How to design and implement a cohort study of student progression rates

Tony Nolan OAM, Information and Intelligence Analyst, Faculty of Business, University of Technology, Sydney

Abstract

Students don't always follow the desired progression pattern. With the range of choices that students now have, they can accelerate or decelerate either by their own design or by act of God. Universities also have capped limits or quotas, so that different types of modelling are needed to truly understand the flexibilities and sensitivities in student progressions.

This session will explore various ways to track student progressions. The complexities of having different data models, the different ways to track students, alternative mathematics used in finite modelling, bridging the old versus new changing definitions and labels, database construction, hyperpanometrics, forecasting and how to solve the defining cohort clustering problem.

Introduction.

The problem facing many educational institutions is the greater requirements for data capacity and then various systemic requirements of modern software and the user's desire for functional information for logistical purposes. The educational environment is very different to a standard market place. It is shaped by finite limits on such variables as intake, progression, sales, human logistics, competitive undertakings, possible outputs, and prices. We do not practise in an infinite environment of unlimited resources or sales capacity. Our client base is also practises a non-standard progression, and a degree of variety in choice of different activities, in service requirements.

Under current commercial programs for sales models and data mining, they are based in the majority of market perceptions, of infinite possibilities. Many also fail to take into account the mixed throughput and cycles of having such a wide range of product and service choice. A student academic community profile for a university year can be made up of many students from different funding groups, different progression and pass rates, different study modes, and at different stages of their studies. It really is a giant mixing bowl of ingredients, from which we have to make the perfect set of pies for the funding model, etc.

There is however a different approach to this problem, by in effect changing the rules and attacking the data from a different methodology. By the recognition of the need for a finite model, and by exploring the possibilities of finite modelling, and using fuzzy logic principles to map movements and sensitivities to logistical changes. It becomes possible to have a series of different types of analysis to give a better picture of the complexities of student profiling. The follow is a mini example taken from an cohort profiling activity undertaken as a pilot study at the Faculty of business, UTS.

Problem:

Current Faculty of Business customers are coming from a market place that expects a wide degree of flexibility in their academic activities. This creates a logistical problem and a paradigm shift because of the newer approach required for program diversity in respect to both academic coverage of multiple disciplines and delivery mode of faculty product. This is further complicated by the increasing degree of competition in the marketplace as well as a reducing budget environment.

Under previous types of modelling of student progression and student choice, activities such as strategic planning, logistical assignment, financial planning, optimal course planning, human resource planning, financial allocations, and just-in-time principles have had their effectiveness restricted because of the quality of data and the lack of data mining techniques and modelling approaches available.

A final compounding factor has been, academic institutions haven't perceived the need to conduct such activities, or have adopted models with limited suitability for our academic environment. Basically we have been borrowing what we have needed in a reactive response, rather than adopting a proactive response towards dedicated modelling activities.

Approach:

Previous modelling approaches have combined all students in a melting pot, and then carved them up into segments of the academic community pie. Segmentation has been within a discipline, mode or course designation. There has also been a philosophical approach to stay focused on the MACRO data levels, which is useful for a current snapshot of the university community, but lacks the coherent specificity for logistical / planning activities. There is also a need to have a data, academic progressions based on optimum pathway suggestions in the Faculty Handbook compared to actual student progression pattens.

The new methodology uses student clustering and focused stratification analysis on standard and enhanced data sets to achieve unique cluster identifiers. Further mapping and difference modelling will allow for better strategic planning and logistical allocation.

Model:

By creating an Cohort model, through a series of new student labels (Appendix 1), and a purpose designed series of filters and chronological reference points, it is possible to profile the study and progression pattens of our customers, and to be able to assess the diversity of alternatives with deviates from the assumed study program within the system.

The Cohort model analysis is based on a newly developed type of mathematics and modelling techniques, that gives a wide range of flexibility in data modelling, not previously available.

