Codebook

Historical Data Archive
(H-DATA)

Summary of Table of Contents

Table of Contents

1 Explanatory Notes

1.1 Release Notes v4

Demscore provides worldwide free access to harmonized data on Democracy, Environment, Migration, Social Policy, Conflict and Representation from several of the world's most prominent social science research institutes. The interdisciplinary nature of Demscore data facilitates large-scale comparative analyses. This is essential to advance adequate policy responses to complex societal challenges associated with the Sustainable Development Goals (SDGs) and beyond, facing Sweden, Europe, and the world today.

With a firm commitment to transparency and openness, Demscore v4 enables users to gain comprehensive insights into various topics across the social sciences. The joint infrastructure ensures data integrity and quality at the highest international standards and maximizes usability in the measurement of contextual data with 25.000 variables across nearly all countries in the world, from 1750 to the present.

This creates critical time- and cost saving advantages in data collection, management, distribution, and not the least for end-users in the scientific community. Demscore's unique approach to translating and merging data scales up to a total of 403.856 variable versions available in the infrastructure, storing more than 10 billion non-missing observations.

This collaborative effort between leading Swedish universities pushes the scale of social science data to a new level and offers unprecedented possibilities for interdisciplinary research and knowledge advancement.

These are the key features of Demscore:

1. **Customized Download:** A fully normalized, joint PostgreSQL database, sophisticated programming, and a user-friendly web-based interface for users to generate custom-designed datasets and codebooks for download.

2. **Translations and Data Merges:** Demscore currently offers more than 1000 merge options between datasets.

3. **Metadata:** Demscore takes information on and organization of metadata to new heights with the inclusion of customized codebooks, a detailed methodology document, and a comprehensive handbook.

4. **Handling of Missing Data:** Demscore pioneers in developing an innovative approach to tackle missing data. Researchers can now account for missing values with increased precision, leading to more robust and reliable analyses.

5. **Merge Scores:** Demscore introduces a unique merge mechanism. This powerful tool enables researchers to combine datasets effortlessly, uncovering connections and patterns that were previously hidden in isolated data silos.

6. **Thematic Datasets:** Demscore provides researchers with curated thematic datasets, each focused on a specific topic. These datasets bring together relevant variables from across the Demscore partners, facilitating in-depth investigations and comprehensive analyses of specific domains.

7. **Interactive Web Portal:** In addition to all the above, Demscore's web portal offers interactive visualization tools, user support and additional information on all partners and data sources.

For more information, please visit `https://www.demscore.se/` or contact contact@demscore.se.

1.2 New in Demscore version 4

A detailed description of changes and additions made for version 4 compared to version 3 can be found in the Methodology Document.

1.3 The Demscore Codebook

The autogenerated Demscore Codebook lists variable entries for those variables chosen by the user along with citation guidelines and licenses per variable.

The meta data is extracted from the codebooks per dataset stored in a table in the Demscore PostgreSQL database with one row per variable for all datasets. This table includes codebook entries, variable tags, labels, and other variable information in LaTeX format used to generate an automated codebook.

Demscore maintains a single set of standard entries for metadata across all datasets, to which all project members contribute their information. Additionally, variables within different datasets may have varying sets of additional information requirements specific to each dataset. These dataset-specific entries are also included, but they are presented as variable-specific metadata beneath the standard entries.

At the outset of the harmonization process, Demscore underwent a thorough variable name cleanup. This involved tasks such as replacing spaces or dots in variable names with underscores and converting all letters to lowercase. Notably, the original tags remain preserved and stored in the PostgreSQL table. Each variable in Demscore is accessible in both short and long forms. The short form comprises the cleaned version of the original variable tag, while the long form starts with the dataset name from which it originates, followed by the cleaned variable name.

For instance, the original name of the variable *MinisterPersonalID* from the H-DATA Foreign Minister Dataset is included as *ministerpersonalid* (short form) and *hdata_fomin_ministerpersonalid* (long form) in Demscore.

In addition, each dataset includes Demscore unit-identifier variables which are named according to the following naming scheme: Beginning with u_, followed by the name of the primary unit and finally the variable tag. The *year*- variable from the COMPLAB SPIN The Out-of-Work Benefits Dataset (OUTWB), which is part of the primary unit u_complab_country_year has the Demscore unit identifier name *u_complab_country_year_year*.

