University at Buffalo
The State University of New York

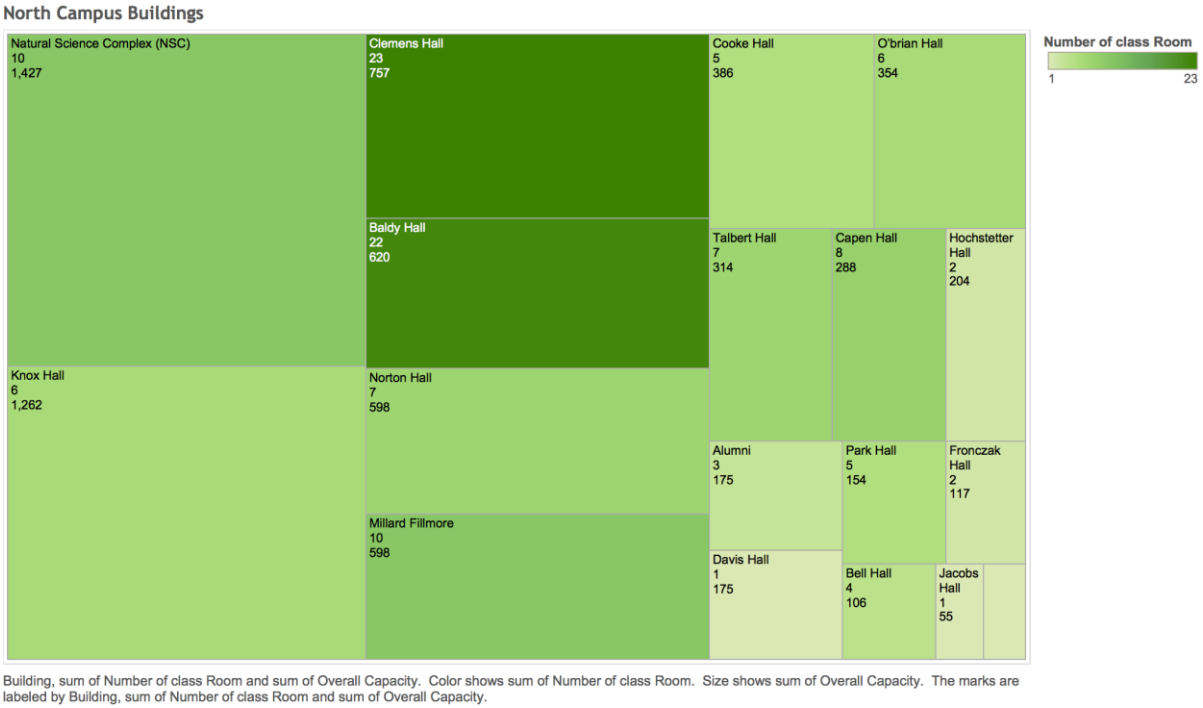
Department of Computer Science and Engineering
School of Engineering and Applied Sciences

CSE 587- Data Intensive Computing
Project II- Problem 3

Provided by:
Alireza Farasat (Person Number: **50060827**)
Rahul Gopalsamy (Person Number: **50163719**)

Introduction:
Classroom scheduling is common problem encountered by Universities and Colleges every semester to efficiently use the available space. In spite of it common occurrence there are still no efficient scheduling algorithms. The enrollment in colleges, popularity of courses offered in a particular semester and several other factors make this problem more complex. It is important to analyze the trend over the years and extrapolate the result to give an efficient scheduling for the following the semester. The historical data can significantly help the planning team to provide more realistic weekly schedule in each semester based on courses' popularity, the number of enrolments, etc.

Classroom Scheduling Problem:
The project description is to get insight into the Classroom scheduling at University at Buffalo North campus over the years. The scheduling of classroom can be viewed as a combinatorial Optimization problem where $x_{ijk} = 1$ represents allocating classroom *i* to the course *j* at the time slot *k.* However, in this report we try to provide some insights without solving a complex optimization problem.
There are a total of 125 buildings in North Campus, University at Buffalo. This include Resident halls, Library, Research space, Stadium, Student Union and many other. Eighteen Buildings have classrooms which are presented below along with the number of rooms, capacity of each room. The following diagram shows the number of class rooms in North Campus buildings.



Building, sum of Number of class Room and sum of Overall Capacity. Color shows sum of Number of class Room. Size shows sum of Overall Capacity. The marks are labeled by Building, sum of Number of class Room and sum of Overall Capacity.

The three parameters of classroom scheduling are:
1) Courses
2) Classrooms
3) Time