Marketing Applications

The delivery of student pipeline profile data could have several applications to the way in which the Faculty promotes our product line. It offers several avenues through which we might improve our cost effectiveness and more accurately tap into our prospective student audience.

Pilot Study:

A data model was developed and pilot study has already been undertaken, within the Faculty of Business, using previous student data, which was adapted into the desired format. A series of outcomes and measurement techniques from the pilot study showed some interesting outcomes, which have already sparked further questions and some interest in expanding the pilot study.

Some Sample Outputs:

Headcount by Bands in Semesters. Individual units by progression status over time. (Appendix 2, 5 & 8)

Individual Student Progression: A new style of report that includes semester loads, pass rates, mode / funding changes, core vs elective ratios. (Appendix 3.)

Chronological Core Unit Analysis. Examines the time a students chooses to study there core units. (Appendix 4.)

Cohort Progression: A time series of student study loads, which identifies pattens of progression, with accretion and band progression. (Appendix 6 & 7.)

Future Developments:

Decision Matrix / Dendrogram. A series of statistical observations, which produces a decision tree on the subject choices of student electives.

4thought Analysis: Using an unique multivariable identification, undertake an data mining activity, which produces a series of variable impacts on projected outcomes.

Hyperpanometric: A multi dimensional approach to factoring all influential variables and modelling

Classroom Profile: Gives an indication of the classroom make up of progression bands, and various accretion possibilities, bleeding and gain of cross faculty funding, etc

Accretion Triggers: A study of what effect summer school and winter school have on accreting student progression.

Appendix 2.

A sample data sheet for modelling

Course Band label

Semester labels

		55	56	57	58	59	60
B01C-0		23	35	45	45	23	24
B01C-1		25	25	36	25	21	23
B01C-2		24	24	21	31	21	24
B01C-3		25	26	21	15	23	26
B01C-4		26	25	45	42	52	35
B01C-5		21	24	14	25	23	21
B01C-6		25	26	25	32	14	14
B01C-7		35	28	12	14	25	29
B01C-8		26	29	45	36	34	28
B01C-9		24	24	35	45	23	25
B01C-10		52	25	26	32	12	32

Flow of bands between time frames.