1.4 Methodology

For details on our methodology please see the Demscore Methodology document available for download on the Demscore website.

1.5 Citations

The Demscore project does not have a formal citation of its own. Hence, when using Demscore, we suggest that you cite the respective projects and datasets. We indicate how every dataset is to be cited in the autogenerated codebook you retreive with your data download, both in the dataset description and the codebook entry for each variable. Most often it is sufficient to cite the dataset a variable originates from, but sometimes there is a variable specific citation listed in the codebook entry in addition to that. For these cases, please also add the variable specific citation to the reference list of your publication. Full references are linked in the codebook entries of the variables and listed in the codebook's bibliography. We suggest you to also cite the Demscore Methodology Document when using data retrieved through Demscore.

1.6 Missing Data

Demscore indicates different types of missingness for observations in the customized datasets:
Missing in original data = Whenever an observation in the original variable is a missing (NA, missing code such as 7777, blank cell), we preserve this missing value. When the original source has special codes for various types of missing, those are preserved.
Missing code: -11111 = Demscore code for observation is missing due to the translation/merge, i.e., missing data due to no data being included for this combination of identifiers in the end Output Unit.
Missing code: -22222 = No observation is merged/translated, but the original data contains information for these identifier combinations elsewhere. For these cases, we use a different code. The

user needs to consult the reference documents (Methodology Document Section 5.1. or the Demscore Handbook) to clarify why the translation to the identifier combinations in the end Output Unit was not possible.

Please note that an observation that is missing in its original output unit does no take the value -11111, but appears as NA/blank cell in the customized dataset.

1.7 Download ID

The download ID can be shared with other users for replication purposes. A user can type the download ID into the Demscore website and retrieve the same download selection and files as the original user. This ID is autogenerated for each download from the Demscore website and will always retrieve the same data, even if the Demscore version was updated in the meantime.

Download ID:

1.8 Unit Identifier Variables

An Output Unit is defined as an output format in which variables can be retrieved from one or more datasets through a strictly defined output grid. A unit table defining this output grid contains unit identifier columns with u_ prefixes and the table is sorted based on these unit identifier columns and has a fixed number of rows. Unit columns are based on the columns that constitute the unit of analysis in a dataset. They are added to the original dataset and marked by a unit prefix (consisting of a u_ and the dataset unit name) before the original variable name. Unit columns can contain slightly modified data, e.g., missing values are replaced by a default value. Sometimes we add additional columns to the unit table, for instance if a dataset includes both a country_id column with a numeric country code, we add the variable storing the full country name to the unit table as well for better readability.

2 H-DATA

The Historical Data Archive (H-DATA) is a hub of historical country-level data running as far back as the French revolution (1789) and offers unparalleled depth of data and temporality, enabling researchers to answer critical questions about the past but to also understand the origins of, and find historical parallels to, present-day problems. H-DATA works to collect, integrate, and curate historical data from Demscore's other modules. By adding this long-term historical dimension, H-DATA makes it possible for researchers to study the path dependency of political institutions where changes are incremental or rare thus making long time-series essential to understanding their causes and consequences. By extending data back into time, H-DATA helps deepen and further our understanding of the conditions of the complex global challenges that we face today. More information is available on the project's website: `https://www.su.se/english/research/research-projects/h-data`

2.1 H-DATA Historical Cabinets

Dataset tag: hdata_cab

Output Unit:

Description: The Historical Cabinets (H-CAB) dataset is an historical extension of the Party Government in Europe Database (PAGED) and covers cabinets formed between 1919-1945 in 12 European countries (Austria, Belgium, Denmark, Finland, France, Germany, Iceland, Ireland, Netherland, Norway, Sweden, United Kingdom). In addition, new data for France was also collected for the period 1945-1958, covering the Fourth Republic which is not included in PAGED. The H-CAB dataset provides data on dates of cabinet formation and dissolution, along with data on the cabinet party composition, allowing for analysis on cabinet formation and duration during the interwar period.

Dataset citiation: When using this dataset, please cite the following:

Teorell, Jan, Johan Hellström Joseph Noonan (2023), Historical Cabinets (H-CAB) Dataset v1.0, `https://www.su.se/english/research/research-projects/h-data`

License: CC-BY-SA 4.0 International
`https://creativecommons.org/licenses/by-sa/4.0/legalcode`

More detailed information on the dataset can be found at the following web page: `https://www.su.se/english/research/research-projects/h-data/datasets-1.610144`

2.1.1 Identifiers

Variables or combination of variables in this section identify rows in the dataset.

