Adam Levick

SI 618 Final Project: Student Evaluations

1. Motivation

In the burgeoning field of Learning Analytics, understanding students through collected educational data is opening up new ways to view the educational systems we take for granted. With access to student end of the semester evaluation data through the Learning Analytics Fellows program at CRLT, I had the opportunity to explore a rich dataset. While it is often more useful to look at difference between similar classes, I wanted to see if I could make comparisons between schools, subjects, and class types.

Research Questions:

1. Which departments are the most liked and most hated?
2. What class levels (100,200,300,400) do students like the most?
3. Do different departments differ in the class levels that are most liked?
4. How does the size of the class relate to how much it is liked?

2. Data Source

For this project I used data provided by the University of Michigan Registrars Office through a Fellows project I am currently participating in with Kevyn Collins-Thompson (esteemed SI Faculty member). This data set contains all responses to student evaluations from the Fall 2013 semester. The data came in a csv file containing variables including:

- Subject: text
- Class ID: int
- Class Section: int
- Course Type: text, Lecture (LEC), Seminar (SEM), Discussion (DIS), Independent study (IND) or Lab (LAB)
- Evaluation response rate (columns for population and responses):
- Question Key (I focused on only a few questions): int
- One column for each of the 6 response types from Strongly agree to Strongly Disagree and NA: int
- The median of the responses (if strongly agree was 5 and strongly disagree was 1): float

There are over 200,000 rows in this file, but the subset I used contained 4061 rows. I have access to this file through the M+Box system.

3.Research Questions and Methods

Python Cleaning: Initial data cleaning was done in python. I identified pertinent fields and created a tsv including the fields mentioned above and calculated a response rate. Only question 1 and question 2 were added to this tsv (though I do not explore q2 in this report), and only when the response number was greater than 10. I also added a "class level" (first integer of class_number) column to be a factor in r. I didn't experience any difficulty inputting the tsv into r. There were no blank fields, but many duplicate ones.
R Cleaning: The input data came to 8956 unique rows, 4061 entries for q1.

