Analysis of iSchool Programs and Courses

by- Asmita Khode, Divyaang Agarwal

Introduction

Often students while looking up for their dream universities and programs don't get access to proper statistics which could help them compare the different options available and make informed decisions. Also, schools and colleges lack information through which they can evaluate their performance over the years and improve the quality in future. Hence, as students at the School of Information Sciences, we tried to study the evolution of the information sciences programs and courses over the years.

We have created several interactive dashboard using Tableau which can benefit new aspirants to have a quick look at the statistics about the iSchool and support their decision-making process. Also, this dashboard can be used by the iSchool officials to evaluate the performance and popularity of the programs and come up with strategies to improve their quality.

Dataset

We have collected data from two major sources - University of Illinois, Division Of Management Information's website and Eddie Web Intelligence Tool. From DMI we collected data regarding the total enrollment by majors, race, gender and residency. Using Eddie we gathered data about the courses like the faculty name, number of students enrolled and number of seat available, etc.

The data on Division Of Management Information's website is available on a semester wise basis. We had to download the data for each semester from Fall 2017 to Spring 2022 and then merge them into a single file for our analysis. The second challenge with this DMI data was that, all the semester's data did not contain the same number of headers so we had to manually consider this and merge all the semester files together.

The Course Schedule universe which we used in the Eddie Web Intelligence tool had a large number of columns so figuring out what columns would be useful for our analysis was bit of a challenge. Next, we had to add query filters to restrict the results to our requirements. We have added the following filters:

1. Course Subject code : This particular filter restricts the course subject code to School of Information Sciences.
2. Section Enrollment Current Info Ind: Since Eddie stores information with different timestamps, it is important to consider the most recent information for each semester and to do so we have set Section Enrollment Current Info Ind as "Y".

3. Section Enrollment : Many course sections are created but are ultimately not used due to some reasons, to filter such sections we have added a filter as Section Enrollment >0.
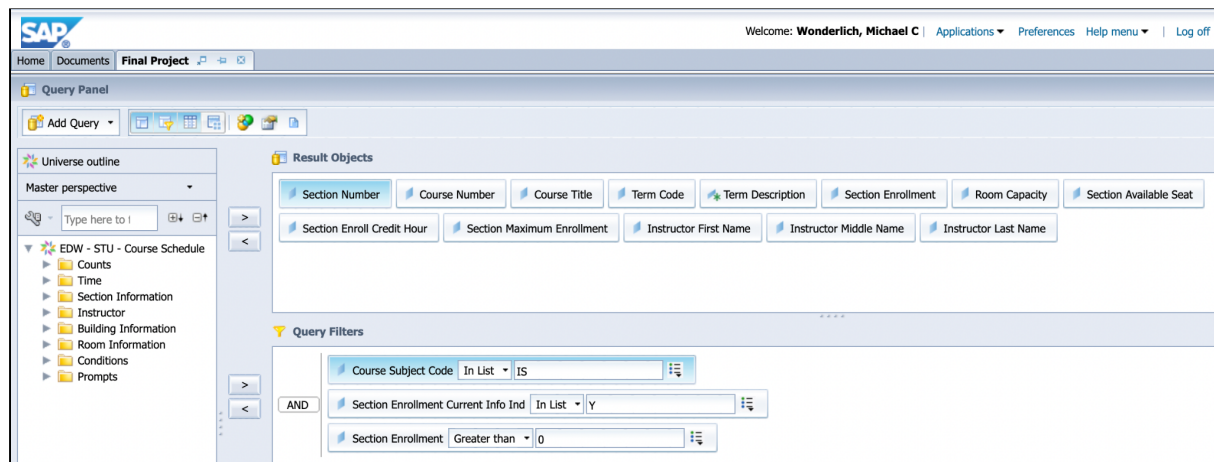

Figure 1: Illustration of the Eddie Web Intelligence tool used to get the courses and instructor data

Data Preparation

The data received from DMI and Eddie were not in ready to use format. So, we used Tableau Prep Builder to prepare our data which we fed to Tableau for creating the desired visualisations. Some of the data preparation steps we took included:

1. In all the data files there is a column 'Term/Term Description' in which data was stored in "Fall 2021" format but we wanted this to be splitted as "Fall" and "2021" separately into two different columns. Thus, we added the "Split terms" step in all the pipelines shown in Figure 2.

