are enrolled in, and the proportion of courses that are iCourses. Additionally, we drop the summer terms from this plot because of the reduced amount of credits that are taken during this term.

Also, we are aware that there are a variety of definitions for determining part-time status (some a bit less obscure than others) but for the sake of simplicity we use a very basic rule for this determination. If you are taking less than 12 credits in a particular term, you are considered to be a part-time student.



Based on the density estimate plot alone, we can see that full-time students are taking fewer iCourses as a proportion of their term schedule. Part-time students on the other hand are interestingly clumping up on both extremes, with small peaks around 0.33, 0.5, and 0.66. These small peaks are likely an artefact of courses being worth on average 3 credits and the fact that we are considering part-time students to be taking less than 12 credits.

It is not exactly clear how different these distributions are or if we can even say that they are different. Before getting into any testing, let's look at the empirical CDF to understand them a little more:



Alright! Now we have something interesting to work with - both of these distributions definitely appear to be different. Let's follow up with a couple tests to determine if we have enough evidence to say that part-time students take more or less iCourses (as a proportion of their entire term schedule) than full time students.

We'll first use the KS test to determine if these two distributions are significantly different. More specifically, we want to test if the distribution of the proportion of iCourses taken in a term (`prop_ico`) is independent of full-time status.

```
## Warning in ks.test(ft, pt): p-value will be approximate in the presence of
## ties
```

```
##
##  Two-sample Kolmogorov-Smirnov test
##
## data:  ft and pt
## D = 0.20042, p-value < 2.2e-16
## alternative hypothesis: two-sided
```

Great - we have more than enough evidence to say that `prop_ico` is NOT independent of full time status. Let's get slightly more specific - now that we know that the distribution of `prop_ico` varies with full-time status, let's test to see if there is any significant difference in the conditional sample means.

```
## Test Statistic for difference in proportions: -80.65933
```

```
## Part-time students take 12.3 % more iCourses as a proportion of their schedule each term
```

Question 6: iCourses fulfilling degree requirements

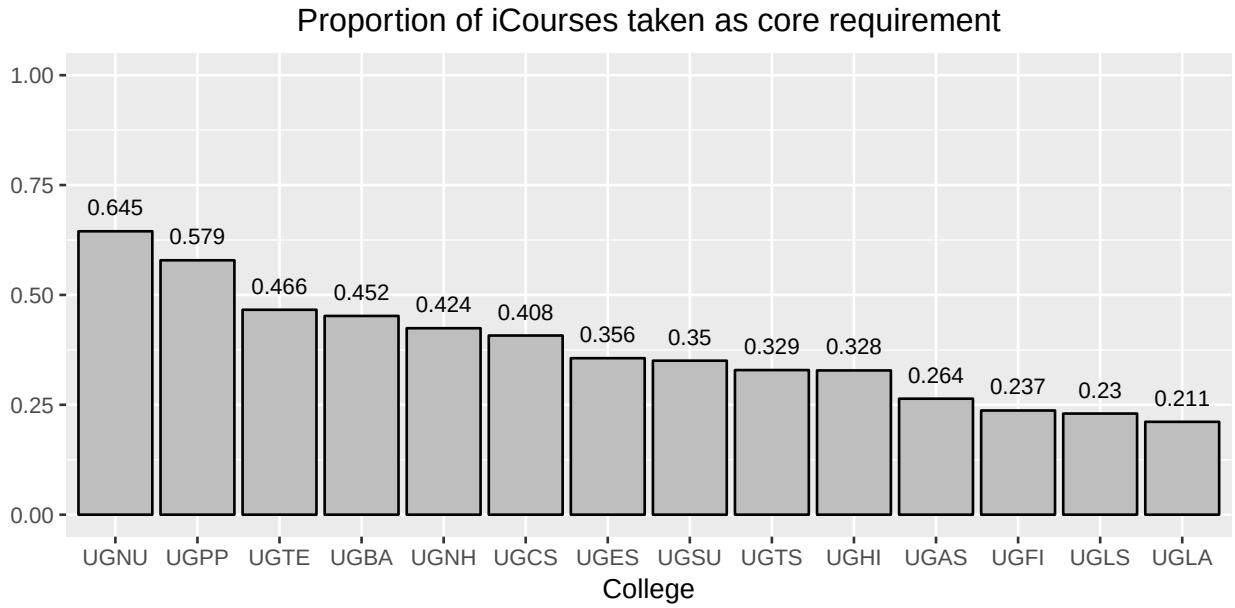
Our final question and perhaps the most interesting one. When it comes to degree requirements, do students view iCourses as an acceptable substitute for in-person courses? Or are students mostly taking iCourses to fulfill more general elective credits? The answer will be important when we start modeling a student's likelihood of taking an iCourse. First, let's calculate the percentage of iCourses that fulfill a core degree requirement in our sample:

```
## 34.5 % of all iCourses taken fulfill a core degree requirement
```