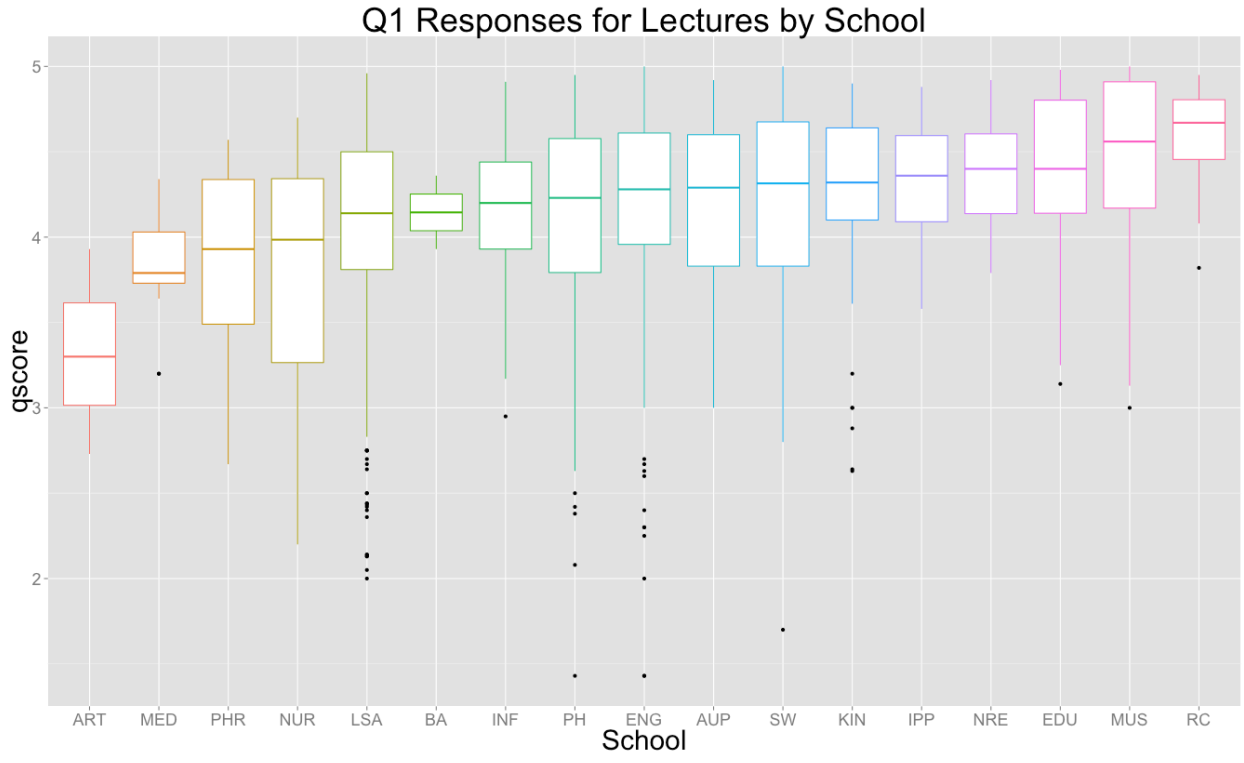
1. **Which departments are the most liked and most hated?**
   In order to compare schools, I took the subset of q1 data, reordered by median score for each school and then graphed the boxplots. I also took a subset of the lectures, because lectures were the only class type that all colleges had. Also, because discussion sections may be affected by the lecture they are discussing, weighing large lecture courses more highly than other class types, I focused on lectures. I also looked at which school had the best response rate for q1 using similar methods.
2. **What class levels (100,200,300,400) do students like the most?**
   I decided to look only at LSA undergrads for this question, as that school had the largest sample size and would be the most useful for the University to understand. I did a basic plot of the class number vs qscore by layering geom point and stats mooth. I then created graphs with a free_x facet grid by the class level (100, 200, 300, and 400), with geom point and stat smooth. I then used the factored class levels and created boxplots for each class type/level. I then looked again at class number, faceting both by class level and class type to look for any interesting in level differences (using point and smooth).
3. **Do different departments (subjects) differ in the class levels that are most liked?**
   For this question I also decided to look at undergrad LSA, since it has a more diverse range of subjects represented than other schools. I graphed qscore vs class level by department, but too many courses came up to make sense of, so I decided to further split the data by looking only at STEM courses (chosen by me). I first faceted by the department to easily relate within a department. I then faceted by level and department to more easily compare between levels of different departments, but did not include this graph in the results.
4. **How does the size of the class relate to how much it is liked?**
   The challenge I ran into with this question was how to scope the data to see trends both over all of the data (100s of students differences), but also smaller differences (10s

of students differences). I capped the shown data to help with this problem (coord_cartesian). I broke data into different groupings and class types to look for trends. I then did a kmeans clustering with 2,3,4,5 which showed a similar trend to the plotted data with the non-normalized data, split into groups by class size (I did not include this graph). The most interesting finding is shown below.

