Courses and Enrollments Data Analysis and Visualization

Christina Lu Jin | 05.20.2021

This is analysis report includes three dashboard each with charts/graphs. The idea was to create more comprehensive visualizations and more in-depth analysis of the courses offered and enrollments at a particular school during year 2013 to 2016. As well as exploring the answers for specific business questions.

Data Overview and Preparation

The snippet below show the top 10 rows of this dataset, which includes 9 variables as follows:

- CourseName: name of the course
- Course Identification: course identification number, all sections of the same course have the same course identification
- CRN: course record number, a unique identifier, each section has a different CRN and no two sections have the same CRN
- PriortoFall16: a binary value indicates whether this course was taught prior to Fall 2016 (being 0) or was it taught in Fall 2016 (being 1)
- Term: The year and semester that course/section was taught
- TotalEnrollment: total number of students enrolled in this section
- NoGrade: number of students who did not receive a grade (received an incomplete in the course)
- WithdraworFailingGrade: number of students who withdrew or failed the course/section
- PassingGrade: number of students who received a passing grade

CourseName	Course Identification	CRN	PriortoFall16	Term	TotalEnrollment	NoGrade	Withdrawor	PassingGrade
History of the World	HIS1234	50001	0	Winter 2014 Quarter	7	0	0	7
History of the World	HIS1234	50002	0	Spring 2015 Quarter	10	0	0	10
History of the World	HIS1234	50003	0	Spring 2016 Quarter	12	0	0	12
History of the World	HIS1234	50004	1	Fall 2016 Semester	5	0	0	5
History of the World	HIS1234	50005	1	Fall 2016 Semester	3	0	0	3
Political Science	RTV2345	50006	0	Fall 2013 Quarter	22	0	5	17
Political Science	RTV2345	50007	0	Spring 2014 Quarter	19	0	4	15
Political Science	RTV2345	50008	0	Spring 2014 Quarter	6	0	0	6
Political Science	RTV2345	50009	0	Summer 2014 Quarter	13	0	2	11

Once receiving the data, I decided to some preparation for easier processing later with visualization tool. The first thing I did was to extract the first three letters, from the course identification column and specify it as course categories, since many different courses may be classified into the same category. Secondly, I noticed that the term column includes both the quarter and the year the sections were offered. Since I will be investigating the data both by year and by quarter, so I take this column and separated it into three attributes – year, quarter, and year and quarter combined. The following snippet shows the result of the preparation.

CourseName	Course Category	Course Identification	CRN	PriortoFall1	Year	Quarter	Year_Quarter	TotalEnrollment	NoGrade	Withdrawn	PassingGrade
History of the World	HIS	HIS1234	50001	0	2014	4-Winter	2014 4-Winter	7	0	0	7
History of the World	HIS	HIS1234	50002	0	2015	1-Spring	2015 1-Spring	10	0	0	10
History of the World	HIS	HIS1234	50003	0	2016	1-Spring	2016 1-Spring	12	0	0	12
History of the World	HIS	HIS1234	50004	1	2016	3-Fall	2016 3-Fall	5	0	0	5
History of the World	HIS	HIS1234	50005	1	2016	3-Fall	2016 3-Fall	3	0	0	3
Political Science	RTV	RTV2345	50006	0	2013	3-Fall	2013 3-Fall	22	0	5	17
Political Science	RTV	RTV2345	50007	0	2014	1-Spring	2014 1-Spring	19	0	4	15
Political Science	RTV	RTV2345	50008	0	2014	1-Spring	2014 1-Spring	6	0	0	6
Political Science	RTV	RTV2345	50009	0	2014	2-Summer	2014 2-Summer	13	0	2	11

Business Questions

In order to provide meaningful analysis towards this dataset, I've came up with the following four business questions:

- How did this school perform over the years from 2013-2016 with given courses?
- What are some popular courses?
- How did the students do over the years?
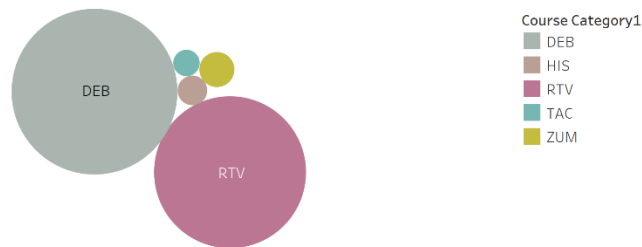- How to improve in later years?