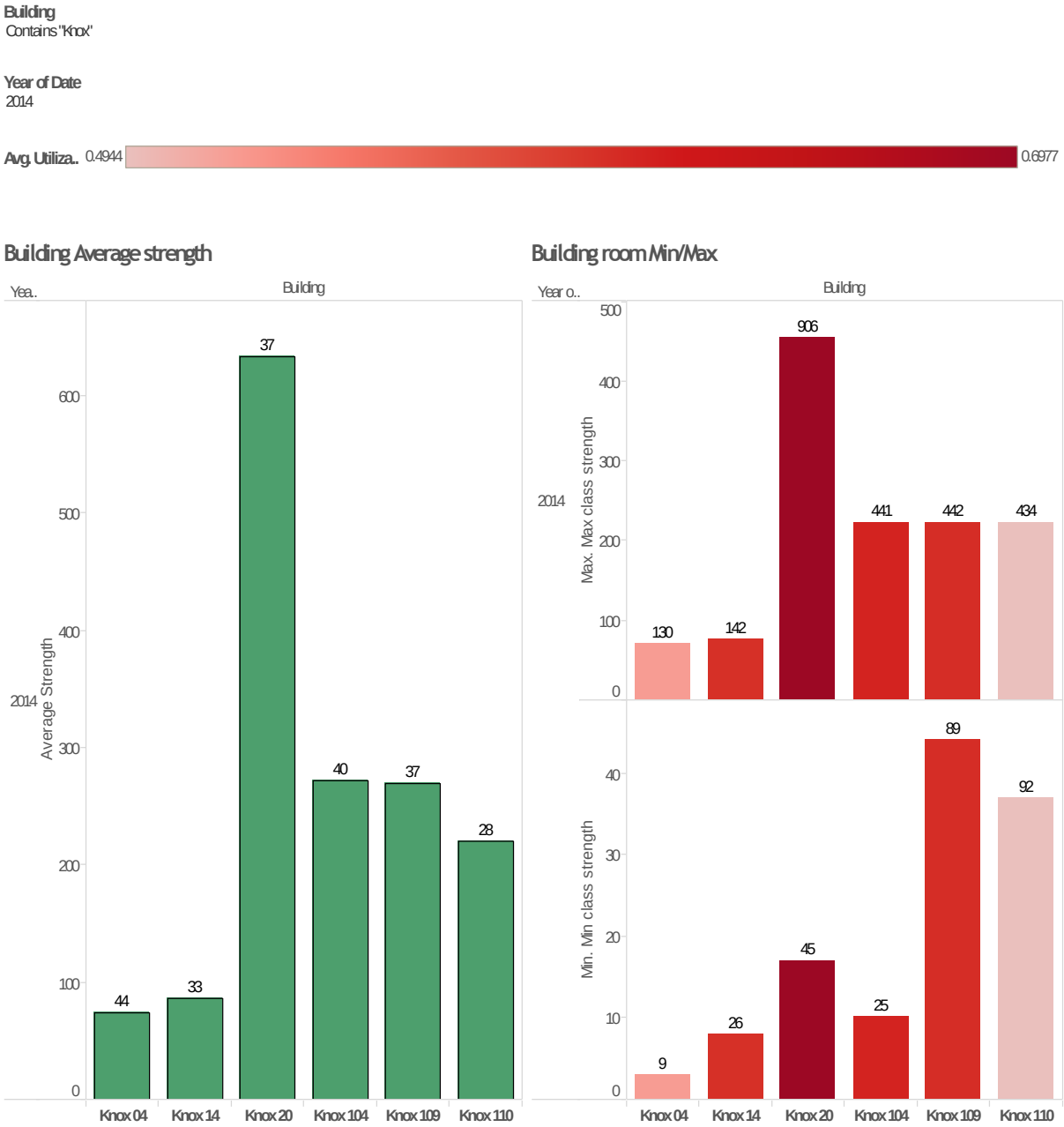
A detailed analysis of above three factors are done over the past years at UB to get a deeper insight.

We begin our analysis with the graphical view of space and number of classrooms in each building in North campus at University at Buffalo.

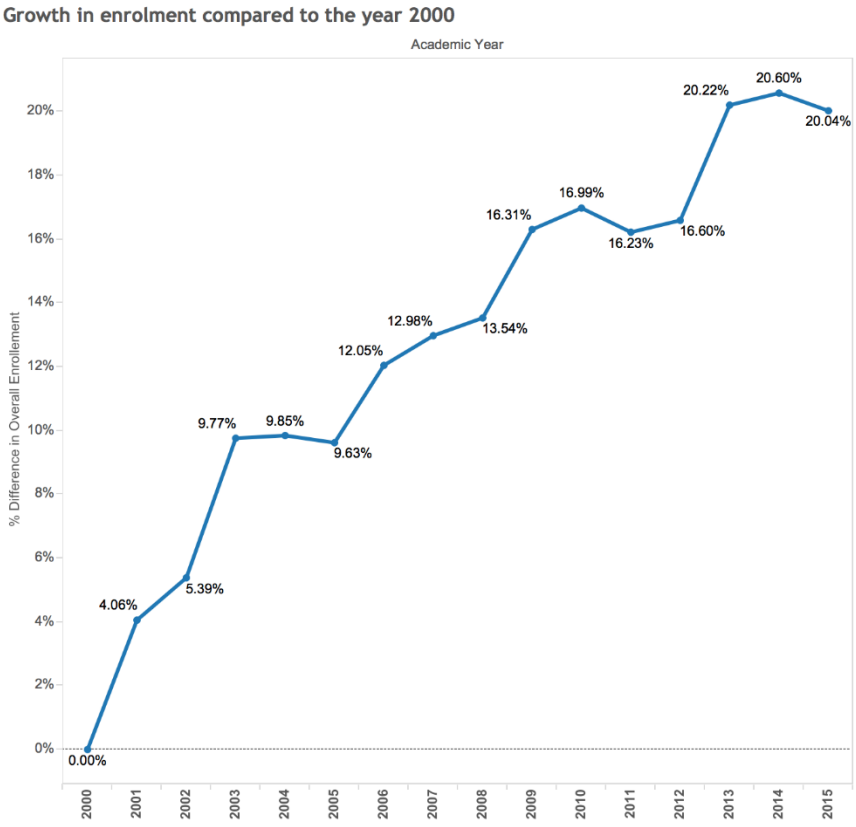
This gives us an overall picture of the classrooms and their capacity in North campus. Next, we would like to know about the enrollment trend in students over the years for University at Buffalo. The overall enrollment for the last 15 years is obtained from the commons dataset.

Building	Number of class Room	Overall Capacity
Alumni	3	175
Baldy Hall	22	620
Bell Hall	4	106
Capen Hall	8	288
Center for the Arts	1	48
Clemens Hall	23	757
Cooke Hall	5	386
Davis Hall	1	175
Fronczak Hall	2	117
Hochstetter Hall	2	204
Jacobs Hall	1	55
Knox Hall	6	1262
Millard Fillmore	10	598
Natural Science Complex (NSC)	10	1427
Norton Hall	7	598
O'brian Hall	6	354
Park Hall	5	154
Talbert Hall	7	314