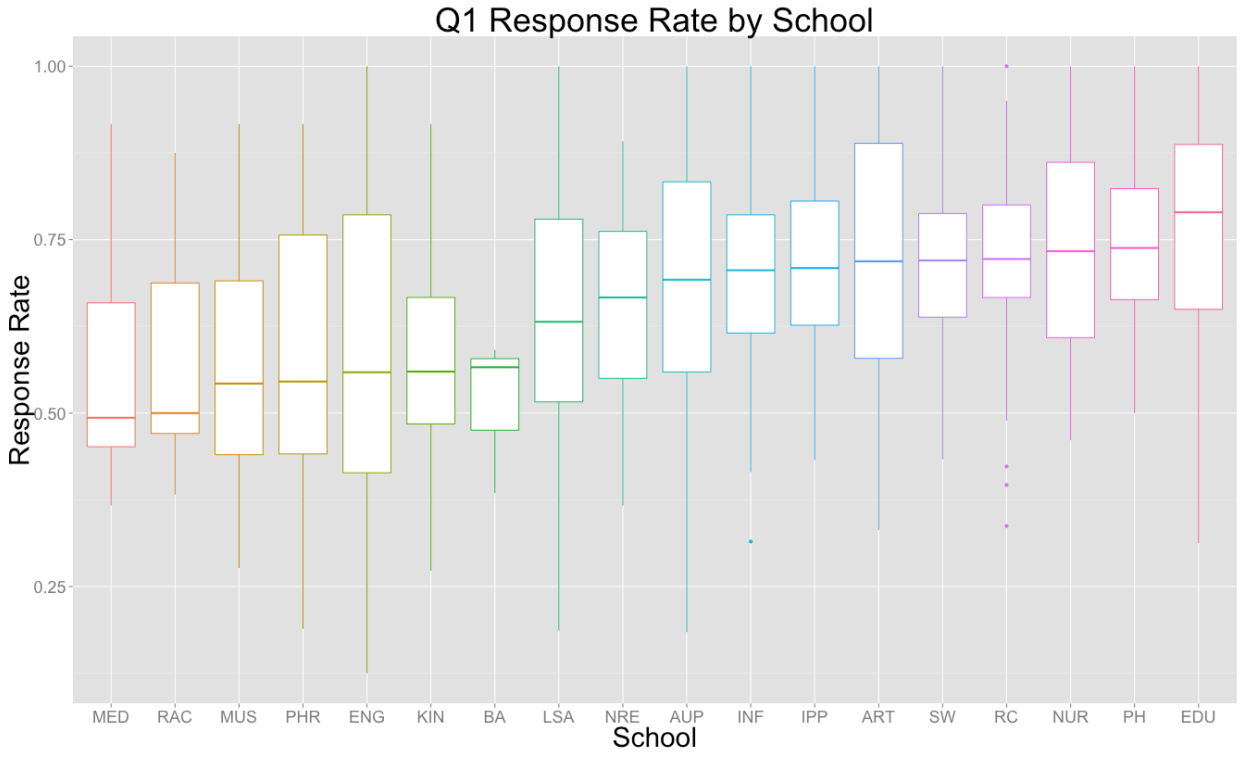
4. Analysis and Results

1. Which departments are the most liked and most hated?

The answer to this question could be extremely useful for the University at large, which could look at programs that do particularly well or poorly in different areas and then try to identify what makes students like or dislike the courses within. While this question is nebulous and maybe not particularly useful (students with different evaluation strategies may be grouped in one school instead of another). I decided to look at lectures for the reasons above, and for that the Residential College (RC) has the highest q1 scores of any school. The RC supports a very specific kind of community, and this visual shows that other schools might benefit from such a community. Their Lectures may also be taught in certain ways. At the same time, the Art school lectures (there are only 3) seem to be disliked by Art students, and perhaps a change of format should be considered.

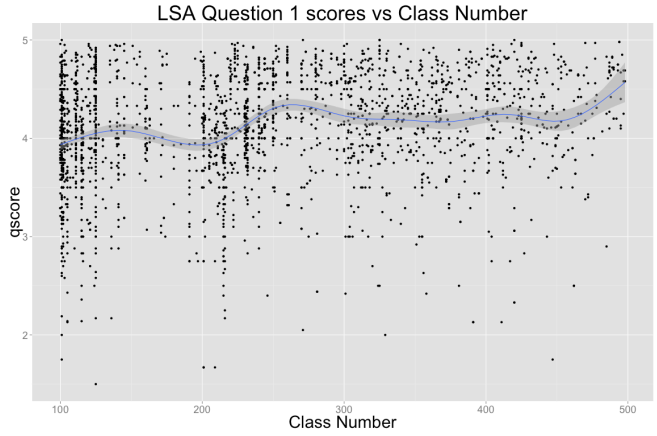


I think it is also useful to see that which schools are able to get the best response rates. The school of Education should be consulted to see what strategies they use to get students to give feedback.