2.1.1.1 Country Name (country)
Long tag: hdata_cab_country
Original tag: country
Dataset citation: Teorell et al. (2023)
Description:
    Name of country.

2.1.1.2 PAGED Country ID (paged_country_id)
Long tag: hdata_cab_paged_country_id
Original tag: paged_country_id
Dataset citation: Teorell et al. (2023)

Description:
 PAGED numeric country ID

2.1.1.3 V-Dem Country ID (vdem_country_id)

Long tag: hdata_cab_vdem_country_id

Original tag: vdem_country_id

Dataset citation: Teorell et al. (2023)

Description:
 V-Dem numeric country ID.

2.1.1.4 CoW Country ID (cow_country_id)

Long tag: hdata_cab_cow_country_id

Original tag: cow_country_id

Dataset citation: Teorell et al. (2023)

Description:
 COW numeric country ID.

2.1.1.5 Cabinet ID (cab_id)

Long tag: hdata_cab_cab_id

Original tag: cab_id

Dataset citation: Teorell et al. (2023)

Description:
 A unique ID following this pattern "{[paged_country_id{]H{[cabinet_number{]".

2.1.1.6 Cabinet Name (cabname)

Long tag: hdata_cab_cabname

Original tag: cabname

Dataset citation: Teorell et al. (2023)

Description:
 Name of the PM consecutively numbered with Roman numerals

2.1.1.7 Election Date (electdate)

Long tag: hdata_cab_electdate

Original tag: electdate

Dataset citation: Teorell et al. (2023)

Description:
 The date of the last election. Date of most recent parliamentary election: (a) two-round elections: date of second round (b) elections held over several days: date of last day.

2.1.1.8 Date In (date_in)

Long tag: hdata_cab_date_in

Original tag: date_in

Dataset citation: Teorell et al. (2023)

Description:
 he start date of a cabinet is determined by which ever of the following conditions first occurs (a) date that PMcabinet was appointed by head of state (b) date of investiture vote in parliament (c) date of general election.

2.1.1.9 Date Out (date_out)

Long tag: hdata_cab_date_out

Original tag: date_out

Dataset citation: Teorell et al. (2023)

Description:

The end date of a cabinet is determined by which ever of the following conditions first occurs (a) day of cabinet resignation (b) day of general election. In cases where governments go in exile due to invasions, the date_out is the date that they left the country. In cases of autocratization, date_out is an approximate date when the country ceased to be a democracy

2.1.2 Cabinet

These variables contail additional information on the cabinet.

2.1.2.1 Days (cab_duration_days)

Long tag: hdata_cab_cab_duration_days

Original tag: cab_duration_days

Dataset citation: Teorell et al. (2023)

Description:

Time in days between cabinet's first and last day in office.

2.1.2.2 Duration (form_duration)

Long tag: hdata_cab_form_duration

Original tag: form_duration

Dataset citation: Teorell et al. (2023)

Description:

Number of days between date_out of previous cabinet and date_in of current cabinet. During periods of long gaps due to autocratization or occupation during WWII form_duration is recorded as missing.

2.1.2.3 Alternative Duration (form_duration_alt)

Long tag: hdata_cab_form_duration_alt

Original tag: form_duration_alt

Dataset citation: Teorell et al. (2023)

Description:

Number of days between date_out of previous cabinet and date_in of current cabinet, unless an election occurs between date_out of the previous cabinet and and date_in of the current one. In these cases, form_duration_alt is the number of days between electdate and date_in.

2.1.2.4 Number of Parties (cab_num_parties)

Long tag: hdata_cab_cab_num_parties

Original tag: cab_num_parties

Dataset citation: Teorell et al. (2023)

Description:

Number of cabinet parties.