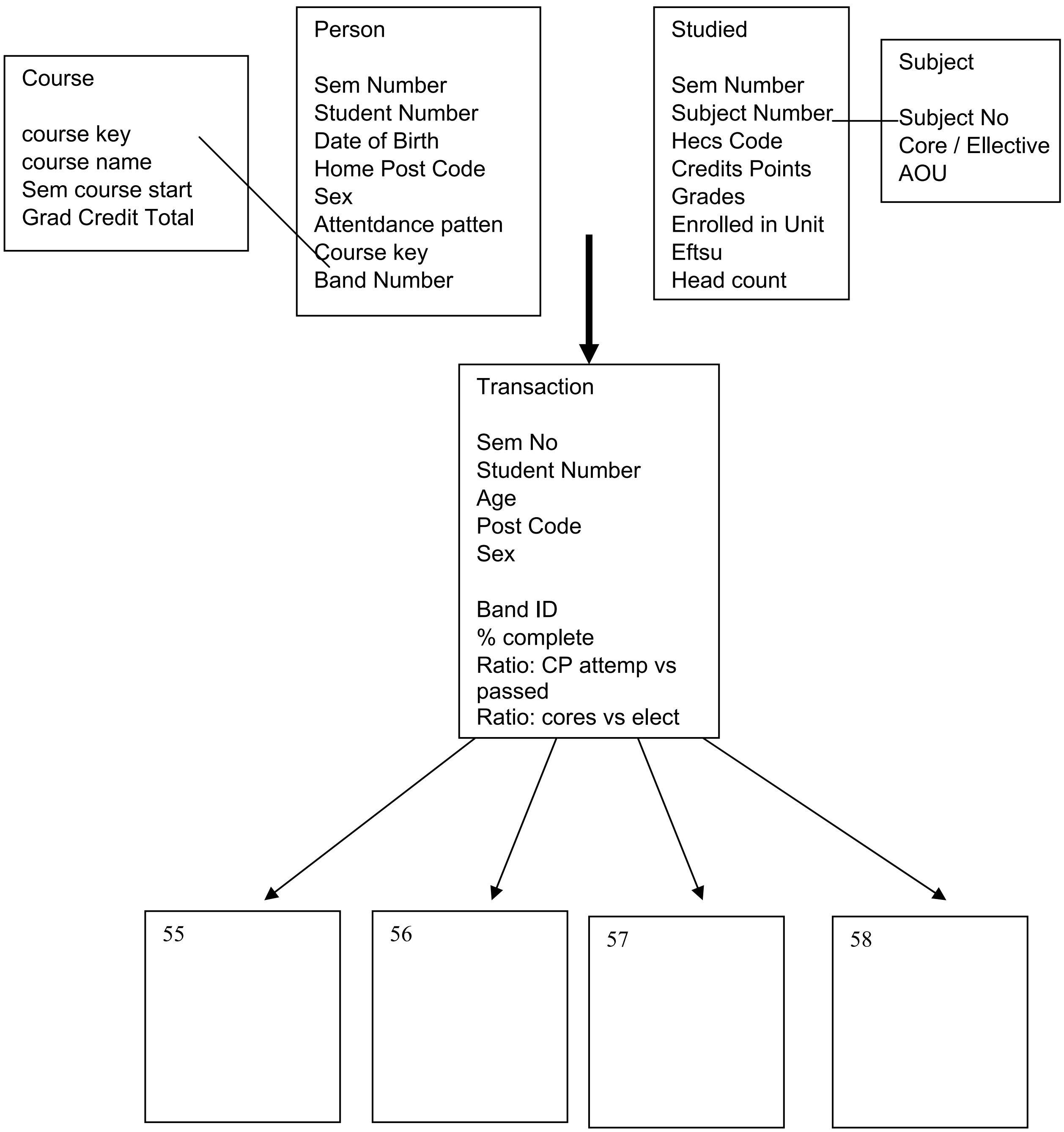
Examine the progression of different bands through time, and check progression rates

Total headcount for each semester

Comment:

This piece of analysis is designed to give an indication of the degree of movement between the bands and the semesters. Although this example is made up of sample data, actual pilot testing demonstrated that once a cohort progresses to 20% completion, ½ the cohort actually jumps to the next percentage band. This is due for further investigation at a later stage.

Appendix 1
Sample Database Structure



Appendix 3.
Student Progression Profile Example

Band	CP Sem accum total	Sum Of Credit Points	C	O
25-B01C-0	9.012	9.012	16	5
27-B01C-1	21.02	12.008	20	
29-B01C-2	29.028	8.008	16	
31-B01C-2	29.046	0.018	12	6
33-B01C-3	47.05	18.004	16	6
35-B01C-3	57.056	10.006	4	12
37-B01C-5	78.062	21.006	22	5
39-B01C-6	89.08	11.018	12	17
40-B01C-6	95.08	6		6
41-B01C-7	107.08	12	6	6
44-B01C-8	117.08	10	10	

Core Units studied in credit points.

Other Units in credit points.

Sum of credit points for that semester load after results.

- Pre decimal point = CP failed / 1000
- Post decimal point = CP passed.

Credit point accumulated total up till each semester - Total passed: Total failed

Relevant band of study.

- New Semester Reference
- Course Code
- Band Identifier.

Comment:

This shows an example of how by using cohort and course codes can be used to cluster credit point data as individual students progression. This allows an examination of load of cores and electives, and also a measure of pass / fail rates.

Appendix 4.

Chronological Core Unit Analysis.

Subject Number.

Head Count

Semester Label.