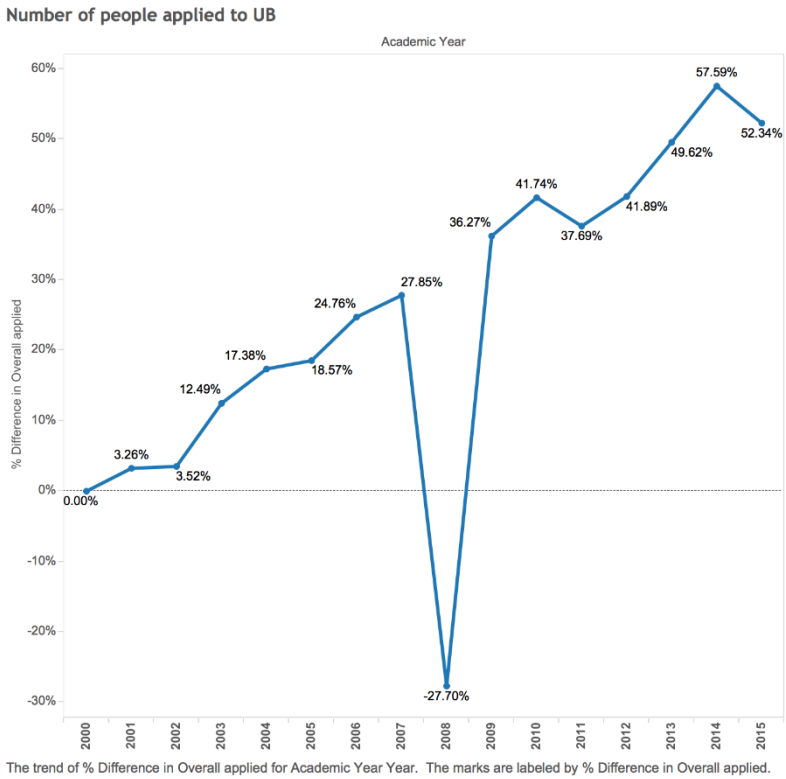
It is worth mentioning that we can compare the minimum and maximum number of students enrolled in each year for each building using dashboard. For example, the following diagram compares the min and max of number of enrolments in 2014 for Knox Hall's rooms. This shows that these rooms are not used efficiently because of the high variance.



On justification for such a high variation can be the increasing trend of enrolments. The percent growth in enrollment rate over the years compared to year 2000 is shown in the graph.



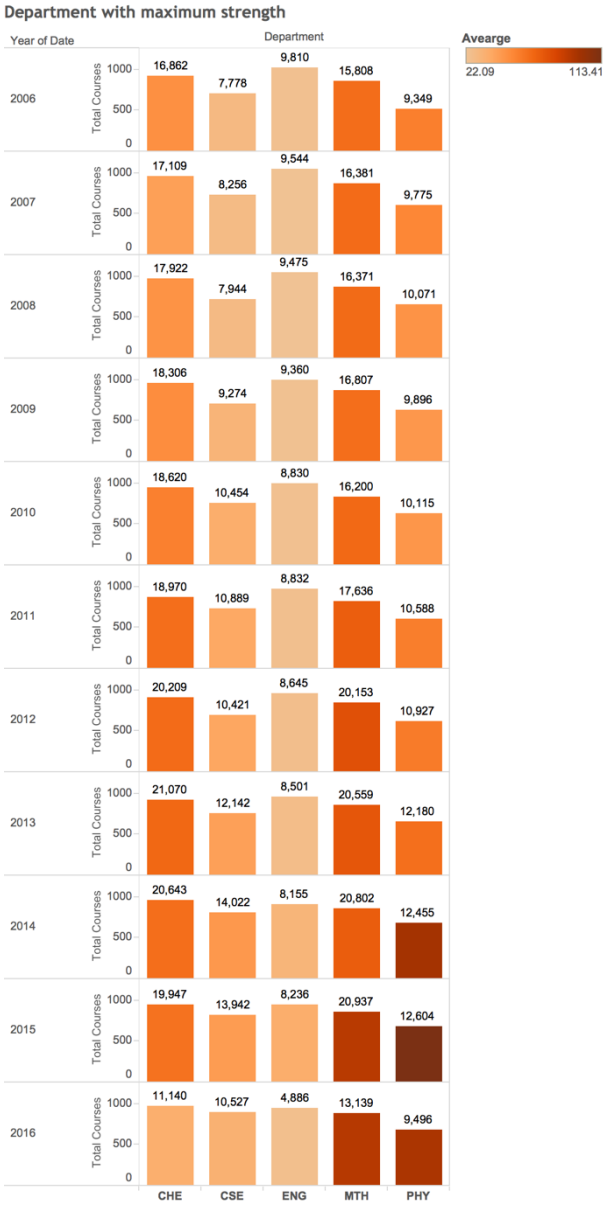
The trend of % Difference in Overall for Academic Year Year. The marks are labeled by % Difference in Overall.



The trend of % Difference in Overall applied for Academic Year Year. The marks are labeled by % Difference in Overall applied.

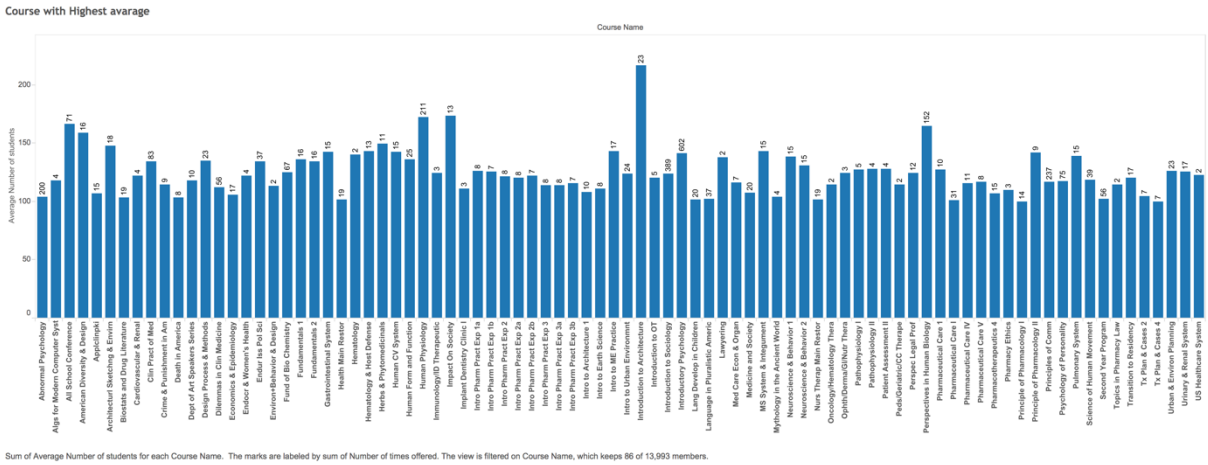
The above two graph shows the increase in number of students over the years in terms of enrollment and people applied.