This is a start, but it will be a little more revealing to see how this varies across colleges. Here's the question we want to get at -

Are there colleges whose students take a majority of their iCourses as required courses?

As with the previous question, we need to be clear on exactly how this is to be interpreted. The college does **not** represent the college to which the iCourse belongs, but rather the college hosting the students' degree program. With that in mind, we can say for example "of all the iCourses taken by students in a degree program offered through the College of Liberal Arts and Sciences, 40% were taken to complete a core degree requirement". Now on to the plot:



Top 3:

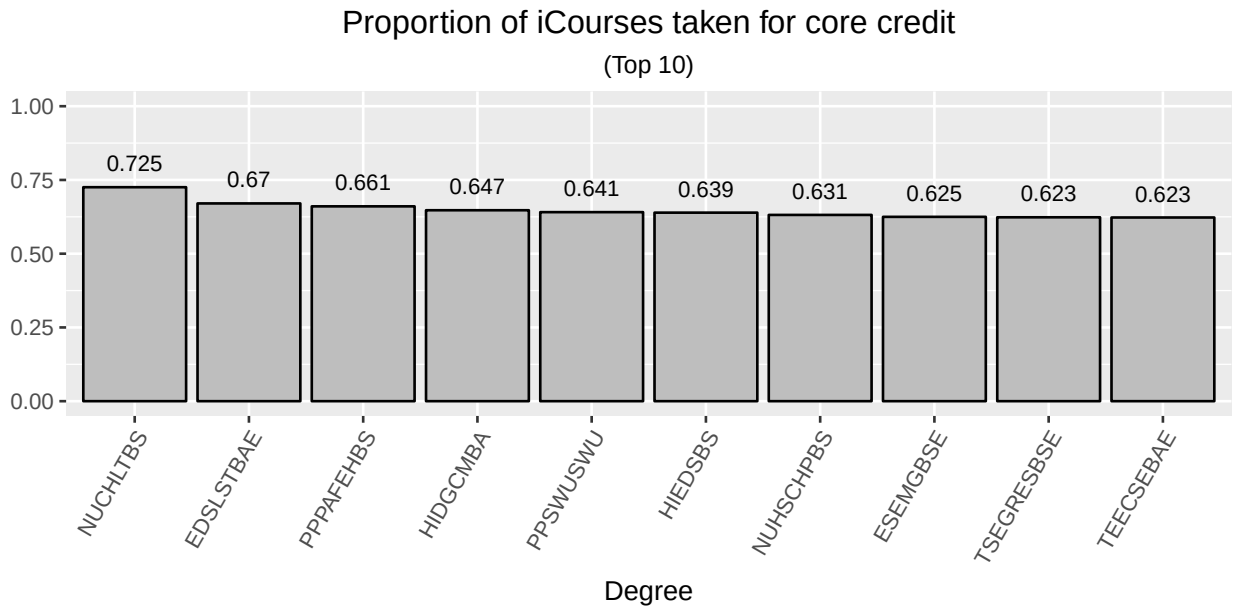
- College of Nursing and Health Innovation (UGNU), 64.5%
- College of Public Service & Community Solutions (UGPP), 57.9%
- Mary Lou Fulton Teachers College (UGTE), 46.6%