2. To get our data based on gender, nationality and residency we created a pipeline as shown in Figure 2 (first one). Firstly, we summed all the columns of races to get the total non-international student enrollment counts. Then we dropped all the columns which were not required. Next, we created three subsets of the input data file, one each for gender, nationality and residency. Lastly, for each category we had to do pivot to get the data in the desired format.

Figure 2: Illustration of the data preparation pipelines developed in Tableau Prep Builder

After executing the Tableau Prep Builder pipelines we got five separate data files as follows:

1. enrollment.csv containing data regarding the total enrollment by major
2. gender.csv containing data regarding the total enrollment by major and gender
3. nationality.csv containing data regarding the total enrollment by major and nationality
4. residency.csv containing data regarding the total enrollment by major and residency
5. eddie.csv containing data regarding the courses, instructors, section enrollment, etc.

For creating visualisations in Tableau, we have joined the first four data file as shown in Figure 3. Eddie data could not be joined with other data files because of granularity issues so we have kept it as a separate data source in Tableau.
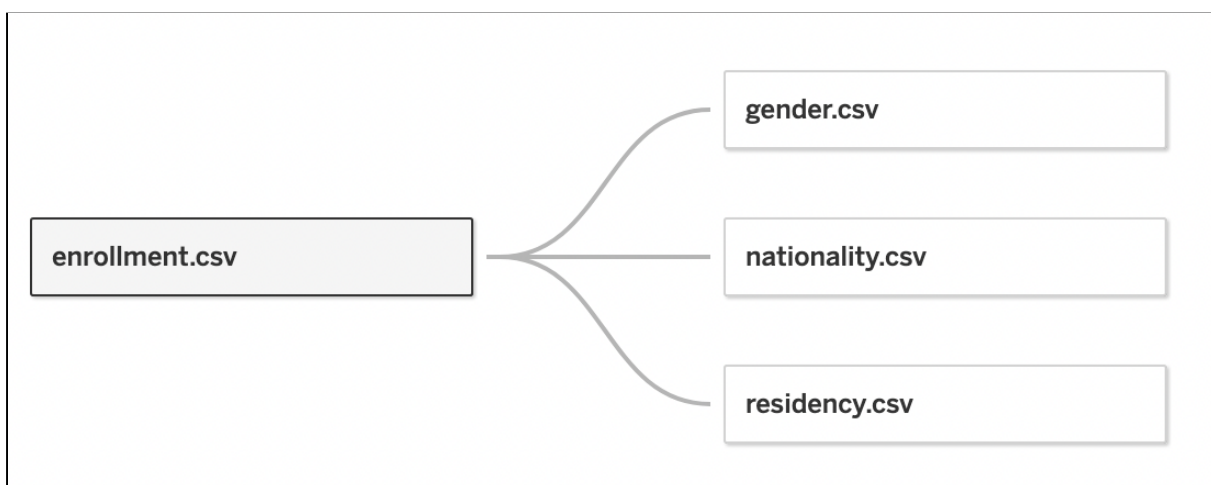