There are various department in the univeristy that offer courses over the years. We would like to get insight into the departments to identify the ones with highest strength and maximum number of courses. Below is the top 5 department with higher number of courses and maximum strength.



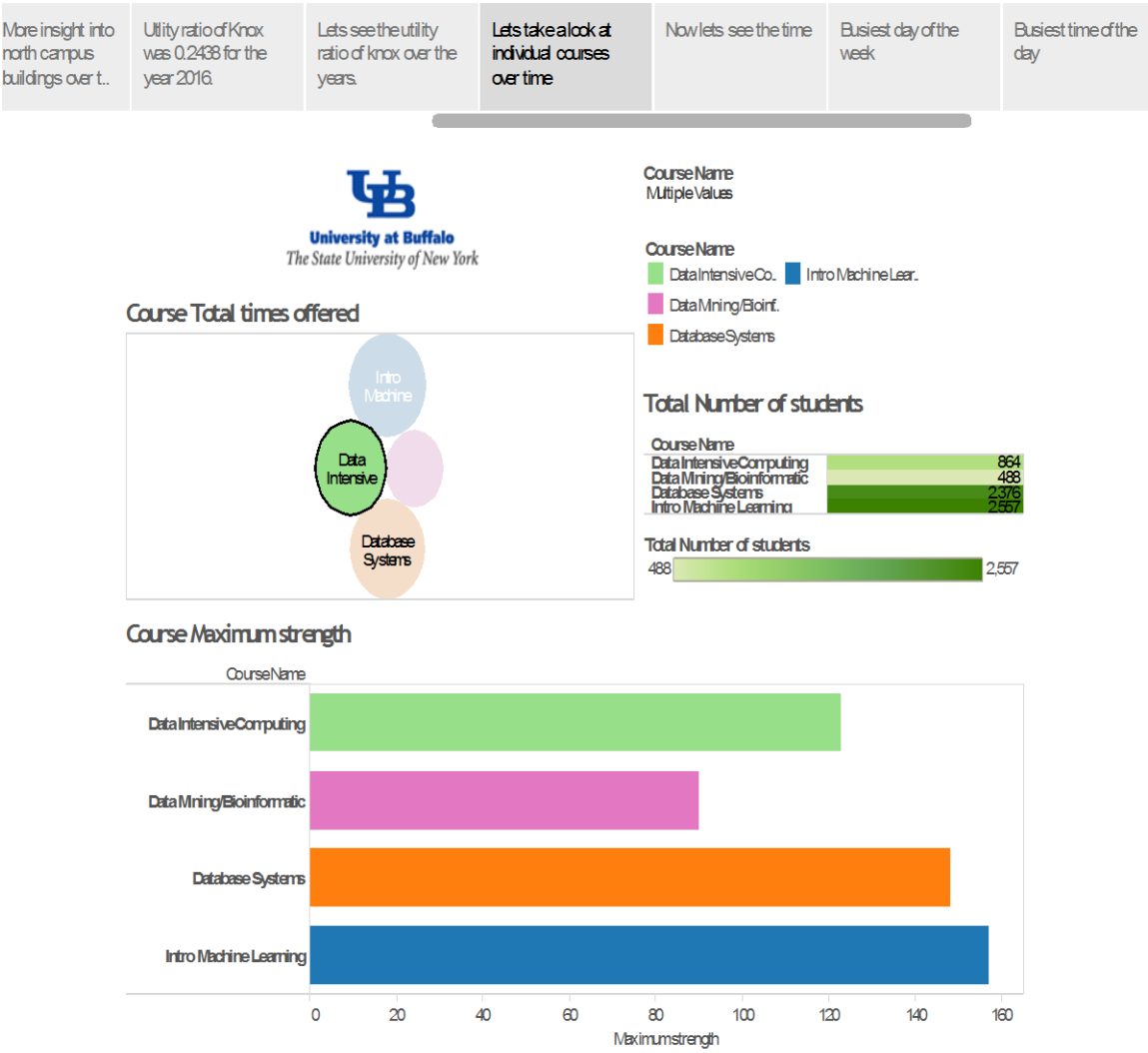
Each department might have courses that are more popular than the other. A department with low overall strength can have a very popular course and department with higher strength might have elective course which will be less taken. Therefore, it is important to get insight on each course.

So we take a look at the courses over the years to get information such as frequency of the course, maximum strength of the course, average number of students enrolled in the course.



Certain courses are very popular with average enrollment of the course over 100. The total number of courses offered from the beginning is 13993. For example, if one is interested to compare the courses related to data analysis, the following diagrams are useful. Data

Intensive Computing is among those highly demanded courses.

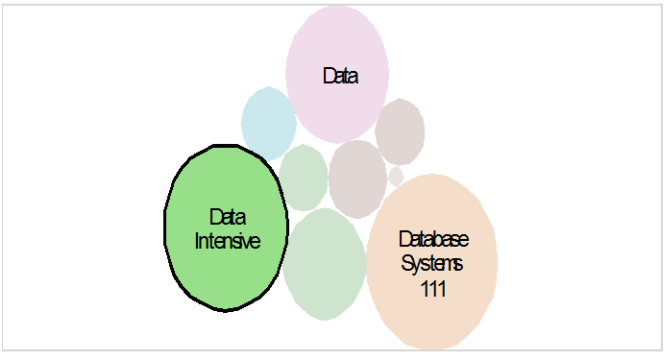


University at Buffalo
The State University of New York

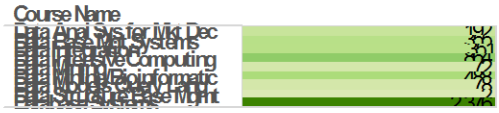
Course Name
Multiple Values

Course Name
Data Anal Sys for .. | Data Intensive Co.. | Data Models (
Data Base Mgt Sy.. | Data Mining | Data Structure
Data Integration | Data Mining/Bioinf. | Database Syst

