Research design and data analysis with big data

Giulio Marini

Contents

The Module introduces main potentialities and hurdles of using (publicly available) big data. Goals are to: 1) envisage a possible research question without resorting to primary data; 2) develop one's dataset out of (multiple) sources in such a way that researcher will approximate a quasi-experimental research design; 3) delivering respective data analysis. Students at first year will have to write a short 1500 words essay discussing an original literature review they will perform about their own topic of PhD (e.g., occurrences of keywords and topics out of searches within Scopus).

Course schedule

Approaching Big Data for Academic Research: potentialities along with first hints of main dilemmas	Drovandi, C. C., Holmes, C., McGree, J. M., Mengersen, K., Richardson, S., & Ryan, E. G. (2017). Principles of experimental design for big data analysis. *Statistical science: a review journal of the Institute of Mathematical Statistics*, *32*(3), 385. Plus original powerpoint
Working around constraints of public repositories. The case of HESA data powered by Tableau	Marini, G. (2024). Brexit and the war for talents: Push & pull evidence about competitiveness. *Higher Education*, 1-16. Plus original powerpoint; dataset & syntax
Merging 3-digits information into a microdata dataset to add critical variables. The case of UK Labour Force Survey	Marini, G., & Henseke, G. (2023). Is a PhD worth more than a Master's in the UK labour market? The role of specialisation and managerial position. *Studies in Higher Education*, *48*(10), 1538-1550. Plus original powerpoint; dataset & syntax
Merging multiple public sources into a dataset to answer a specific research question. The case of promotions in institutions.	Meschitti, V., & Marini, G. (2023). The balance between status quo and change when minorities try to access top ranks: a tale about women achieving professorship. *Gender in Management: An International Journal*, *38*(1), 17-35. Marini, G., & Meschitti, V. (2018). The trench warfare of gender discrimination: evidence from academic promotions to full professor in Italy. *Scientometrics*, *115*(2), 989-1006. Marini, G. (2017). New promotion patterns in Italian universities: Less seniority and more productivity? Data from ASN. *Higher Education*, *73*, 189-205. Marini, G. (2014). Italy's New Requirements for Academic Careers: The New Habilitation and its Worthiness. Plus original powerpoint; dataset & syntax
How to use bibliometric big data to build up a good literature review (about Students' real topic of their PhD).	Some Essays from previous year, plus examples Plus original powerpoint; dataset & syntax. Students belonging to first year also expose their own preliminary ideas for the submission

Bimester: October-November