2.1.2.5 Cabinet Coalition (cab_coalition)

Long tag: hdata_cab_cab_coalition

Original tag: cab_coalition

Dataset citation: Teorell et al. (2023)

Description:

Coalition

0: no

1: yes

2.1.2.6 Cabinet Majority (cab_majority)

Long tag: hdata_cab_cab_majority

Original tag: cab_majority

Dataset citation: Teorell et al. (2023)

Description:
　　Cabinet majority (50percent + 1 seat)
　　0: no
　　1: yes

2.1.2.7 Cabinet Surplus (cab_surplus)

Long tag: hdata_cab_cab_surplus

Original tag: cab_surplus

Dataset citation: Teorell et al. (2023)

Description:
　　Cabinet surplus majority
　　0: no
　　1: yes

2.1.2.8 Government Type (govtype)

Long tag: hdata_cab_govtype

Original tag: govtype

Dataset citation: Teorell et al. (2023)

Description:
　　Government type.
　　1. Minority coalition cabinet: Holds less than 50 percent plus one seat in parliament. Single-party minority cabinets are likewise coded as 'Min'.
　　2. Single-party majority cabinet: Holds 50percent plus one seat in parliament. Is not a coalition.
　　3. Minimal winning coalition: Is turned into a losing coalition by the subtraction of any of the coalition parties, i.e., if it loses a coalition party it holds less than 50 percent plus one seat.
　　4. Surplus majority coalition: Can lose a coalition party and still be winning, i.e. control 50 percent plus one seat or more in the parliament.
　　5. A non-partisan cabinet, e.g. appointed by a president to hold an election

2.1.2.9 Seats (seats)

Long tag: hdata_cab_seats

Original tag: seats

Dataset citation: Teorell et al. (2023)

Description:
　　Total number of seats.

2.1.2.10 Prime Minister Party (pm_party)

Long tag: hdata_cab_pm_party

Original tag: pm_party

Dataset citation: Teorell et al. (2023)

Description:
　　PM's party (short name)

2.1.2.11 Total Seats (cab_seattotal)

Long tag: hdata_cab_cab_seattotal

Original tag: cab_seattotal

Dataset citation: Teorell et al. (2023)

Description:
   Total seats

2.1.2.12   Seatshare (cab_seatshare)
Long tag: hdata_cab_cab_seatshare

Original tag: cab_seatshare

Dataset citation: Teorell et al. (2023)

Description:
   Cabinet seat share

2.1.3   Parties

These variables contain information about the party composition of cabinets.

2.1.3.1   Party Name (cab_party_1)
Long tag: hdata_cab_cab_party_1

Original tag: cab_party_1

Dataset citation: Teorell et al. (2023)

Description:
   ParlGov short name of party 1

2.1.3.2   Party Name (cab_party_2)
Long tag: hdata_cab_cab_party_2

Original tag: cab_party_2

Dataset citation: Teorell et al. (2023)

Description:
   ParlGov short name of party 2

2.1.3.3   Party Name (cab_party_3)
Long tag: hdata_cab_cab_party_3

Original tag: cab_party_3

Dataset citation: Teorell et al. (2023)

Description:
   ParlGov short name of party 3

2.1.3.4   Party Name (cab_party_4)
Long tag: hdata_cab_cab_party_4

Original tag: cab_party_4

Dataset citation: Teorell et al. (2023)

Description:
   ParlGov short name of party 4

2.1.3.5   Party Name (cab_party_5)
Long tag: hdata_cab_cab_party_5

Original tag: cab_party_5

Dataset citation: Teorell et al. (2023)

Description:
   ParlGov short name of party 5

2.1.3.6 Party Name (cab_party_6)
Long tag: hdata_cab_cab_party_6

Original tag: cab_party_6

Dataset citation: Teorell et al. (2023)

Description:
ParlGov short name of party 6

2.1.3.7 Party Name (cab_party_7)
Long tag: hdata_cab_cab_party_7

Original tag: cab_party_7

Dataset citation: Teorell et al. (2023)

Description:
ParlGov short name of party 7

2.1.3.8 Party Name (cab_party_8)
Long tag: hdata_cab_cab_party_8

Original tag: cab_party_8

Dataset citation: Teorell et al. (2023)

Description:
ParlGov short name of party 8

2.1.3.9 Party Name (cab_party_9)
Long tag: hdata_cab_cab_party_9

Original tag: cab_party_9

Dataset citation: Teorell et al. (2023)

Description:
ParlGov short name of party 9

2.1.3.10 ParlGov ID (from ParlGov Dataset) (cab_party_parlgov_id_1)
Long tag: hdata_cab_cab_party_parlgov_id_1

Original tag: cab_party_parlgov_id_1

Dataset citation: Teorell et al. (2023)