Course Total times offered

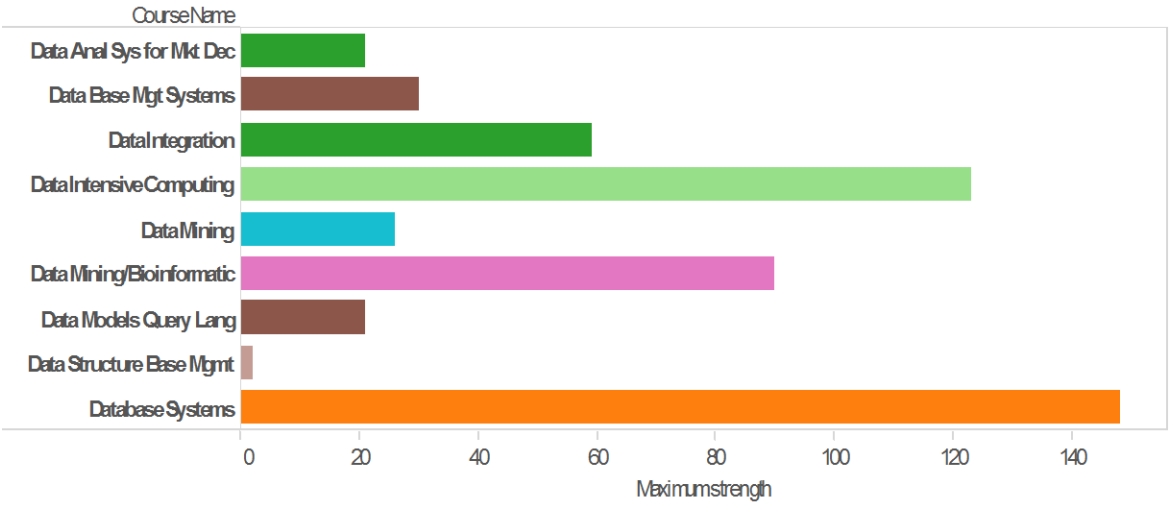


Total Number of students



Total Number of students
2 2,376

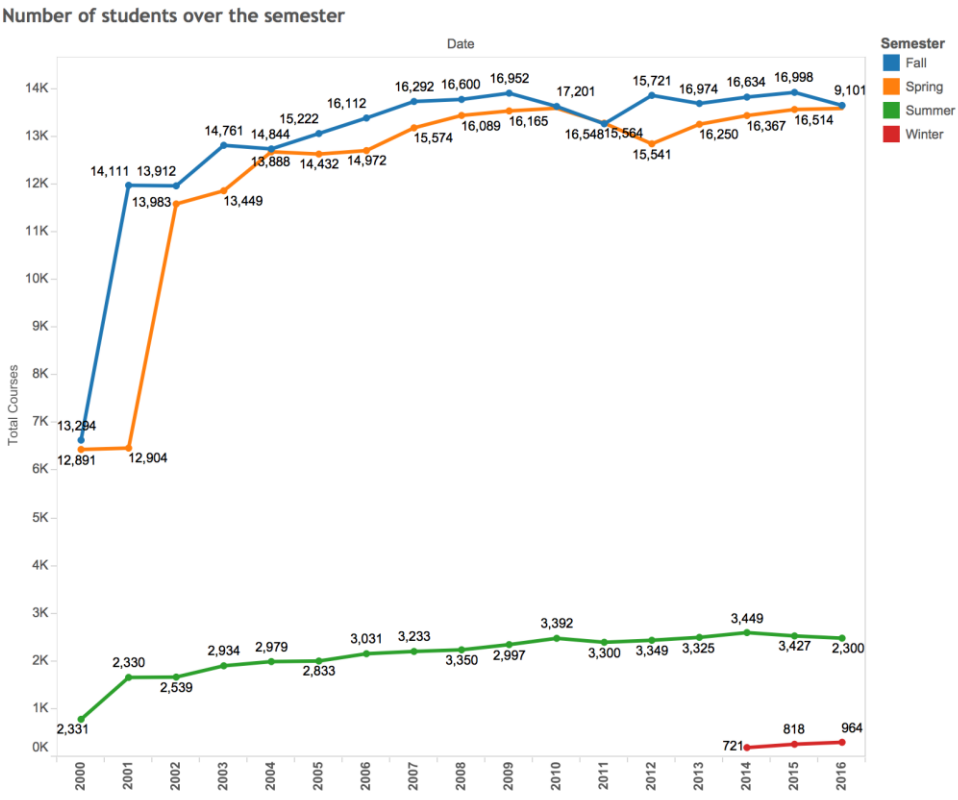
Course Maximum strength



There are four semesters in UB as of 2016.

1) Fall
2) Spring
3) Summer
4) Winter

We would like to know to number of students in each semester to decide the allocation of the room.



