13 June 2023

Readme file for
"Introducing HiSCoD: A New Gateway for the Study of Historical Social Conflict"
by Cédric Chambru and Paul Maneuvrier-Hervieu

The data and code used to support this study have been deposited in the Harvard Dataverse repository for the *American Political Science Review* (https://dataverse.harvard.edu/dataverse/the_review) and is accessible at the following url address: https://doi.org/10.7910/DVN/HGFLGK. Please download all replication materials from that repository.

Content

The replication folder contains all necessary data and code to replicate the results (i.e., tables and figures of the main text and the online appendix) referred in the present article. It contains three sub-folders:

i. "code" contains all functions and R scripts necessary to reproduce descriptive statistics and figures

ii. "data" contains the HiSCoD database as of 13 June 2023. The database is available in both English and French in two distinct formats: csv file (encoding: "UTF-8", sep = ";") and rds file from the R Project.

iii. "figures" contains Figures 1, 2, and 3 from the main text. They are reproduced by running scripts from the "code" folder.

NOTES: To read csv file in your spreadsheet software (e.g. LibreOffice Calc, Microsoft Excel, etc.), please use the tool to import data (e.g. “Data/Get External Data/From Text” in Microsoft Excel), then choose the file and set the encoding and the column delimiters to "UTF-8" and “;” respectively. Please note that using any other encoding, such as “Windows (ANSI)”, will break the file and result in misaligned columns and incorrectly displaying accented characters.

Reproducibility

All figures and statistics have been verified on 13 June 2023 using a machine running on Windows 10 (build 19044.2965) and R version 4.3.0. Packages used to run scripts included: *pacman* (version 0.5.1), *tidyverse* (version 2.0.0), *tidylog* (version 1.0.2), and *extrafont* (version 0.19).

To reproduce the descriptive statistics and figures reported in the paper, please proceed as follows:

1. Open the RStudio project file "01_project_hiscod.Rproj" in the folder "code/code_r", which will automatically set the working directory as ".../replication_folder/code/code_r".

2. Run the R script "02_master_code.R" to automatically install necessary packages, call various functions (e.g. to define a theme for producing figures), and execute scripts "03_hiscod_tables.R" and "04_hiscod_figures.R".

NOTES: If one wants to manually run scripts "03_hiscod_tables.R" and "04_hiscod_figures.R", execute script "02_master_code.R" until line 33, and then open and execute the above scripts.

IMPORTANT: if ones does not want to use RStudio project, open the R script "02_master_code.R" and set the working directory as: setwd("[path_to_replication_folder]/code/code_r"). Then, run line 28 to set all relative paths for other folders (data, figures, etc.).

Additional information

Detailed information on the variables contained in the database are available in the paper and the online appendix (see in particular Table 3 and Tables A.1).

- If you have any questions on the scope, contents, and/or limits of the data, please do not hesitate to contact us.

- If you need more information on the construction and coding of the variables, please do not hesitate to contact us.

For the most recent version of the database, please visit: https://github.com/hiscod/hiscod-project.

You may find additional information on the website of the project: https://www.unicaen.fr/hiscod.

Contributing

If you identify any mistake and/or want to contribute to the HiSCoD project by sharing data, please do not hesitate to contact us. We are very keen on improving its reliability and its scope.