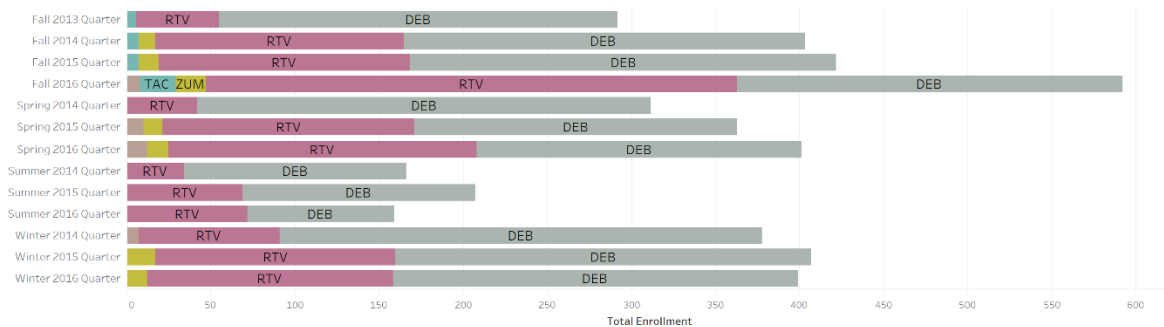## Visualization Dashboard Overview
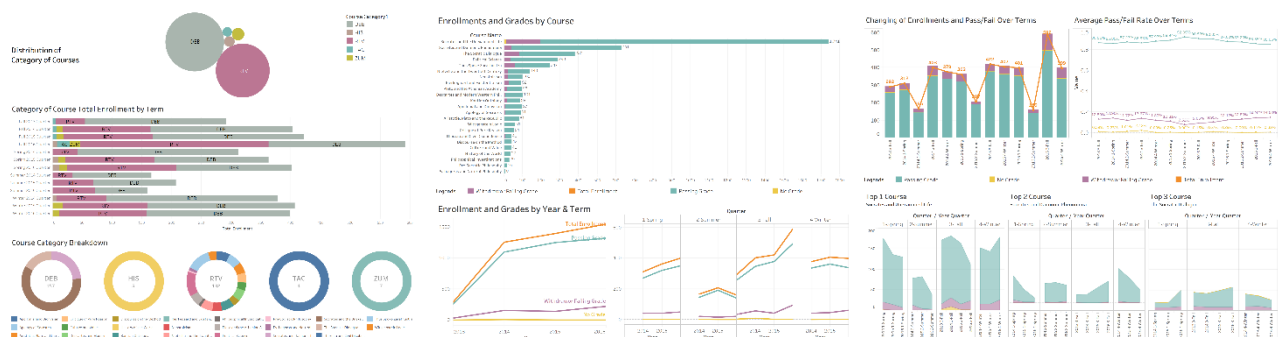
The image above shows a brief overview of the dashboard that I created of this dataset for solving the business questions mentioned in the section above. Data visualization could be really helpful and informative in delivering messages, intuitively showing some hidden patterns or correlations, and providing changes over time or between different categories.
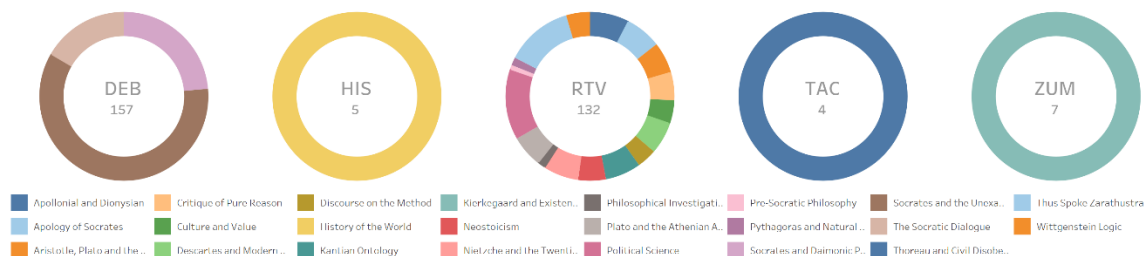
Data Visualization and Analysis – Dashboard Part 1