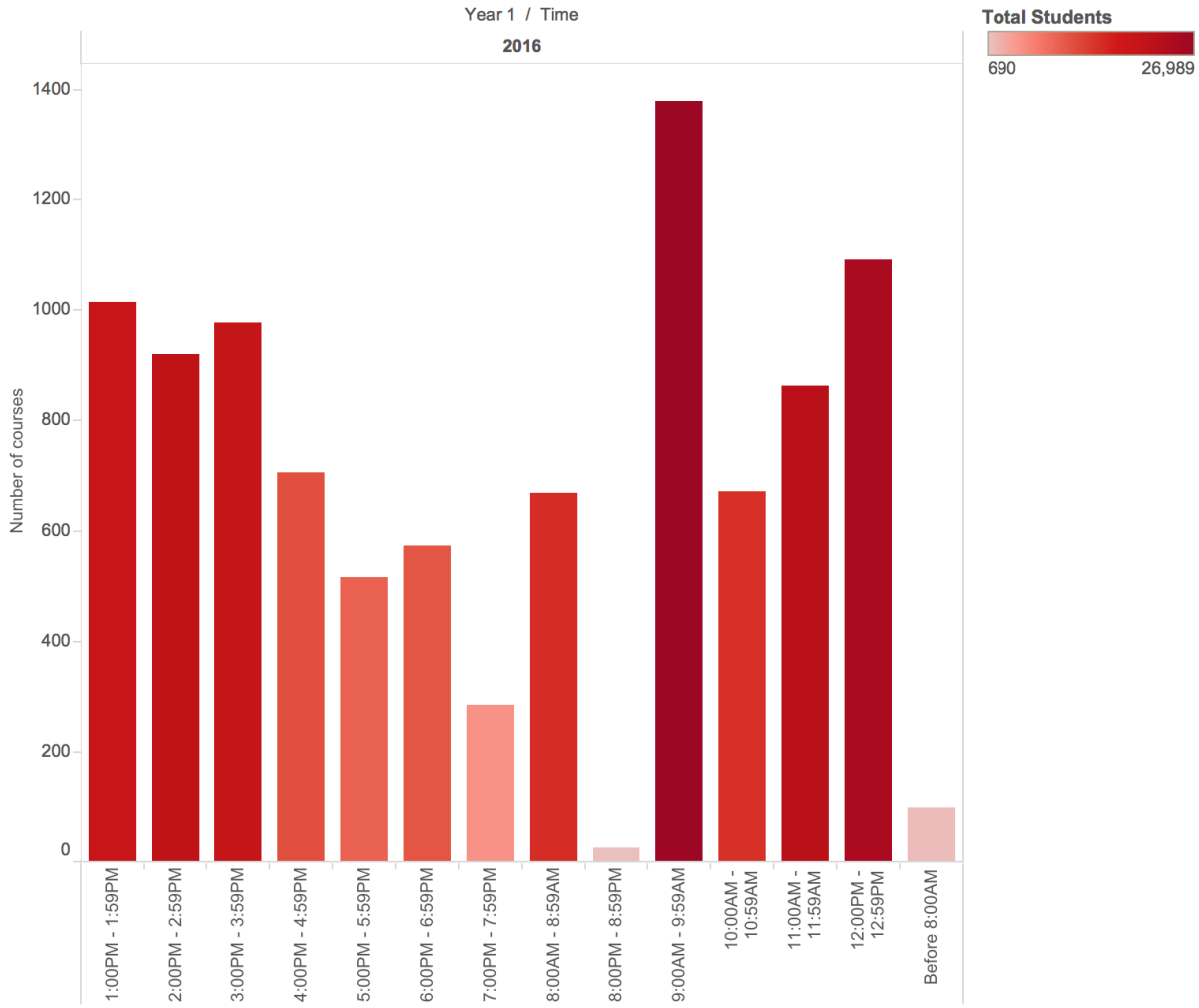
The trend of sum of Total Courses for Date Year. Color shows details about Semester. The marks are labeled by sum of Max strengt.
The view is filtered on Date Year, which keeps 17 of 81 members.

The fall and spring semester has higher number of courses offered and overall strength. The Winter semester is relatively new and has least number of courses and students. Some courses may be offered one of the two semesters, while other courses are offered both the semesters. Thus we would like to know the statistical difference in enrollment in a courses offered over fall and spring. This was done my 2 stage mapreduce program.

```
Intro to Design in Fibers      Fall = Spring
Intro to Design in Metals      Fall = Spring
Intro to Discrete Math Fall = Spring
Intro to Educ Tech     Fall = Spring
Intro to Electronic Mus 1      Fall > Spring
Intro to Eng Drawing   Fall = Spring
Intro to Env. Engr.    Fall > Spring
Intro to Federal Income Taxes Fall = Spring
Intro to Feminist Theory       Fall = Spring
Intro to Financial Accounting Fall = Spring
Intro to Geoenviro Eng Spring > Fall
Intro to Graduate Theory       Fall = Spring
Intro to Health Care   Spring = Fall
Intro to Health Care Org       Fall = Spring
Intro to Health Care Sys       Spring > Fall
Intro to Health Law    Fall = Spring
Intro to Higher Math   Fall = Spring
Intro to Hlth & Hum Srvcs      Fall = Spring
Intro to Human Geography       Spring = Fall
Intro to Illustration  Spring > Fall
Intro to Info Technology       Fall > Spring
Intro to Internat Politics     Fall = Spring
Intro to Judaism       Spring = Fall
Intro to Korean Lit    Fall = Spring
Intro to Law & Legal Proc      Fall = Spring
Intro to Library & Info Study  Fall > Spring
Intro to Ling Analysis Fall = Spring
Intro to MIS   Fall = Spring
Intro to Macroeconomics        Fall = Spring
Intro to Management Accounting Spring > Fall
Intro to Medical APY   Fall = Spring
Intro to Mgmt Info Systems     Fall = Spring
Intro to Mgmt Info Systms      Fall = Spring
Intro to Microbial Genome Ann  Spring > Fall
Intro to Microeconomics        Spring > Fall
Intro to Modern Art    Fall = Spring
Intro to Music Education       Fall = Spring
Intro to NYS Admin Law Fall = Spring
Intro to Nanoelectronics       Fall = Spring
Intro to Native Amer Hist      Fall = Spring
Intro to Native American Women Fall = Spring
Intro to New Media     Fall = Spring
Intro to Old Testament Spring > Fall
```

