

Mapping Course Selection Pathways: A Study on LSE Undergraduate Trends

Andy J. Leo, Benjamin B. Steiner, Jonathan Cardoso-Silva
LSE Data Science Institute

Background

- Investigate student course selection at LSE using anonymised enrolment data.
- Data starting from 2016, provided by the LSE Planning Division.
- Insightful for curriculum planning and enhancing understanding of student goals.

Aim

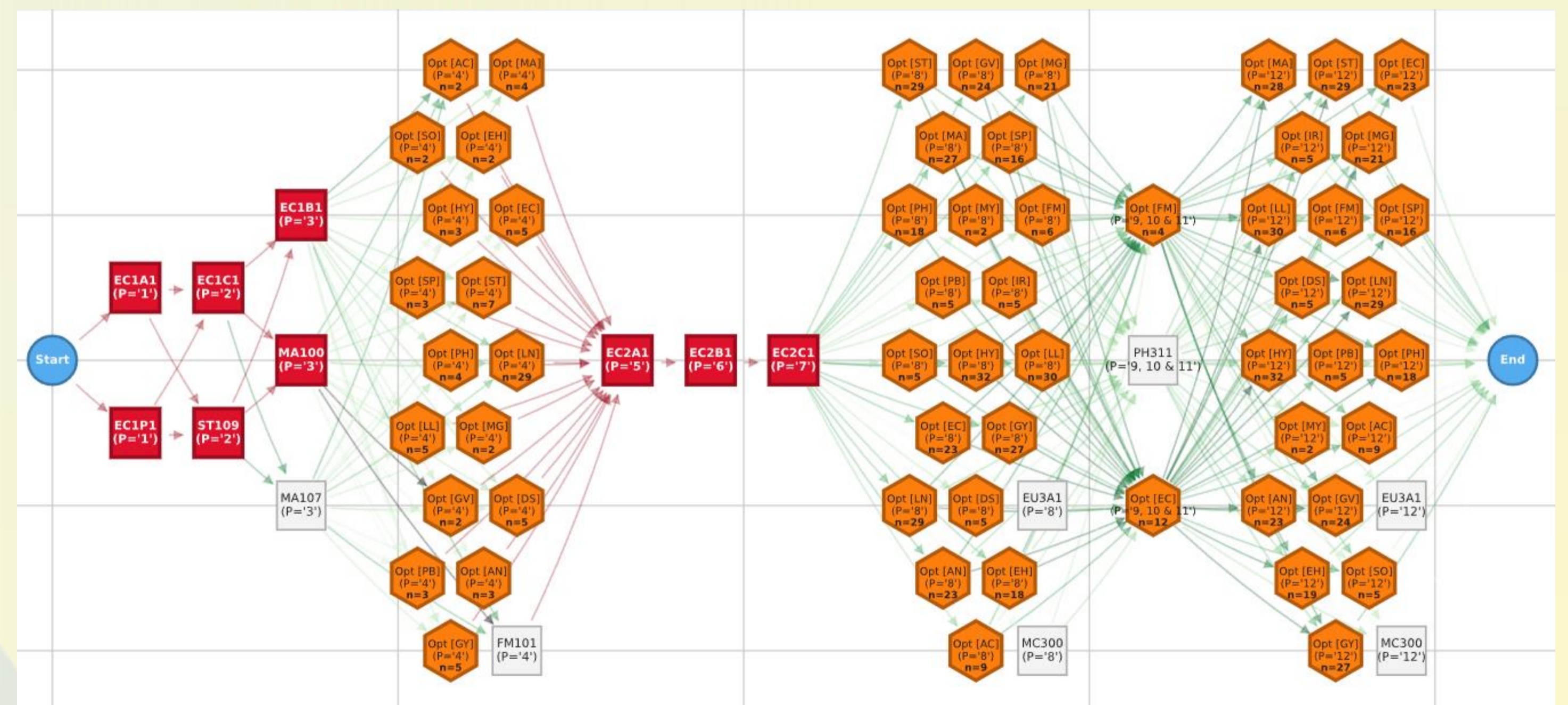
- Using data visualisation and similarity measures like the Jaccard index, analyse course selection trends across departments.
- Try to observe commonly chosen outside options and prevalent selection pathways.