Bottom 3:

- College of Liberal Arts and Sciences (UGLA), 21.1%
- College of Integrative Arts and Sciences (UGLS), 23%
- School for the Future of Innovation in Society (UGFI), 23.7%

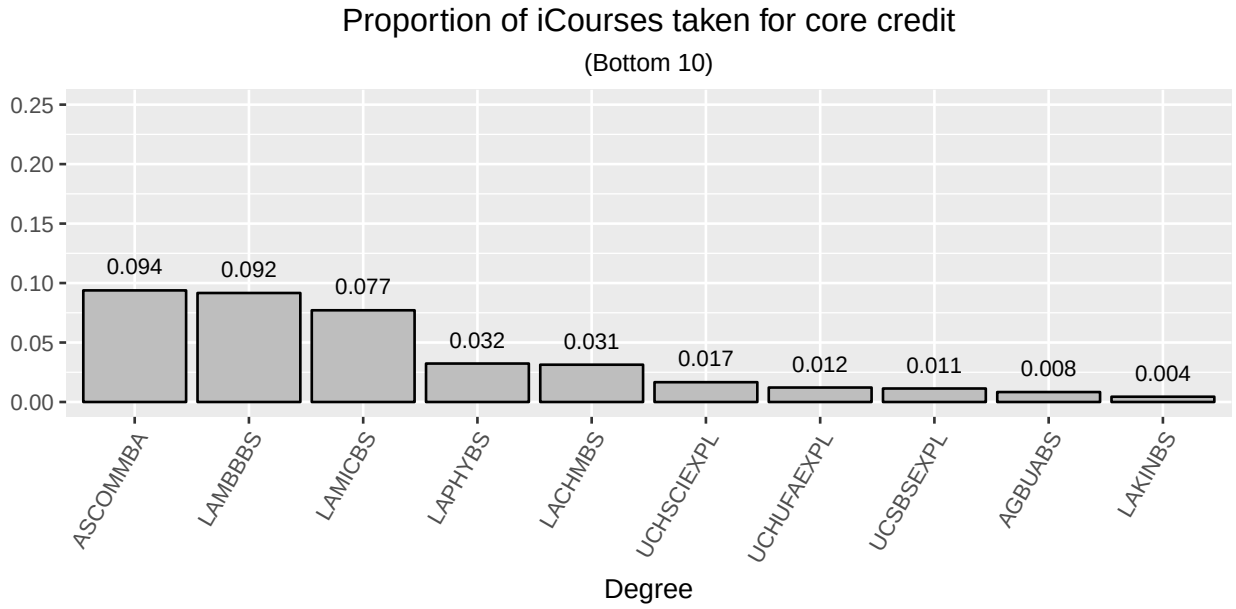
There something very interesting that should be pointed out. First off, UGLS has the second highest amount of ICO enrollments but the second lowest proportion of iCourses being taken as a core requirement. This would seem to suggest that students in these colleges are taking iCourses more often to fulfill elective/general studies requirements. On the other hand, UGES (Engineering) has the lowest amount of ICO enrollments but a little more than a third of those enrollments are for core courses - suggesting the opposite.

Now, let's see how this changes by degree program just as we did earlier. We want to find out which degrees most/least often take iCourses for core credits. Note - we get rid of any degree programs with less than 1000 total course enrollments over the entire time period.