Distribution of Category of Courses

Course Category1
- DEB
- HIS
- RTV
- TAC
- ZUM

Category of Course Total Enrollment by Term

Course Category Breakdown

DEB 157	HIS 5	RTV 132	TAC 4	ZUM 7

Legend:
- Apollonial and Dionysian
- Apology of Socrates
- Aristotle, Plato and the ..
- Critique of Pure Reason
- Culture and Value
- Descartes and Modern ..
- Discourse on the Method
- History of the World
- Kantian Ontology
- Kierkegaard and Existen..
- Neostoicism
- Nietzsche and the Twenti..
- Philosophical Investigati..
- Plato and the Athenian A..
- Political Science
- Pre-Socratic Philosophy
- Pythagoras and Natural ..
- Socrates and Daimonic P..
- Socrates and the Unexa..
- The Socratic Dialogue
- Thoreau and Civil Disobe..
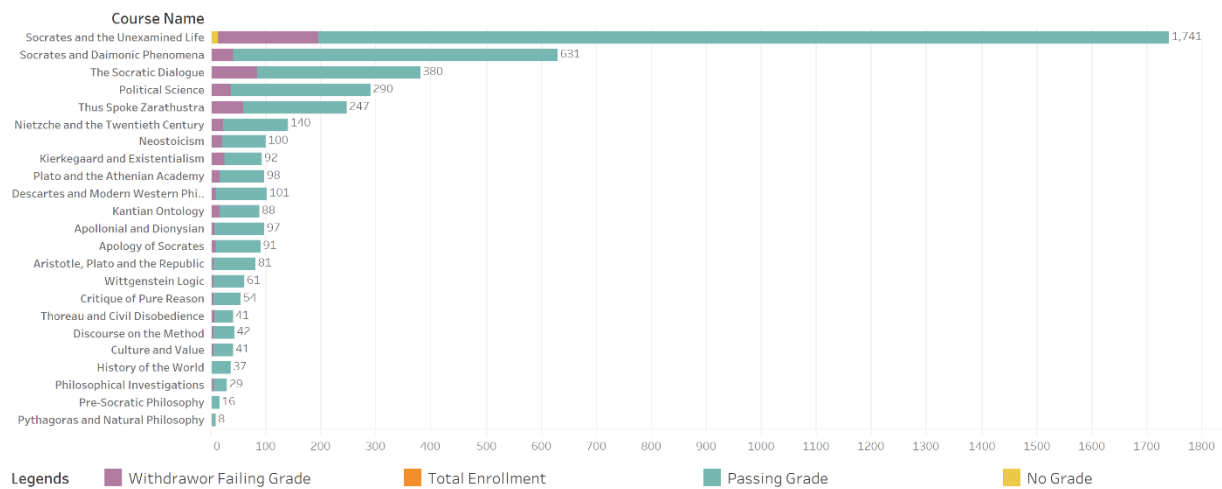- Thus Spoke Zarathustra
- Wittgenstein Logic

This first dashboard displays a general overview of the all the course categories offered in this university. The bubble chart on the very top shows the categorical course situation during 2013 and 2016, color coded by each category. Sizes of circles are indicators of the total number of courses, it indicates that DEB and RTV had large amount of courses than other categories.