Progression Band Categories.

sub	sem	0	1	2	3	4	5	6	7	8	9	10
21125	35	11	31	144	12	7	3	1				
21125	37	62	242	34	9	7	5	2				1
21125	39	41	82	106	14	8	4	2	2	1		
21125	41	91	163	49	24	11	3		1			
21125	43	65	95	92	89	4	10	3	3			
21125	45	149	221	58	26	10	8	3	2			
21130	35	13	24	125	6		2					
21130	37	52	249	19	8	6	1	2				
21130	39	33	58	128	1	5	2	1	1			
21130	41	33	161	22	10	4	1					
21130	43	21	39	40	23	1						
21130	45	69	137	36	20	14	5	4				
25314	35	10	15	29	42	44	37	7	4	1		
25314	37	26	37	32	80	144	22	21	10	5	2	
25314	39	19	37	33	40	49	42	11	8	8		1
25314	41	35	64	47	104	153	31	21	11	12	7	4
25314	43	40	137	128	138	47	57	30	12	8	3	1
25314	45	24	76	51	59	40	27	11	6	3	7	1

Comment:

These shows when the students choose to time table their core units in their progression. This piece of analysis helps us to decipher the students reasons for choosing to study their core units, when they aren't required for their major or minor streams.

Appendix 5.

Head count of 21125 in bands by semester

Subject Number

Course Code + Band

Semester Label

sub	cc-band	25	27	29	31	33	35	37	39	41	43	45
21125	B01C-0	84	11	68	16	81	11	62	41	91	65	149
21125	B01C-1	251	52	210	22	169	31	242	82	163	95	221
21125	B01C-2	1	168	13	145	14	144	34	106	49	92	58
21125	B01C-3		7	9	2	6	12	9	14	24	89	26
21125	B01C-4			3	3	3	7	7	8	11	4	10
21125	B01C-5			1	5	6	3	5	4	3	10	8
21125	B01C-6				1	2	1	2	2		3	3
21125	B01C-7								2	1	3	2
21125	B01C-8								1			
21125	B01C-10							1				

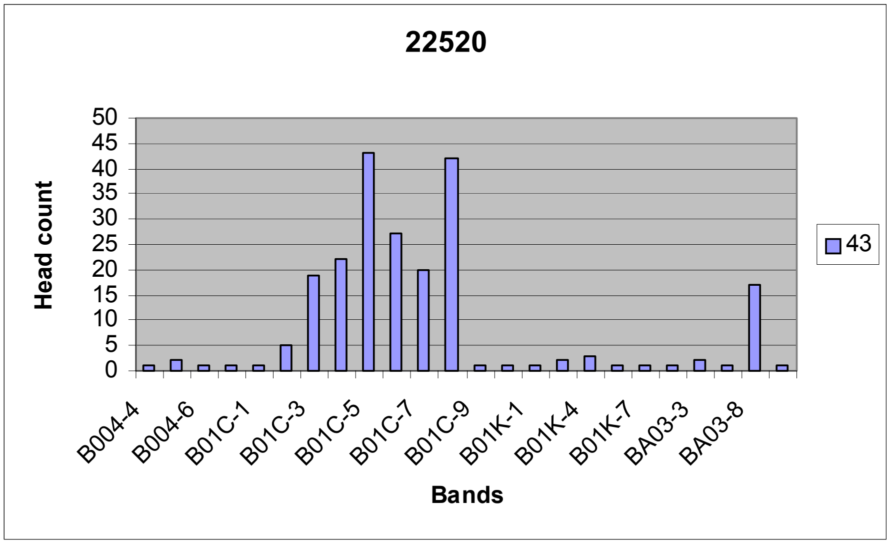
Head count of student's enrolled

Comment:

This shows the subject number 21125, and the semester that students choose to study it, but it is refined by how far the student has progressed in their studies, before they undertake the unit.

Appendix 6.

Subject Cohort Profile.

