Ancient History Unit Enrolment Predictions

Practical and Impractical Predictions

Thomas Demkowski
Macquarie University
thomas.demkowski@students.mq.edu.au — 44610483

Abstract

I focused on predicting student enrolment numbers for all Ancient History units from first to second year, and from second to third year. I explain why most of the units are too impractical (as far as my skill goes) to predict I predicted the numbers for 10 units where I could make a prediction (for 2020 under pre-2020 curriculum). I then compare the pre-2020 curriculum with the new 2020 curriculum with the data I have to compare predictability and similarities in course structure.

Introduction

Macquarie University's data on student enrolments is not particularly useful for producing Business Intelligence insights. In order to improve the quality of its content delivery, we wish to clean this data and create automatic analysis tools in Python applying appropriate statistical techniques to predicting future unit loads and identifying issues with planned staffing levels.
With this context in mind, my analysis focuses on undergraduate student enrolments in the Ancient History department without the need to fully automate any of the university's data. The majority of my analysis (especially when quantifiable) is heavily based upon Dr Brian Ballsun-Stanton's own data cleaning and data automation of the university's data, where I use Python to create tools that access information of interest from Brian's cleaned data.

Objectives

1. For where it is practical, predict higher level (100 to 200, 200 to 300) unit enrolments of the Ancient History department from historical data and the MQ handbook.
2. Explain why (1) is not practical for the majority of Ancient History units under Macquarie University's current historical database
3. Discuss the differences between the new 2020 curriculum with the current regarding predictability

Materials and Methods

Materials used are Dr Brian Ballsun-Stanton's automated data collection and analysis of the university's database, Python for my data extracts from this collection, and Graphviz for my undergraduate visualization.

Three prediction approaches are used. The first studies the *flow* of the Ancient History course structure to predict the Ancient History student distribution should behave throughout first year, second year and third year. The second and third method predict student numbers based on prerequisite-associated units from 2019, where the third is an alternative to the second wherever there is a lack of data.

Mathematical Analysis

We first consider students who are following the Bachelor of Ancient History course structure as of 2019. From the course structure, we can find out what is the probability of any such student enrolling in any of the undergraduate units. If a unit is mandatory, then we suppose there is a 100% chance of the student enrolling or being enrolled in this unit. If a student must choose N out of M units, then there is a N/M probability between each of these M units that the student will enrol in one of them.

Under the current course structure, the 100 level contains one 100% chance unit, five 40% chance units, and six 33% chance units. The 200 level contains seven 30% chance units and twenty 15% chance units. In the 300 level, there is one 100% chance unit, two 50% chance units, and forty-three 7% chance units. From this we can see that there are 35% more 200 level units than 100 level units, and 70% more 300 level units than 200 level units. We also see that it becomes harder to predict student enrolment numbers in the 200 level compared to 100 level, and ditto for the 300 level compared to the 200 level.

Furthermore, suppose, hypothetically, that there are exactly 1000 first year students following the Bachelor of Ancient History course structure (it is realistically between 100 and 200 judging by enrolment numbers in the mandatory AHIS190 unit, however an average of 1500 students will enrol in AHIS units). Let us observe how the unit chances above divide these 1000 "students".

$$\text{100 level:}$$
$$1000 \overset{100\%}{\Longrightarrow} 1000 \overset{40\%}{\Longrightarrow} 400 \overset{33\%}{\Longrightarrow} 132$$
$$\text{200 level:}$$
$$132 \overset{30\%}{\Longrightarrow} 39.6 \overset{15\%}{\Longrightarrow} 5.94$$
$$\text{300 level:}$$
$$5.94 \overset{100\%}{\Longrightarrow} 5.94 \overset{50\%}{\Longrightarrow} 2.97 \overset{7\%}{\Longrightarrow} 0.2029$$

This shows that there is far too much freedom for students under this course structure to predict student enrolment numbers per unit as the students progress to the successive level. In the 100 level, there is a small enough degree of freedom to predict first year student numbers, however we are interested in the transformation from 100 to 200 level, and 200 to 300 level. Unfortunately, the degree of freedom at 200 level and 300 level under this course structure is too unpredictable with this method. We move on to the next method in attempt to complete our first objective.