Description:
ParlGov ID for party 1

2.1.3.11 ParlGov ID (from ParlGov Dataset) (cab_party_parlgov_id_2)
Long tag: hdata_cab_cab_party_parlgov_id_2

Original tag: cab_party_parlgov_id_2

Dataset citation: Teorell et al. (2023)

Description:
ParlGov ID for party 2

2.1.3.12 ParlGov ID (from ParlGov Dataset) (cab_party_parlgov_id_3)
Long tag: hdata_cab_cab_party_parlgov_id_3

Original tag: cab_party_parlgov_id_3

Dataset citation: Teorell et al. (2023)

Description:
ParlGov ID for party 3

2.1.3.13 ParlGov ID (from ParlGov Dataset) (cab_party_parlgov_id_4)

Long tag: hdata_cab_cab_party_parlgov_id_4

Original tag: cab_party_parlgov_id_4

Dataset citation: Teorell et al. (2023)

Description:
      ParlGov ID for party 3

2.1.3.14 ParlGov ID (from ParlGov Dataset) (cab_party_parlgov_id_5)

Long tag: hdata_cab_cab_party_parlgov_id_5

Original tag: cab_party_parlgov_id_5

Dataset citation: Teorell et al. (2023)

Description:
      ParlGov ID for party 6

2.1.3.15 ParlGov ID (from ParlGov Dataset) (cab_party_parlgov_id_6)

Long tag: hdata_cab_cab_party_parlgov_id_6

Original tag: cab_party_parlgov_id_6

Dataset citation: Teorell et al. (2023)

Description:
      ParlGov ID for party 5

2.1.3.16 ParlGov ID (from ParlGov Dataset) (cab_party_parlgov_id_7)

Long tag: hdata_cab_cab_party_parlgov_id_7

Original tag: cab_party_parlgov_id_7

Dataset citation: Teorell et al. (2023)

Description:
      ParlGov ID for party 7

2.1.3.17 ParlGov ID (from ParlGov Dataset) (cab_party_parlgov_id_8)

Long tag: hdata_cab_cab_party_parlgov_id_8

Original tag: cab_party_parlgov_id_8

Dataset citation: Teorell et al. (2023)

Description:
      ParlGov ID for party 8

2.1.3.18 ParlGov ID (from ParlGov Dataset) (cab_party_parlgov_id_9)

Long tag: hdata_cab_cab_party_parlgov_id_9

Original tag: cab_party_parlgov_id_9

Dataset citation: Teorell et al. (2023)

Description:
      ParlGov ID for party 9

2.1.3.19 Party Seats (cab_party_seats_1)

Long tag: hdata_cab_cab_party_seats_1

Original tag: cab_party_seats_1

Dataset citation: Teorell et al. (2023)

Description:
      Party seats lower chamber Party 1

2.1.3.20 Party Seats (cab_party_seats_2)

Long tag: hdata_cab_cab_party_seats_2

Original tag: cab_party_seats_2

Dataset citation: Teorell et al. (2023)

Description:

Party seats lower chamber Party 2

2.1.3.21 Party Seats (cab_party_seats_3)

Long tag: hdata_cab_cab_party_seats_3

Original tag: cab_party_seats_3

Dataset citation: Teorell et al. (2023)

Description:

Party seats lower chamber Party 3

2.1.3.22 Party Seats (cab_party_seats_4)

Long tag: hdata_cab_cab_party_seats_4

Original tag: cab_party_seats_4

Dataset citation: Teorell et al. (2023)

Description:

Party seats lower chamber Party 4

2.1.3.23 Party Seats (cab_party_seats_5)

Long tag: hdata_cab_cab_party_seats_5

Original tag: cab_party_seats_5

Dataset citation: Teorell et al. (2023)

Description:

Party seats lower chamber Party 5

2.1.3.24 Party Seats (cab_party_seats_6)

Long tag: hdata_cab_cab_party_seats_6

Original tag: cab_party_seats_6

Dataset citation: Teorell et al. (2023)

Description:

Party seats lower chamber Party 6

2.1.3.25 Party Seats (cab_party_seats_7)

Long tag: hdata_cab_cab_party_seats_7

Original tag: cab_party_seats_7

Dataset citation: Teorell et al. (2023)