Top 5:

- Community Health (NUCHLTBS), 72.5%
- Educational Studies (EDSLSTBAE), 67%
- Public Service and Public Policy - Emergency Management and Homeland Security (PPPAFEHBS), 66.1%
- Digital Culture - Music (HIDGCMBA), 64.7%
- Social Work (PPSWUSWU), 64.1%

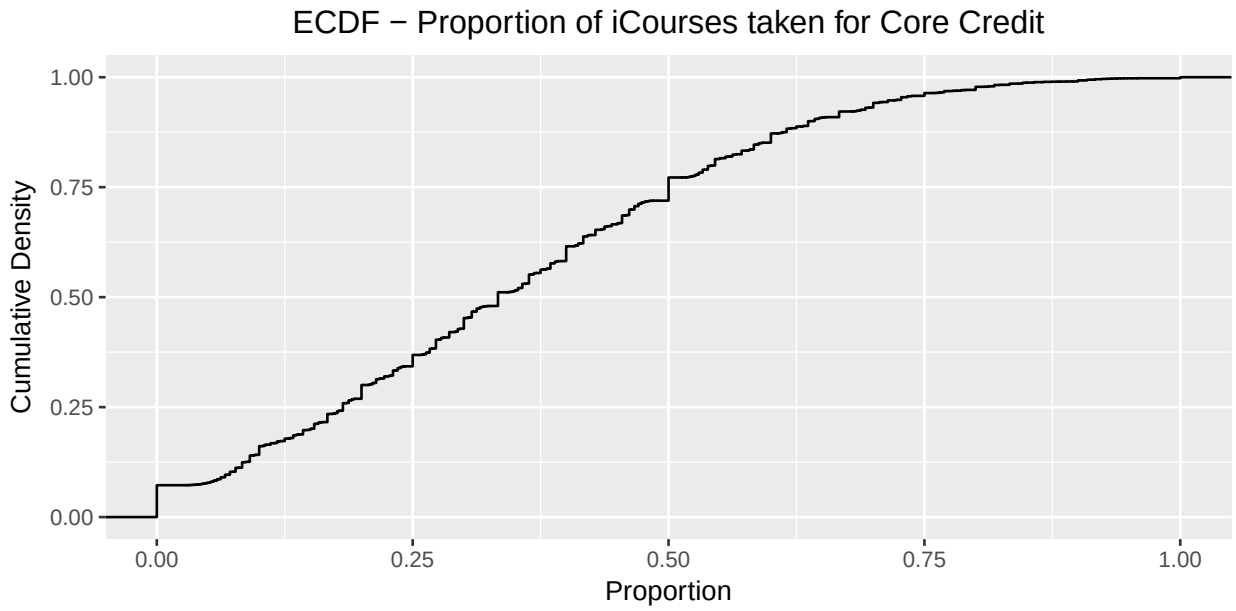


Bottom 5:

- Kinesiology (LAKINBS), 0.4%
- Business Administration (AGBUABS), 0.8%
- Exploratory Social and Behavioral Sciences (UCSBSEXPL), 1.1%
- Exploratory Humanities, Fine Arts, and Design (UCHUFAEXPLR), 1.2%
- Exploratory Health and Life Sciences (UCHSCIEXPL), 1.7%

There seems to be a pretty large difference between the top 10 and the bottom 10, implying that students take iCourses for very different reasons across degree programs.

Finally, one last thing to check out - the distribution of the proportion of iCourses taken for core credit. For each student, we will grab all of the iCourses that they have ever taken and find the fraction of those courses that were required courses. Once that's done for every tsudent, we can take a look at the distribution. One note, though - there will likely be a lot of students who take very few iCourses. The proportion that we calculate for these students may not be the best representation of a students iCourse-taking habits. So for the sake of avoiding this problem, we make the arbitrary decision to cut out students who have taken less than 10 iCourses.



The ECDF gives us a basic idea of this distribution - sort of uniform and right-skewed. Something to note from this graph is that 75% of students who have taken 10 or more iCourses took less than half of them as required courses. In other words, a majority of the students taking iCourses are taking them as electives.