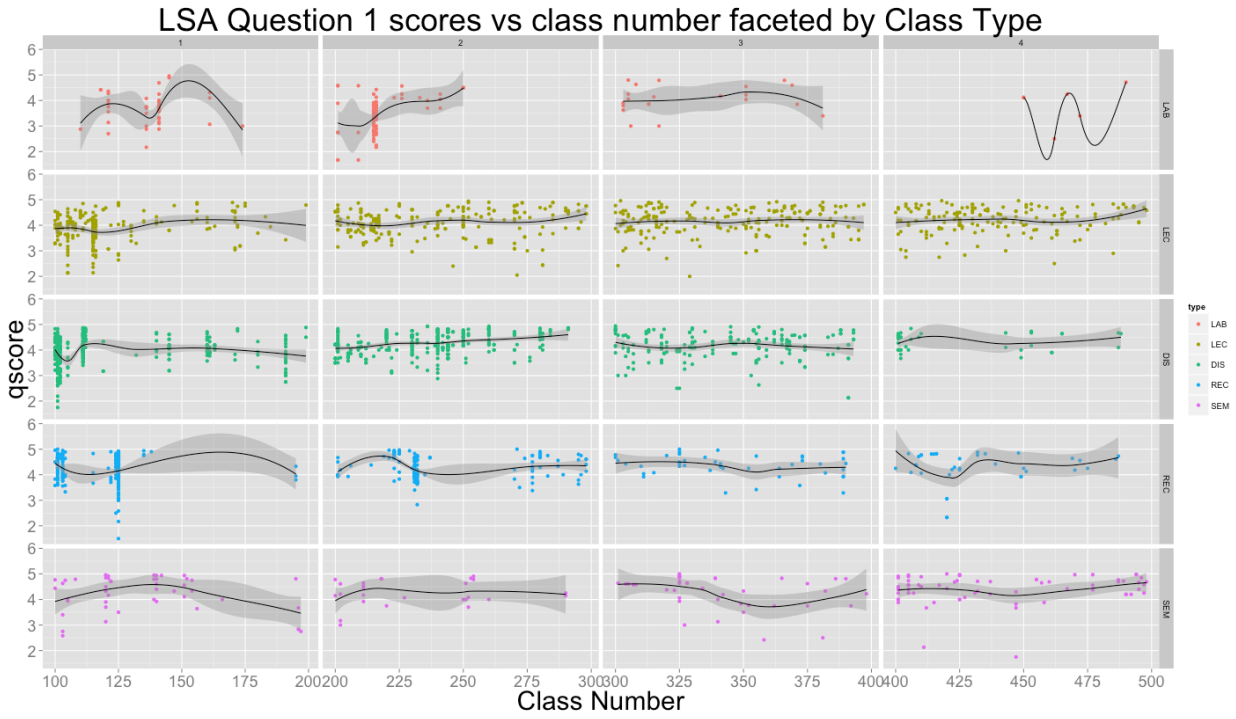


2. What class levels (100,200,300,400) do students like the most?

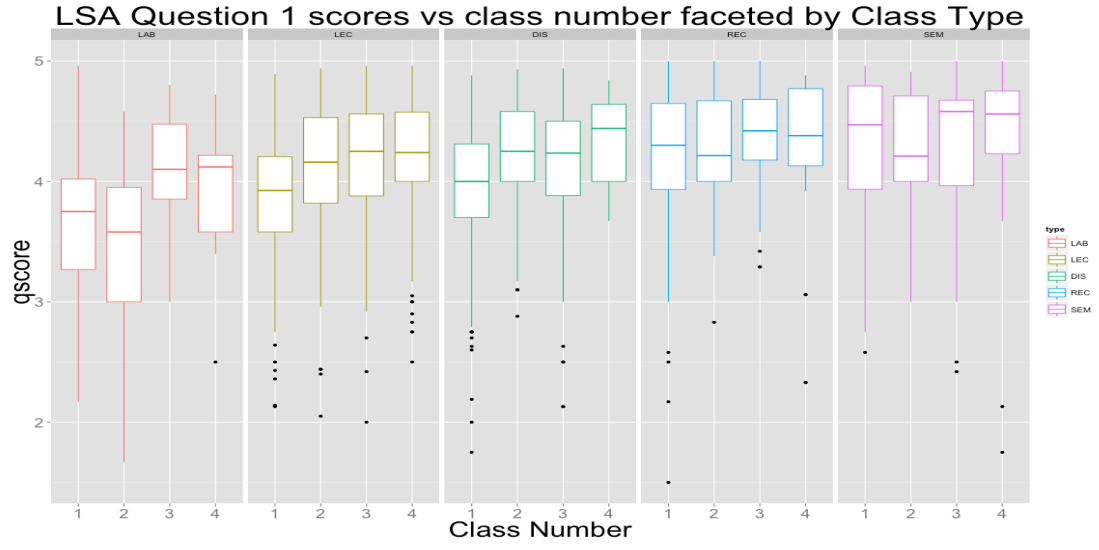
For this question I concentrated on Undergraduate courses in LSA. I first did a basic scatter to compare class number and score. It showed a clear jump from 100 to 200 level classes, and at the upper end of 400 level courses.



Exploring this further, I looked for any interesting in level differences by class type. Things came out fairly flat for most areas, though 200 level lectures showed a clear increase in score the higher they got. We are going to ignore 400 labs.

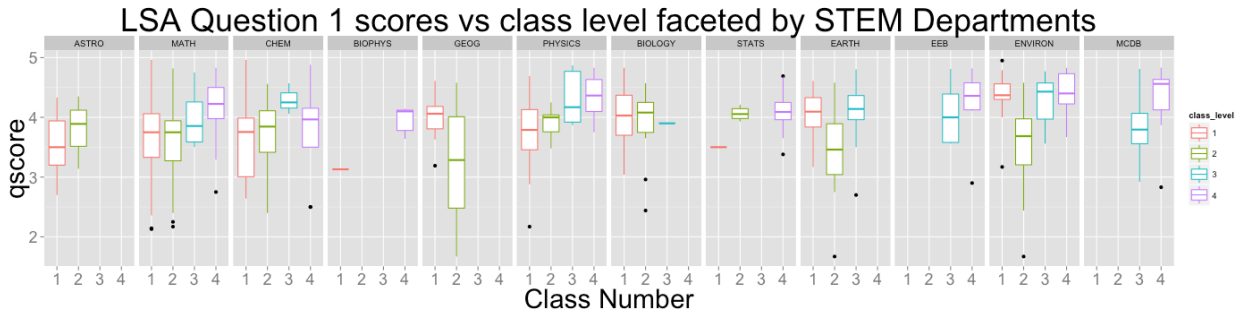


The most useful graph appears below, showing the relationship between class type and level. You can clearly see which class types are more liked in this visual, as well as the general upper trend in qscore as the level of class increases. Notable differences include the 200 level labs (which include stats and orgo classes, large and difficult courses) and the 200 level seminar median is much lower than other levels.



3. Do different departments differ in the class levels that are most liked?

An easy question to answer, I first looked at all departments in LSA at once, but didn't find it to be very useful (and too cluttered). I decided it would be more interesting to look just at STEM departments. Interestingly, both EARTH, ENVIRON, and GEOG seem to have a dip in the 200 level courses, whereas most departments have only an upward trend.



4. How does the size of the class relate to how much it is liked?

Often students will cite large lecture classes as a drawback of going to a larger university. The graph below shows that class size does seem to have an effect on how excellent student's believe a course is, but only up to around 100 students, at which point it flattens out. This could say that for large undergraduate courses, 1 section of 300 students would be no worse than 2 sections of 150.