Description:

Party seats lower chamber Party 7

2.1.3.26 Party Seats (cab_party_seats_8)

Long tag: hdata_cab_cab_party_seats_8

Original tag: cab_party_seats_8

Dataset citation: Teorell et al. (2023)

Description:

Party seats lower chamber Party 8

2.1.3.27   Party Seats (cab_party_seats_9)

Long tag: hdata_cab_cab_party_seats_9

Original tag: cab_party_seats_9

Dataset citation: Teorell et al. (2023)

Description:
   Party seats lower chamber Party 9

2.2   H-DATA Historical Conflicts - Country-Year Level

Dataset tag: hdata_conflict_cy

Output Unit:

Description:
Timespan: 1816-1945
Coverage: 480 wars spanning 130 countries
The Historical Conflict Dataset (HCD) dataset merges multiple existing datasets on wars from 1816-1945 to create a comprehensive list of interstate, intrastate, and extrasystemic wars during this time period. It uses an expanded definition of statehood from the International System(s) Dataset to reduce the Eurocentric bias in defining what constitutes a state. The ambition of this dataset is not to make methodological changes to how conflict is defined or measured, but rather to reclassify war based on a more inclusive definition of statehood and to create a dataset that increases ease of access to war data, particularly for scholars working with country-level historical data between 1816 to 1945.
The data is available in two formats: country-war and country-year. In the county-war datasets has a single case for each country participant and war. The country-year format contains binary variables on whether a country was involved in the different types of conflicts and has one case per country-year.

Dataset citiation: When using this dataset, please cite the following:

Noonan, Joseph Jan Teorell (2023), Historical Conflict Dataset (HCD) Dataset v1.0, `https://www.su.se/english/research/research-projects/h-data`

License: CC-BY-SA 4.0 International
`https://creativecommons.org/licenses/by-sa/4.0/legalcode`

More detailed information on the dataset can be found at the following web page: `https://www.su.se/english/research/research-projects/h-data/datasets-1.610144`

2.2.1   Identifiers

Variables or combination of variables in this section identify rows in the dataset.

2.2.1.1   ISD Country Code (isd_code)

Long tag: hdata_conflict_cy_isd_code

Original tag: isd_code

Dataset citation: Noonan & Teorell (2023)

Description:
   SD numeric country code

2.2.1.2   ISD Country (isd_country)

Long tag: hdata_conflict_cy_isd_country

Original tag: isd_country

Dataset citation: Noonan & Teorell (2023)

Description:
    ISD country name

2.2.1.3 ISD Region (isd_region)

Long tag: hdata_conflict_cy_isd_region

Original tag: isd_region

Dataset citation: Noonan & Teorell (2023)

Description:
    ISD regional classification

2.2.1.4 Gleditsch and Ward Country Code (gw_code)

Long tag: hdata_conflict_cy_gw_code

Original tag: gw_code

Dataset citation: Noonan & Teorell (2023)

Description:
    Gleditsch and Ward (1999) numeric country code

2.2.1.5 Gleditsch and Ward Country Name (gw_character)

Long tag: hdata_conflict_cy_gw_character

Original tag: gw_character

Dataset citation: Noonan & Teorell (2023)

Description:
    Gleditsch and Ward (1999) character country code

2.2.1.6 CoW Code (cow_code)

Long tag: hdata_conflict_cy_cow_code

Original tag: cow_code

Dataset citation: Noonan & Teorell (2023)

Description:
    Correlates of War country code

2.2.1.7 Cow Code (cow_country)

Long tag: hdata_conflict_cy_cow_country

Original tag: cow_country

Dataset citation: Noonan & Teorell (2023)

Description:
    Correlates of War country name

2.2.1.8 CoW Name (cow_character)

Long tag: hdata_conflict_cy_cow_character

Original tag: cow_character

Dataset citation: Noonan & Teorell (2023)

Description:
    Correlates of War character country code

2.2.1.9 V-Dem Country (v_dem_country)

Long tag: hdata_conflict_cy_v_dem_country

Original tag: v_dem_country

Dataset citation: Noonan & Teorell (2023)

Description:

V-Dem country name

2.2.1.10   V-Dem Country Code (v_dem_character)

Long tag: hdata_conflict_cy_v_dem_character

Original tag: v_dem_character

Dataset citation: Noonan & Teorell (2023)