The bar chart on the middle shows the total enrollment of each term, color coded by course categories. From the chart we can see that DEB and RTV do have significantly more enrollment than all others. Between the two, DEB has even more enrollments than RTV for most of the terms except Fall 2016. Also, couple other observations are, firstly there was an increase among most terms over the years, except summer quarter; secondly, summer quarters generally has the lowest enrollments; and lastly, there is a huge spike happened in Fall 2016.
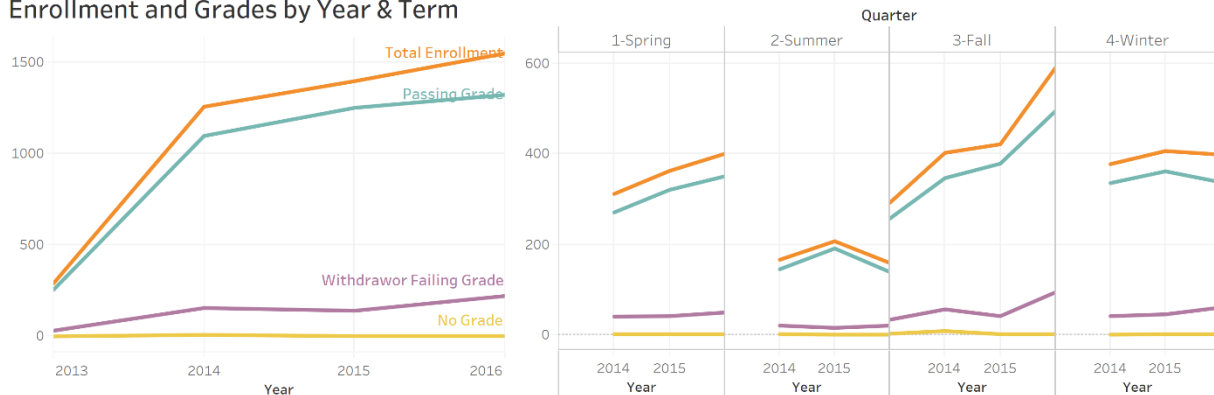
The donut chart, on the bottom, displays the breakdown of the courses and sections composition within each category, each color indicates an individual course, and the numerical value in the middle shows the total number of sections. The donuts show that courses are very unevenly distributed, HIS, TAC, and ZUM only have 1 course, DEB has 3 courses, while RTV has 17 distinct courses. And the course "Socrates and the Unexamined Life" has almost 100 sections, which is the most among all courses.

Data Visualization and Analysis – Dashboard Part 2

### Enrollments and Grades by Course

Course Name	
Socrates and the Unexamined Life	1,741
Socrates and Daimonic Phenomena	631
The Socratic Dialogue	380
Political Science	290
Thus Spoke Zarathustra	247
Nietzche and the Twentieth Century	140
Neostoicism	100
Kierkegaard and Existentialism	92
Plato and the Athenian Academy	98
Descartes and Modern Western Phi..	101
Kantian Ontology	88
Apollonial and Dionysian	97
Apology of Socrates	91
Aristotle, Plato and the Republic	81
Wittgenstein Logic	61
Critique of Pure Reason	54
Thoreau and Civil Disobedience	41
Discourse on the Method	42
Culture and Value	41
History of the World	37
Philosophical Investigations	29
Pre-Socratic Philosophy	16
Pythagoras and Natural Philosophy	8

Legends   ◼ Withdraw or Failing Grade   ◼ Total Enrollment   ◼ Passing Grade   ◼ No Grade

Enrollment and Grades by Year & Term

The second dashboard explores the changes and relationship between different measurements of academic outcomes, which includes number of students with passing grade, withdraw/failing grade, and those with no grade. The first stacked bar chart shows the composition of these three measurements for each and every single course, and the added-up value of them all is the total enrolment. The chart displays the drastic difference in enrollment between "Socrates and the Unexamined Life", this course itself has almost as much enrollment as the sum of all other courses, which determines this course to be the most popular one. Some courses with high withdraw/failing and no grade rate are The Socratic Dialogue (21.1%), Thus Spoke Zarathustra (21%), and Socrates and the Unexamined Life (11%).