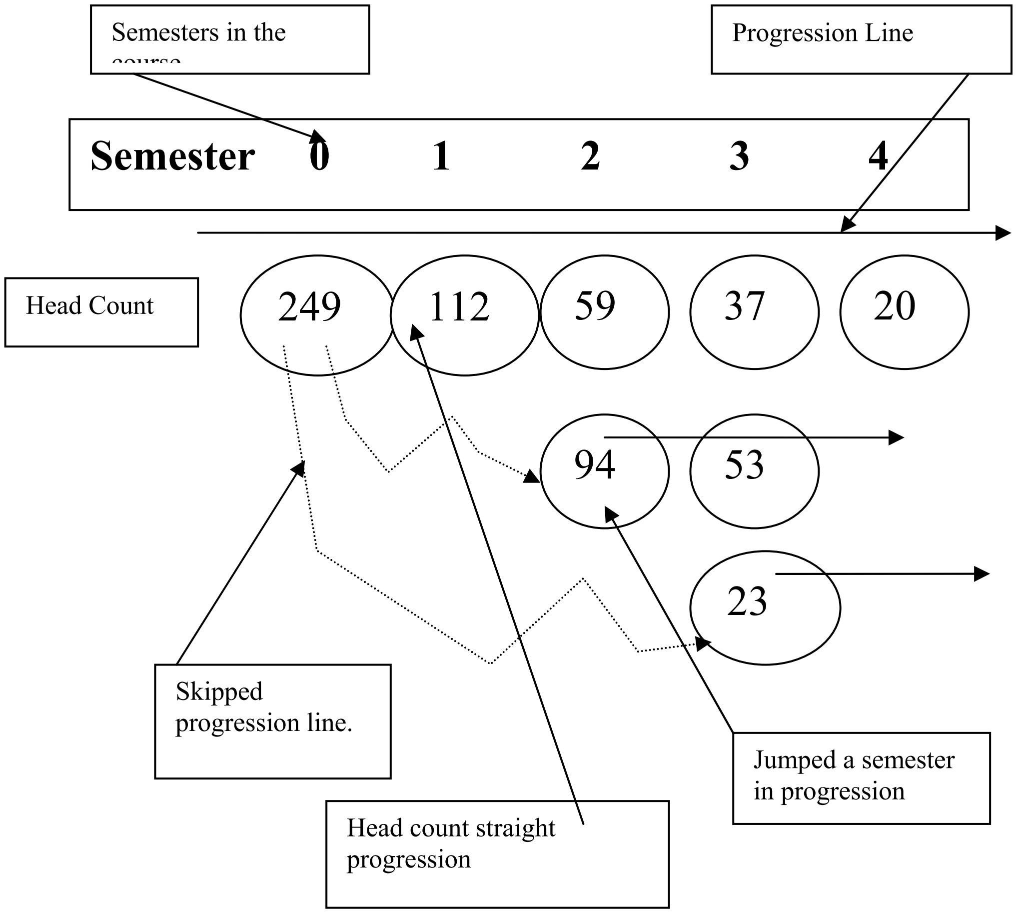


Comment:

This piece of analysis is designed to give an indication of the variety of students from different cohorts who make up a classroom or student profile. This is of significance interest to see how many students are taking the subject or class from different courses or faculities.

Appendix 7.

Cohort progressions, straight and skipping.



Comment:

This diagram shows the break up of a specific Cohort, and the progression rates through the bands. How many students have a standard progression rate, and how many students take an accelerated path.

Appendix 8.

Study pattens in bands.

course	CountOf Band	B0	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10
B056	104		X									
B056	61			X								
B056	50	X										
B056	28			X	X							
B056	26		X	X								
B056	22	X	X									
B056	21		X	X	X							
B056	20		X		X							
B056	14				X							
B056	13		X	X		X						
B056	13	X		X								
B056	11											
B056						X						

Various band structures.

Count of students in that band patten.

Course Code

Comment:

This diagram shows the any pattens of students skipping progression bands, and searches for any pattens. This would then be flitted for winter or summer school effects.