Description:

V-Dem character country code

2.2.1.11   V-Dem Code (v_dem_code)

Long tag: hdata_conflict_cy_v_dem_code

Original tag: v_dem_code

Dataset citation: Noonan & Teorell (2023)

Description:

V-Dem numeric country code

2.2.1.12   Year (year)

Long tag: hdata_conflict_cy_year

Original tag: year

Dataset citation: Noonan & Teorell (2023)

Description:

Year

2.2.2   War classification variables

Variables in this section aim to contribute to updateing existing classifications of wars at the state level (interstate, intrastate, and extrasystemic) with a broader definition of statehood that reduces the Eurocentric bias in determining what constitutes a state and simoultaneously reclassify war based on a more inclusive definition of statehood.

2.2.2.1   Interstate War (inter_war)

Long tag: hdata_conflict_cy_inter_war

Original tag: inter_war

Dataset citation: Noonan & Teorell (2023)

Description:

Indicator for interstate war participation (1 = interstate war participation during the year)

2.2.2.2   Count Inter (count_inter_war)

Long tag: hdata_conflict_cy_count_inter_war

Original tag: count_inter_war

Dataset citation: Noonan & Teorell (2023)

Description:

The number of interstate wars

2.2.2.3   Interstate War Name (inter_war_name)

Long tag: hdata_conflict_cy_inter_war_name

Original tag: inter_war_name

Dataset citation: Noonan & Teorell (2023)

Description:

Name of the interstate war(s)

2.2.2.4 Intrastate War (intra_war)

Long tag: hdata_conflict_cy_intra_war

Original tag: intra_war

Dataset citation: Noonan & Teorell (2023)

Description:

Indicator intrastate war participation (1 = intrastate war participation during the year)

2.2.2.5 Count Intra (count_intra_war)

Long tag: hdata_conflict_cy_count_intra_war

Original tag: count_intra_war

Dataset citation: Noonan & Teorell (2023)

Description:

The number of intrastate wars

2.2.2.6 Intrastate War Name (intra_war_name)

Long tag: hdata_conflict_cy_intra_war_name

Original tag: intra_war_name

Dataset citation: Noonan & Teorell (2023)

Description:

Name of the intrastate war(s)

2.2.2.7 External Participants (external_participant_intra_state)

Long tag: hdata_conflict_cy_external_participant_intra_state

Original tag: external_participant_intra_state

Dataset citation: Noonan & Teorell (2023)

Description:

Indicator for states which participated in another state's intrastate war

2.2.2.8 External Participants Side (external_participant_intra_state_side)

Long tag: hdata_conflict_cy_external_participant_intra_state_side

Original tag: external_participant_intra_state_side

Dataset citation: Noonan & Teorell (2023)

Description:

Indicator which classifies external_participant_intra_state as either participating in a war as (1) against the government, (2) in support of the government

2.2.2.9 Extra War (extra_war)

Long tag: hdata_conflict_cy_extra_war

Original tag: extra_war

Dataset citation: Noonan & Teorell (2023)

Description:

Indicator for extrasystemic war (between as state and a non-state group outside its own territory) participation (1 = extrasystemic war participation during the year)

2.2.2.10 Count Extra War Name (count_extra_war)

Long tag: hdata_conflict_cy_count_extra_war

Original tag: count_extra_war

Dataset citation: Noonan & Teorell (2023)

Description:

The number of extra systemic wars

2.2.2.11 Extra War Name (extra_war_name)

Long tag: hdata_conflict_cy_extra_war_name

Original tag: extra_war_name

Dataset citation: Noonan & Teorell (2023)

Description:
  Name of the extra systemic war(s)

2.3 H-DATA Historical Conflicts - War Level

Dataset tag: hdata_conflict_war

Output Unit:

Description:
Timespan: 1816-1945
Coverage: 480 wars spanning 130 countries
The Historical Conflict Dataset (HCD) dataset merges multiple existing datasets on wars from 1816-1945 to create a comprehensive list of interstate, intrastate, and extrasystemic wars during this time period. It uses an expanded definition of statehood from the International System(s) Dataset to reduce the Eurocentric bias in defining what constitutes a state. The ambition of this dataset is not to make methodological changes to how conflict is defined or measured, but rather to reclassify war based on a more inclusive definition of statehood and to create a dataset that increases ease of access to war data, particularly for scholars working with country-level historical data between 1816 to 1945.
The data is available in two formats: country-war and country-year. In the county-war datasets has a single case for each country participant and war. The country-year format contains binary variables on whether a country was involved in the different types of conflicts and has one case per country-year.

