

Transparent and Reproducible Social Data Science – A reflection on tools and techniques suitable for lifecourse and equality research

Vernon Gayle, University of Edinburgh

Across a range of social science disciplines there is increasing interest in research transparency and the reproducibility of empirical results. The theme of this presentation is undertaking transparent and reproducible research using statistically orientated methods. It will outline the concepts associated with undertaking transparent and reproducible research. A focus of the paper will be the 'elements that form 'the workflow', and the practical activities that are routinely undertaken within research projects. We will introduce some tools and protocols from neighbouring fields such as computing and e-research, that have proved effective in undertaking transparent sociological research. Against the backdrop of our experiences in undertaking reproducible lifecourse and equality research, we posit a short set of practicable guidelines to help researchers undertake transparent and reproducible statistically orientated social data science.

Keywords: Transparent; Reproducible; Duplication; Replication; Lifecourse; Equality; Workflow; Jupyter notebooks.