
Plan Overview

A Data Management Plan created using DMPonline

Title: Improving ethnic minority polling

Creator: Nicole Martin

Principal Investigator: Nicole Martin

Data Manager: Nicole Martin

Affiliation: University of Manchester

Template: ESRC Template Customised By: University of Manchester

ID: 58533

Last modified: 09-06-2020

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

Improving ethnic minority polling

Manchester Data Management Outline

1. Will this project be reviewed by any of the following bodies (please select all that apply)?

- Funder

2. Is The University of Manchester collaborating with other institutions on this project?

- No - only institution involved

3. What data will you use in this project (please select all that apply)?

- Acquire new data
- Re-use existing data (please list below)

4. Where will the data be stored and backed-up during the project lifetime?

- University of Manchester Research Data Storage Service (Isilon)

5. If you will be using Research Data Storage, how much storage will you require?

- < 1 TB

6. Are you going to be working with a 3rd party data provider?

- Yes

YouGov will collect data from an online survey.

7. How long do you intend to keep your data for after the end of your project (in years)?

- 0-4 years

Questions about personal information

Personal information, also known as personal data, relates to identifiable living individuals. Special category personal data is more sensitive information such as medical records, ethnic background, religious beliefs, political opinions, sexual orientation and criminal convictions or offences information. If you are not using personal data then you can skip the rest of this section.

Please note that in line with [data protection law](#) (the General Data Protection Regulation and Data Protection Act 2018), personal information should only be stored in an identifiable form for as long as is necessary for the project; it

should be pseudonymised (partially de-identified) and/or anonymised (completely de-identified) as soon as practically possible. You must obtain the appropriate [ethical approval](#) in order to use identifiable personal data.

8. What type of personal information will you be processing (please select all that apply)?

- Special categories and criminal convictions
- Anonymised personal data

9. Please briefly outline how you plan to store, protect and ensure confidentiality of the participants' information.

All data will be fully anonymised at all times. The PI will not have access to any personal identifying information.

10. If you are storing personal information (including contact details) will you need to keep it beyond the end of the project?

- Not applicable

11. Will the participants' information (personal and/or sensitive) be shared with or accessed by anyone outside of the University of Manchester?

- No

12. If you will be sharing personal information outside of the University of Manchester will the individual or organisation you are sharing with be outside the EEA?

- Not applicable

13. Are you planning to use the personal information for future purposes such as research?

- No

14. Who will act as the data custodian for this study, and so be responsible for the information involved?

Nicole Martin

15. Please provide the date on which this plan was last reviewed (dd/mm/yyyy).

10-06-2020

Assessment of existing data

Provide an explanation of the existing data sources that will be used by the research project, with references

This project will make use of multiple datasets. Some already exist such as the UK household panel survey, *Understanding Society*, election results and data published in Census tables. Other data will be opinion polling collected or collated for this project from the polling company YouGov.

Provide an analysis of the gaps identified between the currently available and required data for the research

The novel data collection from YouGov is needed in order to achieve the research aims of the project - (i) to assess the sampling bias in online samples of ethnic minorities in Britain, and (ii) to have adequate sample size and data specific to recent or future electoral events.

Information on new data

Provide information on the data that will be produced or accessed by the research project

Understanding Society: the data comes from surveys administered to households and individuals. It has geographical identifiers, and key variables of interest to the study. It is structured in individual- and household-level datafiles.

Census tables: Census data is published on the distribution of different ethnic groups across different electoral units (constituencies, local authorities, wards). It can be linked to other data through standardised identifiers.

Polling data collected by YouGov: this data will contain poll responses and respondent characteristics, including geographical identifiers which will permit the statistical methods used in this project.

Quality assurance of data

Describe the procedures for quality assurance that will be carried out on the data collected at the time of data collection, data entry, digitisation and data checking.

The survey company YouGov have their own internal quality control procedures which will ensure data quality.

Backup and security of data

Please describe the data back-up procedures that you will adopt to ensure the data and metadata are securely stored during the lifetime of the project.

Data will be stored using the University of Manchester Research Data Storage facility which has back up procedures in place.

Management and curation of data

Outline your plans for preparing, organising and documenting data.

At the end of the project, replication datasets will be prepared with documentation following best practice as identified by the UK Data Service.

Difficulties in data sharing and measures to overcome these

If you expect obstacles to sharing your data, explain which and the possible measures you can apply to overcome these.

License conditions prohibiting data sharing may be an issue for *Understanding Society* data. This can be overcome by using the UK Data Service and creating a replication dataset with the same license conditions as the data used to create the replication dataset. I do not foresee obstacles to sharing the other data used in the project.

Consent, anonymisation and strategies to enable further re-use of data

Make explicit mention of the planned procedures to handle consent for data sharing for data obtained from human participants, and/or how to anonymise data, to make sure that data can be made available and accessible for future scientific research.

All data used this project will be fully anonymised, and has been collected with informed consent of participants that data will be used for reanalysis.

Copyright and intellectual property ownership

Please state who will own the copyright and IPR of any new data that you will generate.

The PI will own copyright and IPR of any new data generated.

Responsibilities

Outline responsibilities for data management within research teams at all partner institutions

The PI as sole investigator will be responsible for all aspects of data management.