Figure 3: Illustration of Tableau data source

Analysis

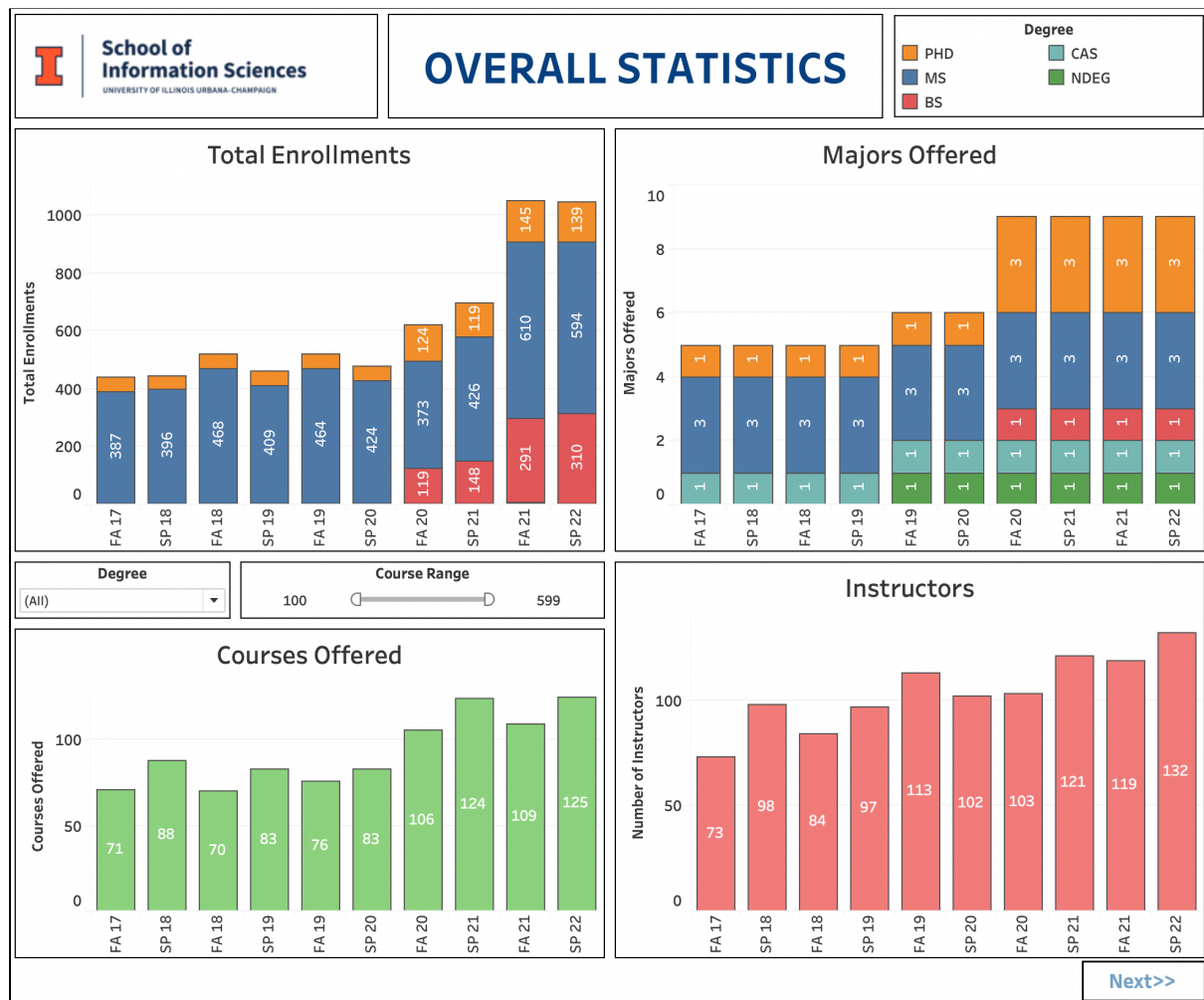**Dashboard 1 - Overall Statistics**

Figure 4: Dashboard 1 - Overall Statistics

Dashboard 1 depicts the overall statistics from Fall 2017 to Spring 2022 based on total enrollments, total courses offered, total majors on level of education and also the total number of instructors. We can filter total enrollments by degree offered like BS, MS, PHD , CAS and NDEG. Also, we can also set a course number range to get a bar chart for the number of courses offered in that particular range.

In Figure 4, we can see that year on year overall enrollments, courses offered and instructors have been increasing. One factor that might account to this increase is the addition of new majors by the School of Information Sciences which might have helped captivate new enrollments and hence increase in instructor's appointments. Also, there is a great increase in number of enrollments in year 2021 as compared to previous year which can be investigated for correlation with pandemic.