The line chart on the bottom consist of two parts, left one shows the changing of each of the measurements over the years, though the right one compares the changes by terms of each year. The only term that took place in 2013 was the fall term, all other terms started from 2014. Obviously, the overall number of enrollments, passing grade and withdraws slightly grows throughout the years. But in between each quarter, total enrolment constantly increases from 2014 to 2016 spring, the number of withdraw and failing grades didn't change much. Summer

and Winter terms have similar patterns, though winter term has almost double the amount of summer. Fall is the term that had the most significant increase, the totally enrollment had a 102.74% increase from 2013 to 2016, there was a huge jump happened from fall 2015 to fall 2016, this could possibly due to the change of the structure of courses took place in term fall 2016. Though, what draws a concern is that there was a greater increase in the number of withdraw/failing grades too. This could be due to either that students are not very familiar with the new course structure, or the courses became harder and resulting in lower grades.

Data Visualization and Analysis – Dashboard Part 3
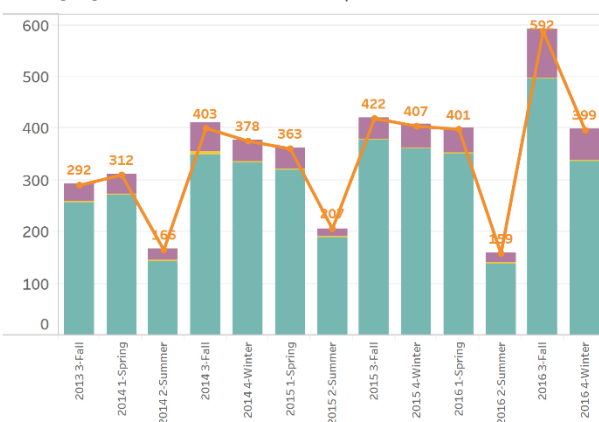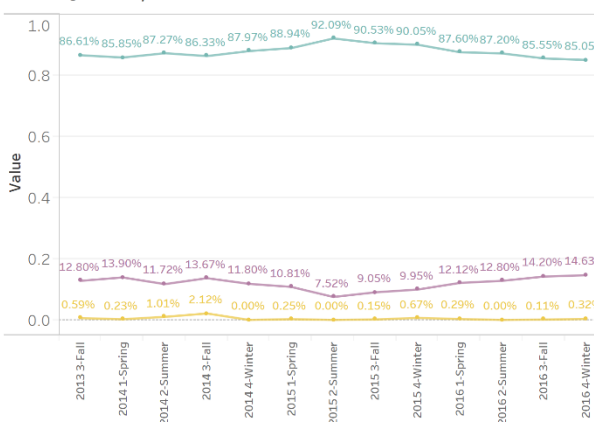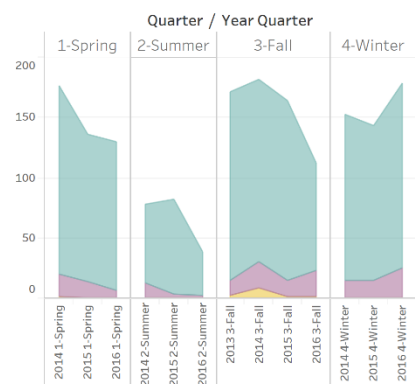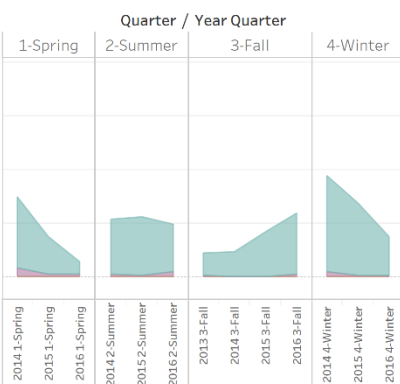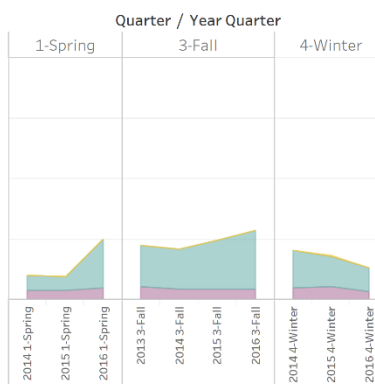
This third part of the dashboard showcases the change and rate of these three measurements over the terms and also for the top three popular courses. The top left chart is a combination of stacked bar chart, displays the number of passing, withdraw and no grade, and line chart, displays the total number of enrollments (which is the sum of the three). Here we can clearly see the dips at summer quarters, as well as the surge in term fall 2016. The line chart on the top right shows the average passing, withdraw/failing, and no grades rate. These rates are extra

