

SPPQ README Guide

Essential Information:

1. Title and Author(s) of article
2. Compute environment dependencies:
 - a. operating system
 - b. statistical software version
 - c. packages used in analysis
3. List of all files within your verification package in SPPQDataverse and descriptions of each file
4. Must be saved as either a .TXT or .PDF
5. Full data citations and access instructions for all original data sources in the README or Codebook.

README Template

[Article Title]

[Author(s)] – [Affiliation] – [E-mail] – [Twitter (optional)]

Compute Environment:

[Operating system]

[Statistical software versions]

[Estimated run time for all analyses]

[Packages used in analysis]



R PRO-TIP

As part of the base package the command "R.Version" lists the R version and the computer platform.

After a script has been run, the command "(.packages())" will list all of the packages that have been loaded during the session.



STATA PRO-TIP

The 'version' command will tell you which version you are currently using. The command 'dis "c(machine_type)"' will also tell you if you are using a 64 or 32 bit machine type. Please also be sure to include the type of Stata being used (IC, SE, or MP).



STATA PRO-TIP

You can open your .do file on a clean install of Stata that does not have any user packages (such as your University virtual lab, etc.) and run your script to see if there are any errors that pop up due to missing packages.

Alternatively, you can also use the “-ado-” command which will list all user programs that have been installed on your Stata version, however these are not necessarily packages from the list that are actually used in your script.

IMPORTANT



If your analyses were run on a high-performance compute (HPC) environment, please provide the following information:

The script(s) used to schedule the resources (i.e. SLURM or PBS scripts, etc.). These scripts include important information such as the memory needed, the time needed and the number of processes to run (nodes, tasks, etc.).

List of files:

(Note: if there is a folder hierarchy you can list the files in that structure here)

DATA FILES

- Datafile.dta - main analysis dataset constructed from original source data; this data is used to generate the results within the manuscript
- Appendixdatafile.dta - data used to generate results within appendix
- Originalsourcedata.dta - original source data used to compile main analysis data set.
Source: AuthorLastName, AuthorFirstName. Title of Dataset, Year. Name of distributor [distributor], date of distribution as YYYY-MM-DD. DOI or persistent URL
- Originalsourcedata1.dta - original source data used to compile main analysis data set.
Source: AuthorLastName, AuthorFirstName. Title of Dataset, Year. Name of distributor [distributor], date of distribution as YYYY-MM-DD. DOI or persistent URL
- Originalsourcedata2.dta - original source data used to compile main analysis data set.
Source: AuthorLastName, AuthorFirstName. Title of Dataset, Year. Name of distributor [distributor], date of distribution as YYYY-MM-DD. DOI or persistent URL

IMPORTANT



If original source data cannot be shared within Dataverse due to access restrictions or proprietary nature of the data, please provide in-depth instructions for requesting access to data and the exact file names you requested. Please also make sure to inform the editors that you cannot share specific data due to copyright and access restrictions.

CODE

- Mainanalysis.do - uses datafile.dta to construct Figure 1, Figure 2, and Table 3 from the manuscript
- Appendixanalysis.do - uses Appendixdatafile.dta to construct Appendix Table 1, Table 2, and Figure 1
- Constructanalysisdata.do - uses original sourcedata.dta, or original sourcedata1.dta, and original sourcedata2.dta to construct datafile.dta, which is used to create manuscript results.



PRO-TIP

If using programs that do not have a coding environment (such as Excel, ArcGIS, QGIS, GRASS, etc.), you must instead provide detailed, step-by-step instructions for creating results using these programs.

DOCUMENTATION

- README.txt - This file, which lists all files and compute environment dependencies necessary to verify the results from this manuscript
- Codebook.PDF - codebook for datafile.dta, the analysis dataset used to generate results from this manuscript.
(Note: If citations for original source data are not included in your README, they must be included in your codebook. See the SPPQ Codebook Guide for more information and examples).

Examples of Complete and Successful READMEs

1. Gill, Jeff, 2020, "Replication Data for: Measuring Constituency Ideology Using Bayesian Universal Kriging", <https://doi.org/10.15139/S3/7NNASB>, UNC Dataverse, V2, UNF:6:g4eRjAsIUYrkiMSLD4V5yg== [fileUNF]
2. Sung Eun Kim; Johannes Urpelainen; Joonseok Yang, 2020, "Replication Data for: State Policy and Lobbying in a Federal System: Evidence from the Production Tax Credit for Renewable Energy, 1998-2012", <https://doi.org/10.15139/S3/KUT29A>, UNC Dataverse, V1, UNF:6:8vOafbQnnFLzjmCERFoNWg== [fileUNF]
3. Fisk, Colin A., 2019, "Replication Data for: No Republican, No Vote: Undervoting and Consequences of the Top-Two Primary System", <https://doi.org/10.15139/S3/YCSYUN>, UNC Dataverse, V1, UNF:6;jX9EXk JWcpsV1zAmZXP yFg== [fileUNF]
4. Nemerever, Zoe; Butler, Daniel, 2020, "Replication Data for: The Professionalism Advantage: Attracting, Fostering or Retaining Knowledge?", <https://doi.org/10.15139/S3/MPTNKT>, UNC Dataverse, V1
5. Wolak, Jennifer, 2020, "Replication Data for: Why Do People Trust their State Government?", <https://doi.org/10.15139/S3/TGKTEJ>, UNC Dataverse, V1, UNF:6:MK75wjgwAx12ETHhLqAIPw== [fileUNF]