Methodology

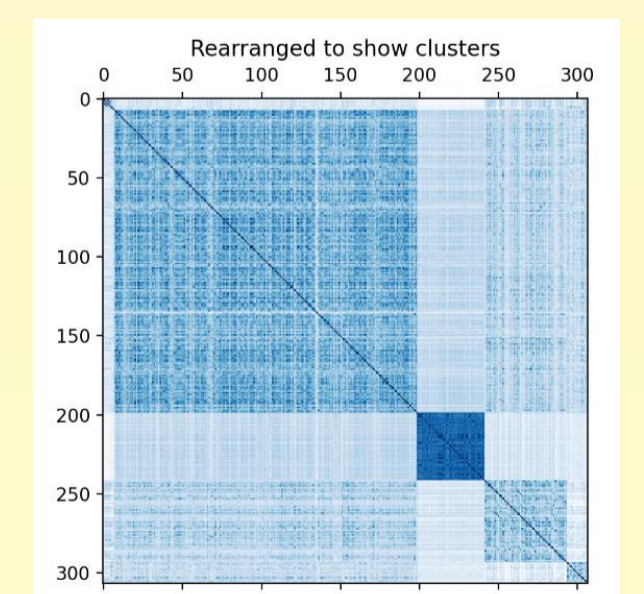
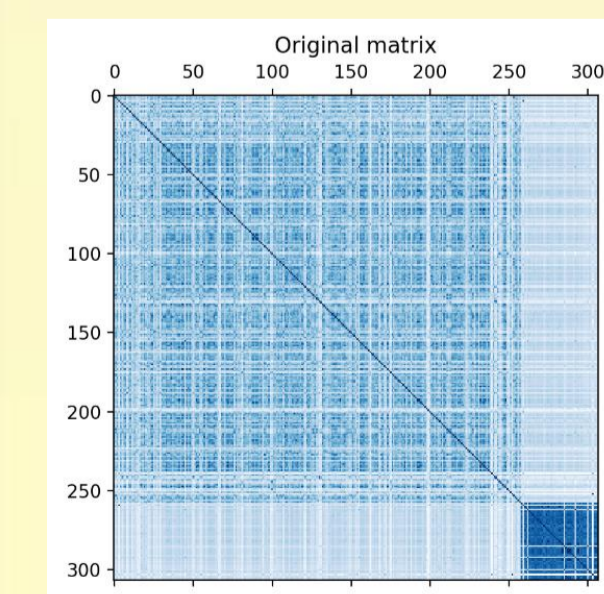
- Original data in two .xlsx files, combined and converted to .db for efficiency.
- Data cleaning and processing performed in Python using Pandas.
- Visualise possible pathways (using NetworkX) for degree programmes.
- Student course selection for programme pairs compared using the Jaccard similarity coefficient, a value between 0 and 1 (higher value – greater similarity).
- Create symmetric coefficient matrices in order to perform spectral clustering analysis.
- Spectral clustering used to cluster students with similar course selection.

Results

- Graphed pathways for programmes with similarities (like Economics/Finance).
- Cluster plots shed light on common course pathways and interdisciplinary interest, potentially signalling the organic formation of emergent degree programmes.



	EFRO00HO	EFRO00QQ	EFRO00FX	EFRO00QH	EFRO00XE	EFRO00WU	EFRO00HP	EFRO00QE	EFRO00FQ
EFRO00HO	1.00	0.44	0.59	0.22	0.50	0.24	0.50	0.73	0.23
EFRO00QQ	0.44	1.00	0.35	0.38	0.35	0.40	0.29	0.37	0.35
EFRO00FX	0.59	0.35	1.00	0.15	0.56	0.22	0.56	0.69	0.17
EFRO00QH	0.22	0.38	0.15	1.00	0.10	0.21	0.10	0.16	0.15
EFRO00XE	0.50	0.35	0.56	0.10	1.00	0.22	0.56	0.50	0.17
...
EFWXPHQQ	0.24	0.15	0.23	0.04	0.23	0.08	0.23	0.24	0.07
EFWXQH0H	0.11	0.24	0.10	0.31	0.16	0.23	0.10	0.05	0.16
EFWXQ0HE	0.24	0.15	0.23	0.04	0.23	0.08	0.23	0.24	0.07
EFWXQ0UH	0.24	0.15	0.23	0.08	0.23	0.08	0.23	0.24	0.07
EFWXQ0UH	0.30	0.20	0.29	0.04	0.29	0.14	0.29	0.30	0.11



Example course selection similarity coefficient matrix, programme pathway, and spectral clustering plots shown above.