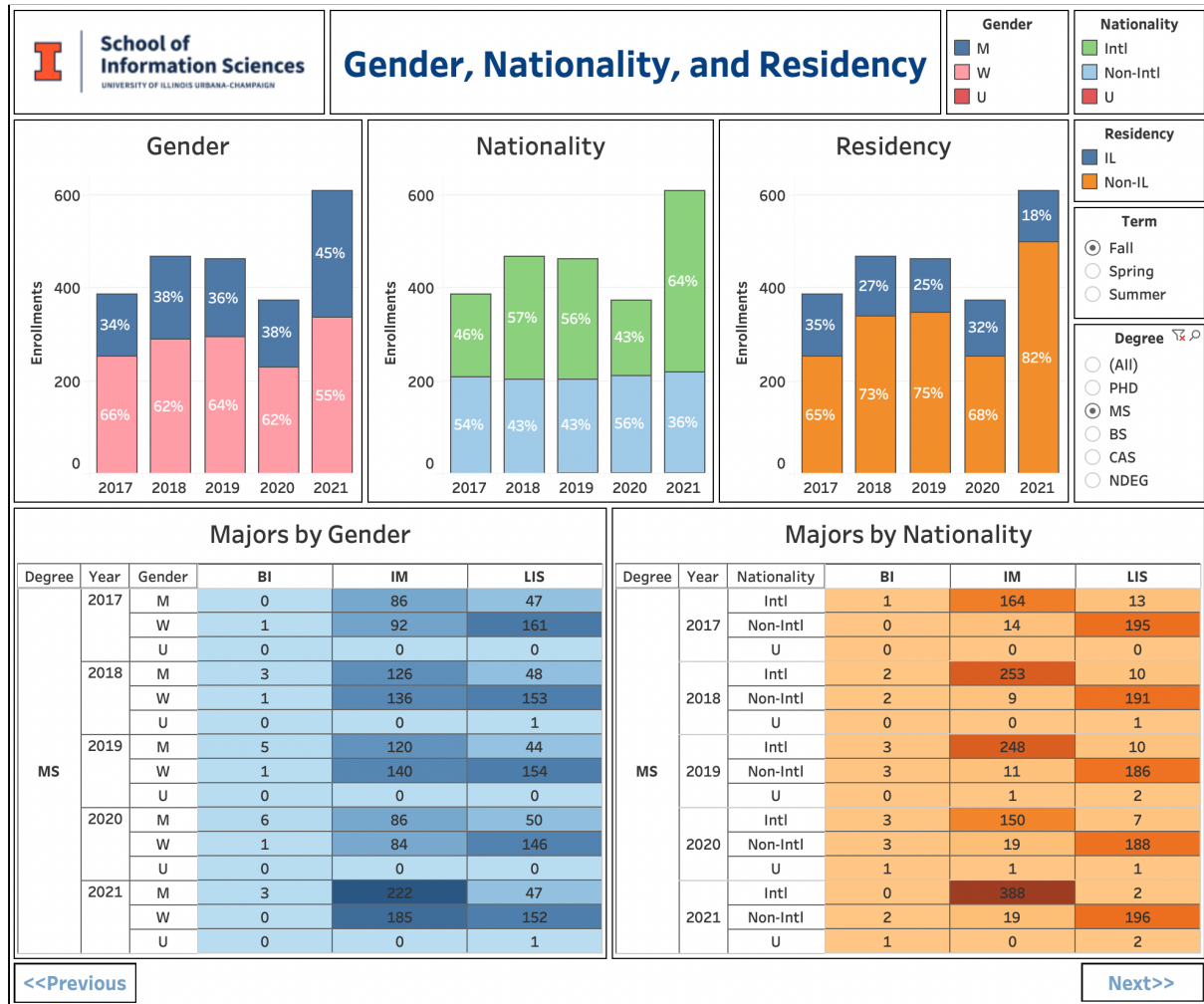**Dashboard 2 - Gender, Nationality, and Residency**

Figure 5: Dashboard 2 - Gender, Nationality, and Residency

Dashboard 2 illustrates the enrollment statistics by gender, nationality and residency on yearly basis for a particular term and degree as selected in the filter. All the five visualisations in the dashboard react to the Term and Degree filter.