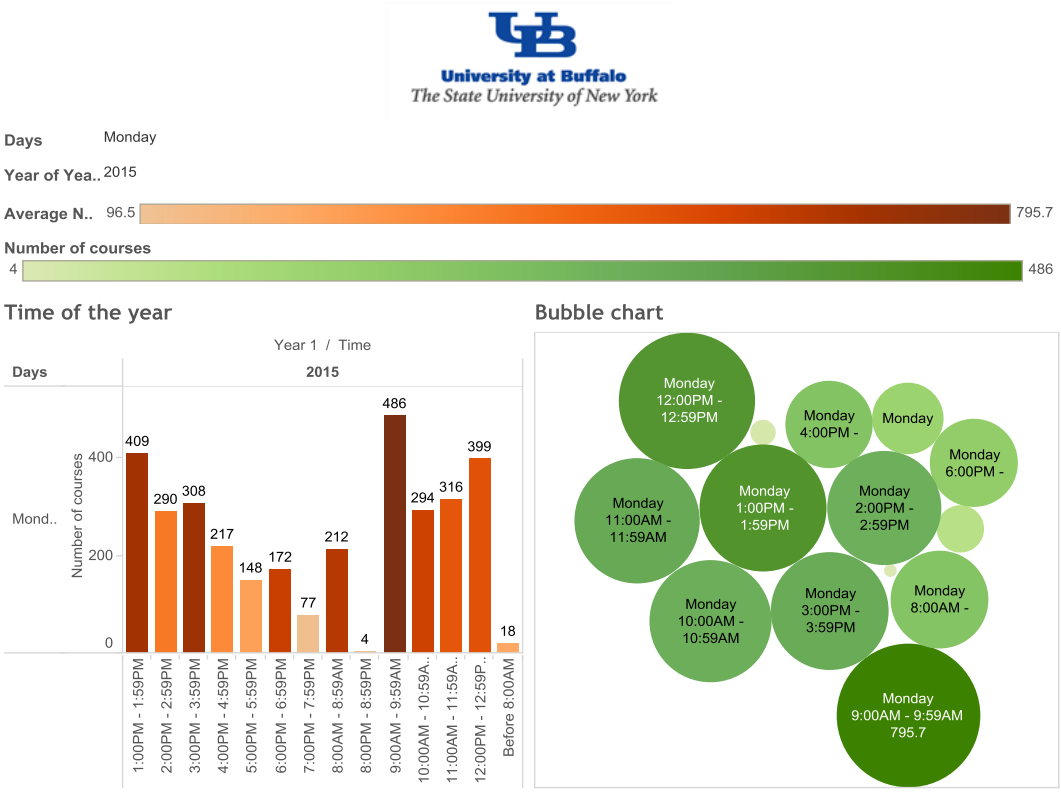
Time is important in scheduling problem. So we have to get an intuition on time. The busiest time of the day is duration where most number of classes are scheduled.

Busiest time of the day



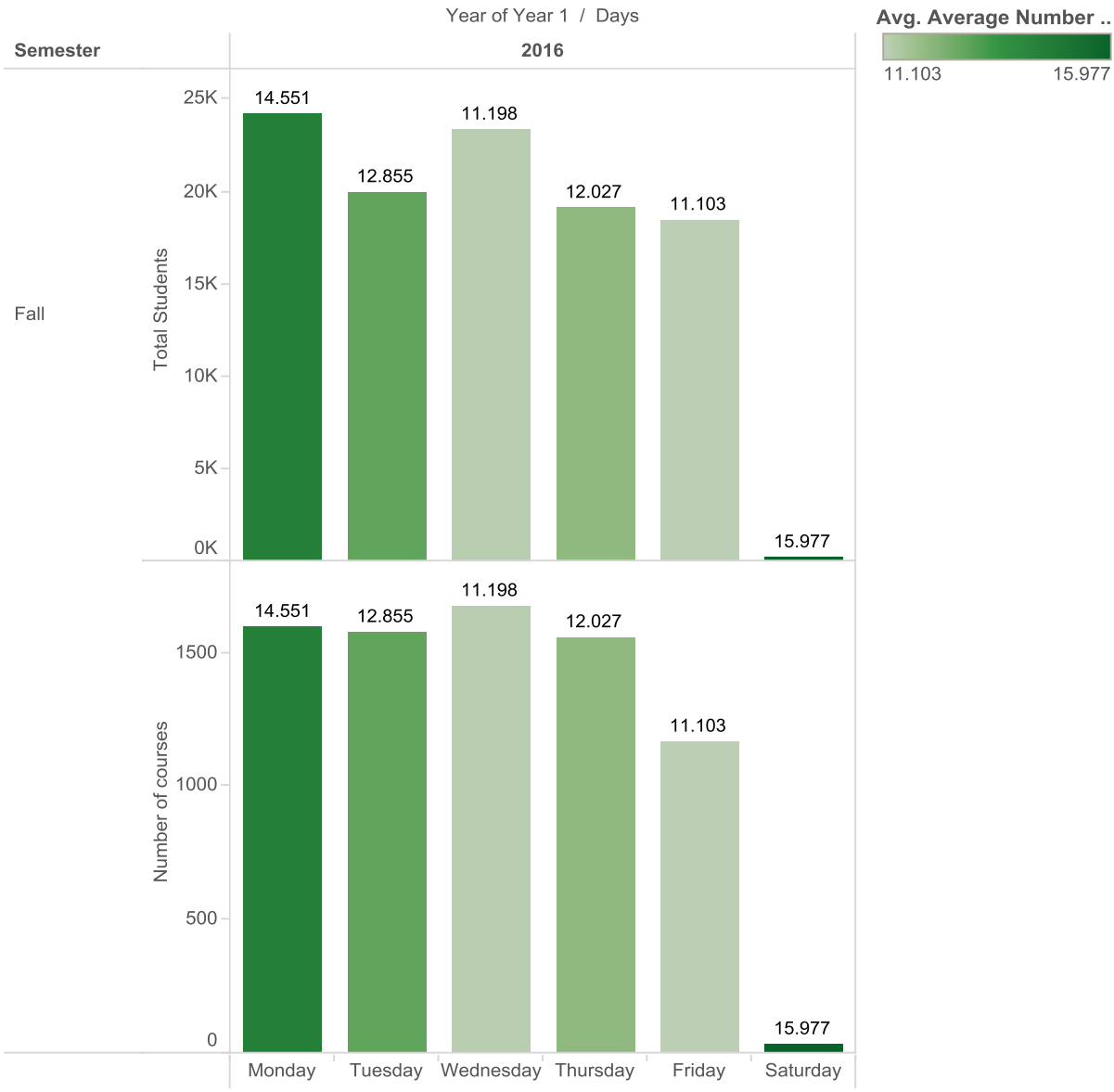
Sum of Number of courses for each Time broken down by Year 1 Year. Color shows sum of Total Students. The view is filtered on Year 1 Year, which keeps 2016.