The following method is independent on students following the Ancient History course structure. We attempt to predict Ancient History units using the prerequisite. First we note that under the Python data analysis, there is only one Ancient History unit with co-requisite requirements, namely, AHIS389, which only states the need for 6cp in AHIS or AHST units at 300 level; therefore we ignore this exception. We also ignore NCCW cases, as this was not included in Dr Brian Ballsun-Stanton's data analysis of the MQ handbook. It should be noted that both of these methods do not consider the possibilities of students entering a class without the need for meeting the prerequisite unit requirements, which can be possible via special approval. Our method assumes that this is negligible.
As seen in Figure 1, we see that there are 21 prerequisite-associated units. For every unit with a prerequisite unit, we may predict the number of students that will be enrolled either of two ways. The first is more accurate, relying on the average passing grade rate (from years 2016 to 2019) of the prerequisite units as a multiplier of the number of students enrolled in the unit in 2019. (check real quick if I can do that from 2016 to 2019) The other way is to less accurate and will directly give the sum of 80% (Pareto's principle) student enrolment numbers for every prerequisite unit of that unit. The latter is to be used whenever we lack the data on completion rates. Now, out of all of the prerequisite-associated units, there are only 10 Ancient History units that *have* prerequisite units. The predictions for these will be given by Table 1 along with the calculation method.

2020 curriculum comparison One big distinction between the upcoming course structure and the current is that the new introduces specialized courses, where Ancient History students have a sub-course structure behind their course (it is possible that I am mistaken here, although have not encountered evidence for this). Another large distinction is that the unit codes will change and there are likely added, removed or modified units; similar to what is happening to the mathematical department. Due to lack of data on the new curriculum, we shall only made some remarks on the subject in the results.

Results

We have evaluated three different methods for solving our first objective. The first was much more general, designed to show the *flow* of the course structure as first year students progress throughout the course. Here, we found that there is too much freedom of flow between the units where it rapidly divides students into different classes of their choice; so much that even for our over-assumed 1000 student count case.
The lines between each unit represents a prerequisite unit link. Every unit that is a prerequisite or has a prerequisite is highlighted in yellow. This representation shows the lack of mandatory structure enforced onto the undergraduate units.

AHIS370's prediction is likely unreliable under this model.

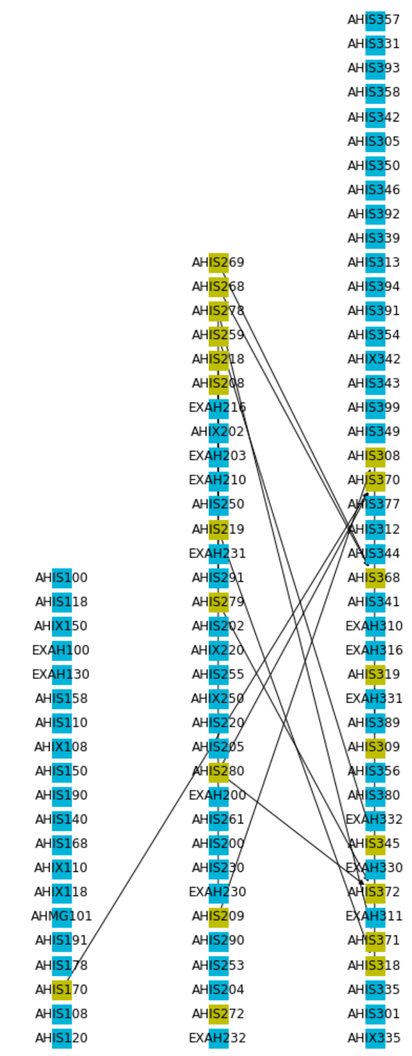

Figure 1: All Ancient History undergraduate units in 2019

Below we see our student enrolment predictions for 2020:

Unit	2019	2020 prediction
AHIS372	7	37
AHIS279	17	12
AHIS318	11	19
AHIS370	6	84
AHIS345	5	40
AHIS308	10	16
AHIS209	33	36
AHIS368	0	33
AHIS319	3	15
AHIS219	29	5 (2nd approach)
AHIS309	2	5
AHIS371	7	7

Table 1: Prediction of calculable units

As we can see, most of the predictions are intuitive. However, codes 370, 345, 209 and 368 are suspiciously counter-intuitive. This is most likely the result of most students from the prerequisite units not obeying the Ancient History course structure. Every one of these listed 300 level units that were possible to predict unfortunately are units of *high freedom of choice*; that is, under our *flow* distribution, students following the course structure have a 7% chance of entering them.

Conclusions

There are not enough prerequisite-linked units in the Ancient History department to reliably predict future student enrolment numbers for most units. For the units that can be predicted, some raise suspicion of inaccuracy due to students having too much freedom for our numbers to be reliable. As for the others, there should be an increase in student numbers per class, except for AHIS219, which should have a strong decrease due to a lack of students who are eligible for it.

The new 2020 curriculum may distort our predictions since all of these units in the table will have new unit codes, and some may be removed or modified. New units may possible to add to the table in the future. However, the new specialized sub-course structures should add much greater practicality to the *flow* analysis approach for future prediction models.

Acknowledgements

This work is heavily based upon Dr Brian Ballsun-Stanton's large data collection and analysis that he built from Macquarie's database. None of this mathematical analysis would have been possible without his work.