For the dashboard shown in Figure 5 we have filtered data by Fall term for the Masters program. For this particular term and program we see that the overall percentage of women enrolled is higher as compared to men throughout the years from 2017 to 2021. Also we see there is somewhat of a balance between national and international students with a minor difference in overall percentage from 2017 to 2020 but we notice a sudden jump in percentage of international students in 2021. Moving to gender ratio by majors, the Information Management program seems to have a comparable number of men and women whereas women dominate in numbers for Library and Information Sciences. Also, majority of students enrolled in Information Management program are international students but on other hand non -international students mostly prefer Library and Information Sciences.

Similarly, we can select different options available in the filters and make inferences for different scenarios.
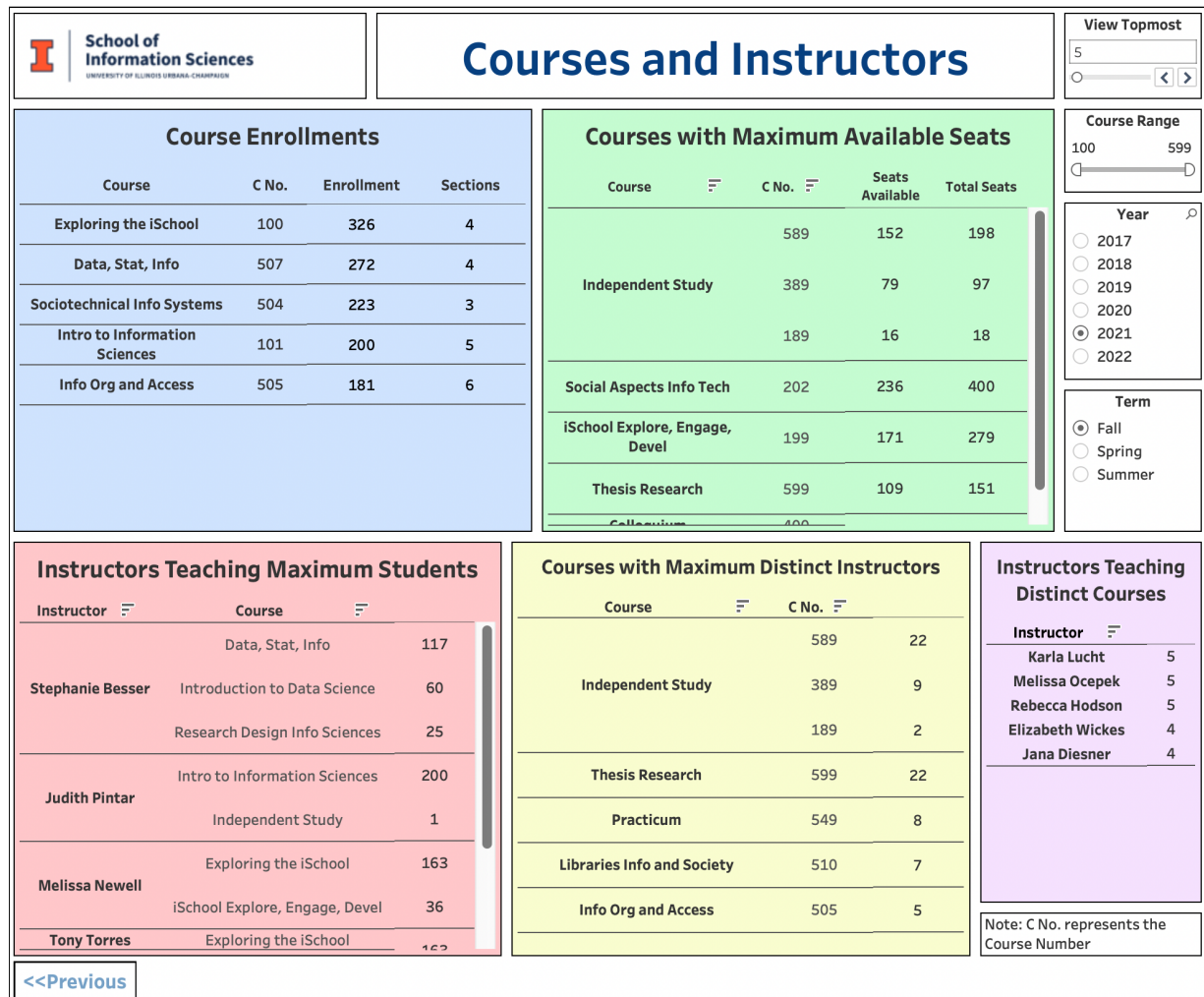
Dashboard 3 - Courses and Instructors

