Plan Overview

A Data Management Plan created using DMP Tool

Title: Doctoral Dissertation Research: Voting and the behavioral economics of housing in an affordability crisis

Creator: Michael Hankinson

Affiliation: Harvard University (harvard.edu)

Principal Investigator: Michael Hankinson

Data Manager: Michael Hankinson

Funder: National Science Foundation (nsf.gov)

Funding opportunity number: 15-571

Template: NSF-SBE: Social, Behavioral, Economic Sciences

Last modified: 07-08-2024

Copyright information:

Doctoral Dissertation Research: Voting and the behavioral economics of housing in an affordability crisis

Roles and responsibilities

The Data Management Plan should outline the rights and obligations of all parties as to their roles and responsibilities in the management and retention of research data. It must also consider changes to roles and responsibilities that will occur should a principal investigator or co-PI leave the institution.

Co-PI Michael Hankinson will be responsible for all data management, monitoring, and custodianship in perpetuity once the exit poll surveys are collected.

Expected data

The Data Management Plan should describe the types of data, samples, physical collections, software, curriculum materials, and other materials to be produced in the course of the project. It should then describe the expected types of data to be retained.

Existing data used for this project will consist of publicly available US Census data, the San Francisco voter file, and the San Francisco tax assessor's database. Each document is freely available through the City of San Francisco government. Data from the voter file and tax assossor's database will be merged with the new data collected through the exit poll surveys at the individual level. This merged csv file will then have unique identifiers removed before being publicly shared.

Period of data retention

SBE is committed to timely and rapid data distribution. However, it recognizes that types of data can vary widely and that acceptable norms also vary by scientific discipline. It is strongly committed, however, to the underlying principle of timely access, and applicants should address how this will be met in their DMP statement.

Our team will hold access to the data for a two month period, allowing for the data to be merged, cleaned, and de-identified prior to sharing. Afterwards, the data, metadata, and relevant R code will be freely available. There are no expected embargo periods for political/commercial/patent or publisher reasons.

Data format and dissemination

The Data Management Plan should describe data formats, media, and dissemination approaches that will be used to make data and metadata available to others. Policies for public access and sharing should be described, including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements. Research centers and major partnerships with industry or other user

communities must also address how data are to be shared and managed with partners, center members, and other major stakeholders.

The exit sureys will be coded and merged into a csv file for ease of analysis. Metadata of analytical steps and question wording will be shared in pdf format for ease of presentation. R code will be annotated for those seeking to replicate the analysis. All files will be made publicly available within 2 months by posting to the Harvard Dataverse Network. No permission restrictions will be placed on the data. Data will be deindentified to preserve anonymity. There are no foreseeable ethical or privacy issues. The final data management process will be vetted by the Harvard IRB.

Data storage and preservation of access

The Data Management Plan should describe physical and cyber resources and facilities that will be used for the effective preservation and storage of research data. These can include third party facilities and repositories.

Metadata will include the full research plan and exit survey wording for comprehension of responses.The metadata will also include a walkthrough of the annotated R code to facilitate replication of the analysis.

Additional possible data management requirements

More stringent data management requirements may be specified in particular NSF solicitations or result from local policies and best practices at the PI's home institution. Additional requirements will be specified in the program solicitation and award conditions. Principal Investigators to be supported by such programs must discuss how they will meet these additional requirements in their Data Management Plans.

Question not answered.