attributes that I've created using the existing values, passing rate equals to the number of students with a passing grade of that term divided by the total number of enrollment of the same term, and times a hundred percent to convert it to a percentage value. Similar procedures for the withdraw/failing rate and no grade rate. From the trend we can see that the no grade rate was the highest for term fall 2014, but overall hold a pretty constant and rate, not much up and downs. While looking at the passing rate, there was an upward increase trend from term fall 2013 to summer 2015, but a downward decreasing trend appeared after that. Whilst, withdraw/failing rate had a completely opposite pattern of trend with passing rate, it was decreasing before summer 2015, and start increasing after that. Which make sense, since these two factors are positively linearly correlated.

The bottom graph is a combination of three area charts, each shows the term comparison of the three measures of the top 3 most popular courses. The most popular course, Socrates and the Unexamined Life, on the left most, is showing a general decrease in total enrollments except winter term. The high withdraw/failing rate is also a concern in this class, looks like there was an increase in the withdraw rate for this course especially in fall and winter terms, this could due to many reasons such as the difficulty in the course, the unfamiliarity of the course structure, the strict course policy made by the professor. Moving on the second popular course, Socrates and Daimonic Phenomena, in the middle, shows a drastic decrease in both spring and fall terms. Summer term didn't change much over the years, but the enrollment for the fall term had a huge increase especially after year 2014. The withdraw/failing rate though looked pretty low for this class, this show that the course material is probably a little easier to understand, and the course structure is more friendly with the students. Lastly, the third popular course, The Socratic Dialogue, on the most right, this course was never offered in the summer quarter, since there is no data for summer. For the rest of the three terms, a sign of increase had been shown for spring and fall terms, and a decrease has shown for the winter term. This course has a moderate withdraw/failing rate, and the number of withdraws didn't change much over the years.

Business Insights and Additional Takeaways

To answer the business questions that we proposed in early sections, I came up with the following sights:

How did this school perform over the years from 2013 -2016 with given courses?

- There was some increase in enrollments over the years but not very significant. Fall term of 2016 had a jump in enrollments, investigate the change of course structure for that term. Consider implementing it in later terms, as well as branding and promoting the school and the courses.

What are some popular courses?

- DEB category including Socratic Topic are quite popular, consider opening up more course or sections to allow more students joining.

	How did the students do over the years?	• A decent amount of withdraw and failing rate, and there was a sign of increasing starting from term Summer 2015. Consider either making the course easier or giving more options of extra bonus point to lower withdraw / failing rate and to not discourage upcoming students.
	How to improve in later years?	• Pull some surveys among all current students about how they like and don't like about the course and professor, and then improve courses accordingly.

Besides the findings and insights above, I also had concluded some additional takeaways as follows that I would like to make suggestions to the university:

- The university should consider offering more courses in different categories, instead of just focusing on two, for the variety of interest from students.
- It is observed that Summer terms generally have less enrollments comparing to the other three. University should reconsider to allocate human resources and facilities wisely to have more resources in Spring, Fall and Winter terms.

References

1) Smoak, A. Create Multiple KPI Donut Charts in Tableau. Retrieved from https://www.youtube.com/watch?v=5hMY95vgnlg
2) Smoak, A. Stacked Bar Chart Totals in Tableau. Retrieved from https://www.youtube.com/watch?v=CylyFY0UIj0&t=17s
3) How to combine text from two or more cells into one. Retrieved from https://support.microsoft.com/en-us/office/combine-text-from-two-or-more-cells-into-one-cell-81ba0946-ce78-42ed-b3c3-21340eb164a6
4) https://www.youtube.com/watch?v=5hMY95vgnlg&t=1s