Figure 6: Dashboard 3 - Courses and Instructors

Dashboard 3 showcases a lot of information about courses and instructors like we can get the course enrollments, courses with maximum available seats, instructors who have taught maximum number of students and courses which have been taught by maximum different instructors.

All the visualisations in this dashboard have a parameter control "View Topmost" whose value can be set by the user. The purpose of this parameter control is to restrict the number of rows displayed in the dashboard. For example, in all the tables in Figure 6 only 5 rows are displayed, these five rows are the topmost five outcomes for the respective categories.

This dashboard also contains filters for course range, year and term. Again, all the visualisations present in this dashboard react to these filters. Figure 6 shows results for Fall 2021 with course numbers ranging from 100 to 599.

Keeping term as Fall , course range between 100 and 599 and topmost records to be fetched to five we have found that from 2017 till 2021 Info, Org and Access and Data Stats and Info have always been amongst the popular courses in terms of Course Enrollment. Also since 2019, Karla Lucht is teaching maximum number of distinct courses as compared to other instructors. This is just one of the many inferences we can make from this dashboard. The user can select different terms, year and course range in the filters to get the results for their desired configuration.

Some points to note regarding Dashboard 3:

1. Dashboard 3 has option to display data for Fall 2022. Since the data for Fall 2022 is changing everyday, the information portrayed for this term may not be reliable and is restricted to the date we downloaded the data, i.e. around mid April.

2. Sometimes the course range filter slider does not show the full course number ranges, in such cases we can simply click on the lower or upper course numbers in the fillter and manually the type the desired course number.

Future Work

With this project we would like to dig deeper into analysing admission criteria for ischool programs by studying parameters like number of applicants, gpa, exam scores, work experience of applicants and acceptance rate etc. which would further help aspirants to analyse their application while applying for programs at iSchool and also help iSchool to keep a check on quality of students that are admitted to different programs. We also look forward to seek suggestions and feedback from iSchool authorities and implementing them in our project.

Glossary

We have abbreviated several terms in the dashboards. Here are the full forms for those terms:

1. M - Men
2. W - Women
3. U - Unknown
4. Intl - International
5. Non-Intl - Non-International
6. IL - Illinois
7. Non-IL - Non Illinois
8. BI - Bioinformatics

9. I - Informatics
10. IM - Information Management
11. IS - Information Sciences
12. LIS - Library & Information Sciences

Attachments

The following file have been attached with this report:

1. DWBI-Project.twbx - Tableau Packaged Workbook
2. dwbi-data-prep.tfl - Tableau Flow
3. DMI.xlsx - Data downloaded from DMI website
4. Eddie.csv - Data downloaded from Eddie
5. enrollment.csv - Tableau Prep Builder output file
6. gender.csv - Tableau Prep Builder output file
7. nationality.csv - Tableau Prep Builder output file
8. residency.csv - Tableau Prep Builder output file
9. eddie.csv - Tableau Prep Builder output file

We have emailed the Eddie file to mcwonder on 6 May, 2022, the name of the file is Final Project.