The time between 9.00 and 9.59 AM is the busiest time of the day. The average number of course is around 700. The following diagram shows one sample on Mondays in 2015. Monday 9:00-9:50 AM is the busiest time.



The busiest day is also as follows:

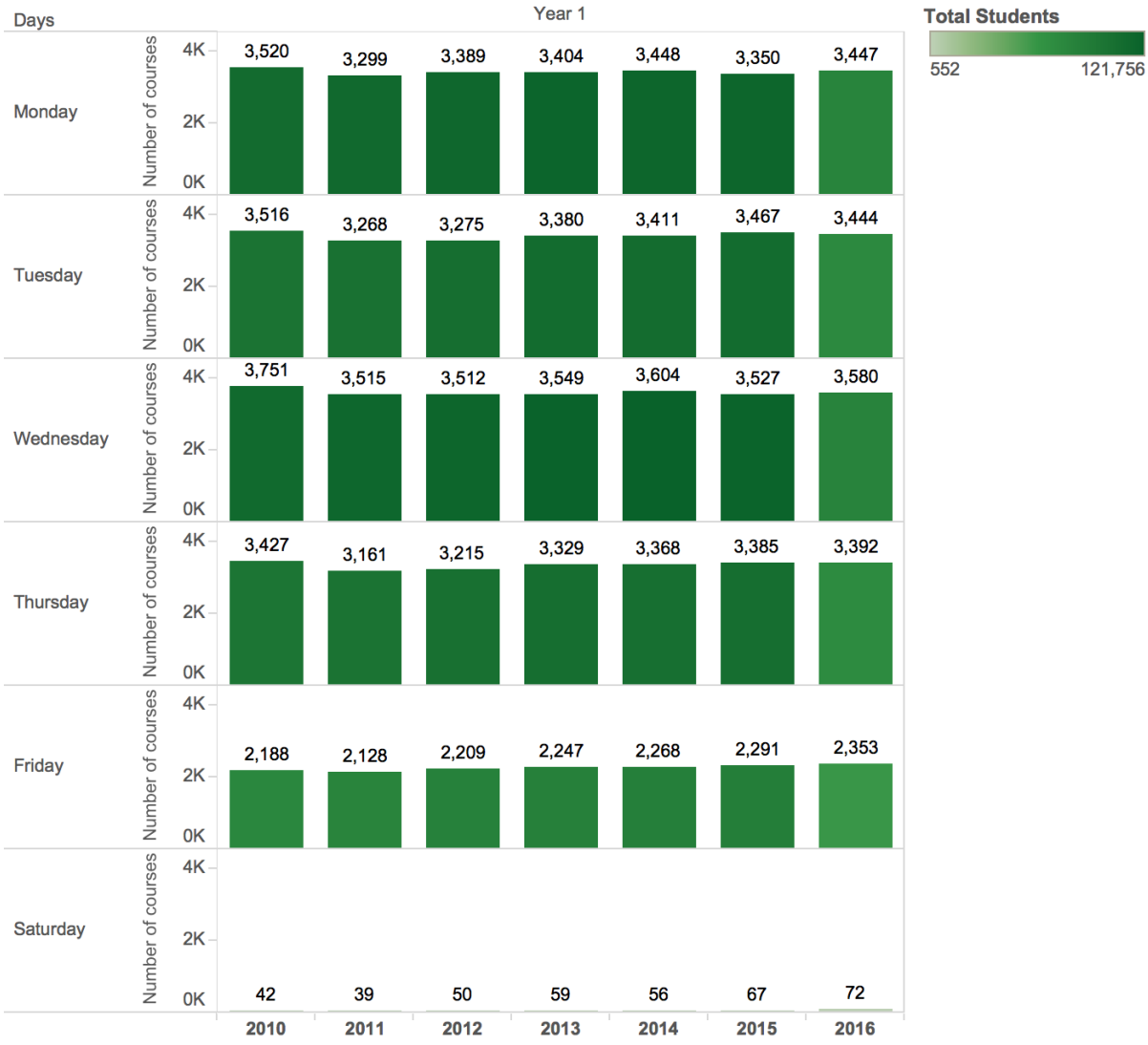
Busiest day



Sum of Total Students and sum of Number of courses for each Days broken down by Year 1 Year vs. Semester. Color shows average of Average Number of students. The marks are labeled by average of Average Number of students. The data is filtered on Year 1 Year, which keeps 2016. The view is filtered on Semester, which keeps Fall.

We would also like to take a look at variation in number of courses over the week to see if any particular day is significantly busier than the others. The below graph shows that Friday and Saturday have very less number of courses compared with other days. Amongst Monday Tuesday Wednesday and Thursday there are no significant difference.

Busiest day of the week



Sum of Number of courses for each Year 1 Year broken down by Days. Color shows sum of Total Students. The marks are labeled by sum of Number of courses. The view is filtered on Year 1 Year, which keeps 7 of 25 members.

Further a five dashboards are created with Tableau desktop for the end user to interact with the data we have processed using MapReduce.
There are different features that one can consider in class scheduling. For example, the CSE department has number of courses in which more than 50 students register. The similar analysis can be done for other departments.

Busiest time of the day						

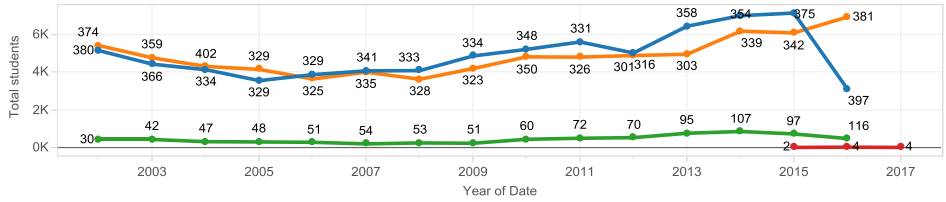
University at Buffalo
The State University of New York

Department CSE

Semester Fall Spring Summer Winter

Year
2000 to 2017

Department Information



Class size over 50