Nicholas Lichtsinn

IST 707

Homework 2

**Introduction:**

**General:**

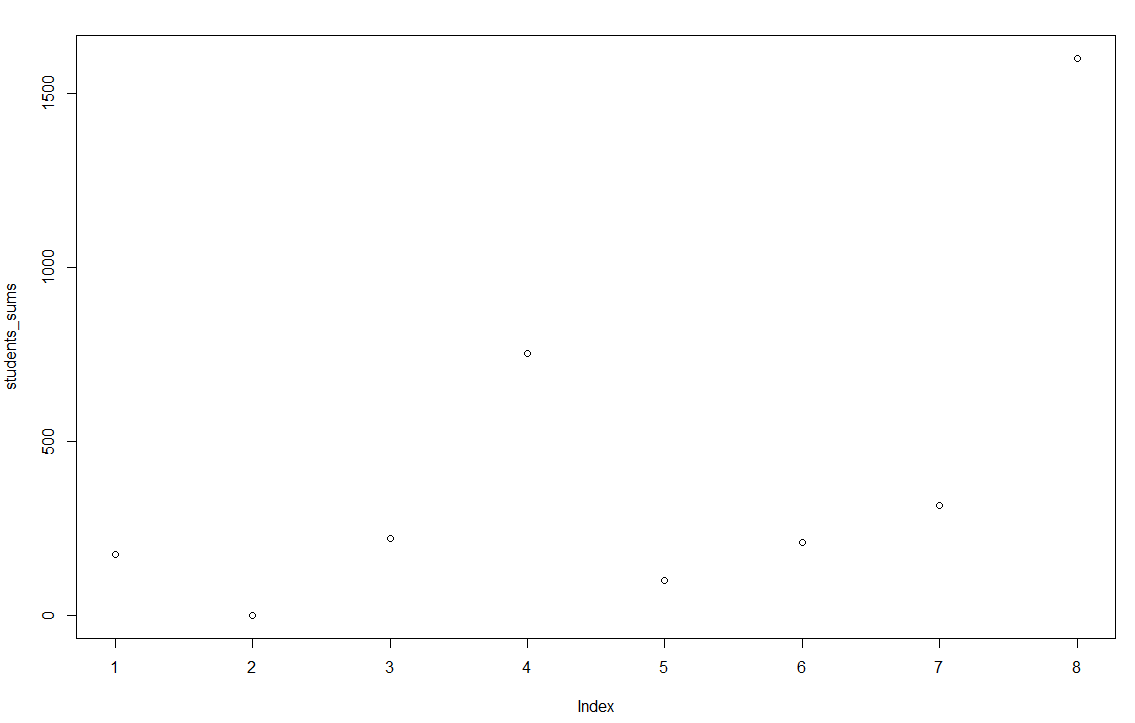
Students' education varies from school to school yet we look at degrees from each school as equal. In reality each student's education is a result of the effort and work they put in for each class and what they learned from the course. Looking at this dataset, there are patterns that emerge to show the level of education each student is receiving.

**Analysis and Models**

**About the Data:**

This dataset consists of responses from students at five schools (A, B, C, D and E) that are implementing the same math course this semester, with 35 lessons. There are 30 sections of students split between the five schools. This data was recorded when the students were about three quarters of the way through the 35 lesson semester.

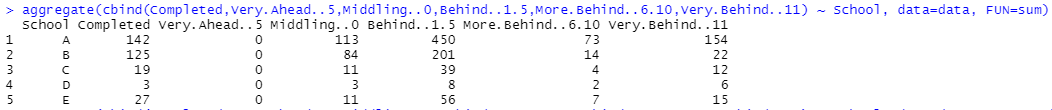
For each section, students answered if they were: very ahead (more than 5 lessons ahead), middling (5 lessons ahead to 0 lessons ahead), behind (1 to 5 lessons behind), more behind (6 to 10 lessons behind), very behind (more than 10 lessons behind), or completed (finished with the course). There were no missing values, and no clear outliers, however there were no students that fell into the “very ahead” category being more than five lessons ahead.

Here is a scatter plot of the number of students in each category 2 - very ahead, 3 - middling, 4 - behind, 5 - more behind, 6 - very behind, 7 - completed, 8 - total students.

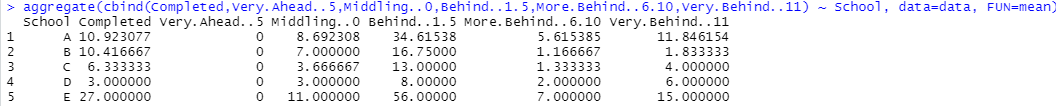
This shows that there are a majority of students in the “Behind” category being 1-5 lessons behind schedule, this is a worrying trend to see across all schools.

This means that you write about each variable, **visualize** each variable (as feasible), and talk about what the variable represents. Tables are great for this as well.

**Analysis of the model:**



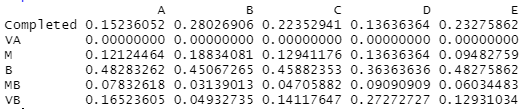
This table shows the breakdown of the different schools and what categories their students fall into. We can see the size of each school and then look at the education from each school from the percentage of students who are middling or completed the course.



Using the mean number of students for each category split by schools shows the breakdown of students in each section.

**Results**

Here is the percentage breakdown of students in each category split by schools A through E.



This shows that although school A has a higher number of student that have completed the course schools B and E have a much higher percentage of their students completing the course.

Schools A, C, and E have a high percentage of students who are “Very Behind” but school D has a much higher percentage than any other in this category indicating that their methods might not be the best. All 5 schools have a good amount of their students in the behind category which may be an indicator of how all schools are teaching their classes.

**Conclusions**

Based on these results I would send my students to School B as they have the highest percentage of students who have completed the course and the lowest percentage of students in the “More Behind” and “Very Behind” categories. A on paper looks like it has a lot of students completing the course but just has a lot more students in general than other schools.

This could be a difference in teaching style if some schools are teaching with live instruction prior to at home work and other schools teaching having students complete individual work before class time.

There is probably some sort of error in collecting the data as there are no students in the “Very Ahead” category. This could also be because of where they are in the semester. Any student who is going to work ahead will have already finished all of the material.