Dataset citiation: When using this dataset, please cite the following:

Noonan, Joseph Jan Teorell (2023), Historical Conflict Dataset (HCD) Dataset v1.0, `https://www.su.se/english/research/research-projects/h-data`

License: CC-BY-SA 4.0 International
`https://creativecommons.org/licenses/by-sa/4.0/legalcode`

More detailed information on the dataset can be found at the following web page: `https://www.su.se/english/research/research-projects/h-data/datasets-1.610144`

2.3.1 Identifiers

Variables or combination of variables in this section identify rows in the dataset.

2.3.1.1 ISD Country Code (isd_code)

Long tag: hdata_conflict_war_isd_code

Original tag: isd_code

Dataset citation: Noonan & Teorell (2023)

Description:
  SD numeric country code

2.3.1.2 ISD Country (isd_country)

Long tag: hdata_conflict_war_isd_country

Original tag: isd_country

Dataset citation: Noonan & Teorell (2023)

Description:
ISD country name

2.3.1.3 ISD Region (isd_region)

Long tag: hdata_conflict_war_isd_region

Original tag: isd_region

Dataset citation: Noonan & Teorell (2023)

Description:
ISD regional classification

2.3.1.4 Gleditsch and Ward Country Code (gw_code)

Long tag: hdata_conflict_war_gw_code

Original tag: gw_code

Dataset citation: Noonan & Teorell (2023)

Description:
Gleditsch and Ward (1999) numeric country code

2.3.1.5 Gleditsch and Ward Country Name (gw_character)

Long tag: hdata_conflict_war_gw_character

Original tag: gw_character

Dataset citation: Noonan & Teorell (2023)

Description:
Gleditsch and Ward (1999) character country code

2.3.1.6 CoW Country Code (cow_code)

Long tag: hdata_conflict_war_cow_code

Original tag: cow_code

Dataset citation: Noonan & Teorell (2023)

Description:
Correlates of War country code

2.3.1.7 CoW Country Code (cow_country)

Long tag: hdata_conflict_war_cow_country

Original tag: cow_country

Dataset citation: Noonan & Teorell (2023)

Description:
Correlates of War country name

2.3.1.8 CoW Country Name (cow_character)

Long tag: hdata_conflict_war_cow_character

Original tag: cow_character

Dataset citation: Noonan & Teorell (2023)

Description:
Correlates of War character country code

2.3.1.9 V-Dem Country Code (v_dem_country)

Long tag: hdata_conflict_war_v_dem_country

Original tag: v_dem_country

Dataset citation: Noonan & Teorell (2023)

Description:
    V-Dem character country name

2.3.1.10 V-Dem Country Name (v_dem_character)

Long tag: hdata_conflict_war_v_dem_character

Original tag: v_dem_character

Dataset citation: Noonan & Teorell (2023)

Description:
    V-Dem character country code

2.3.1.11 V-Dem Code (v_dem_code)

Long tag: hdata_conflict_war_v_dem_code

Original tag: v_dem_code

Dataset citation: Noonan & Teorell (2023)

Description:
    V-Dem numeric country code

2.3.2 War classification variables

Variables in this section aim to contribute to updateing existing classifications of wars at the state level (interstate, intrastate, and extrasystemic) with a broader definition of statehood that reduces the Eurocentric bias in determining what constitutes a state and simoultaneously reclassify war based on a more inclusive definition of statehood.

2.3.2.1 War Name (war_name)

Long tag: hdata_conflict_war_war_name

Original tag: war_name

Dataset citation: Noonan & Teorell (2023)

Description:
    Name of conflict

2.3.2.2 War Type (war_type)

Long tag: hdata_conflict_war_war_type

Original tag: war_type

Dataset citation: Noonan & Teorell (2023)

Description:
    War classifications which
    consist of three different types of war:
    1. interstate: A war between two states
    2. intrastate: A war between a state and a non-state actor, inside of its territory.
    3. extrasystemic: A war between a state and a non-state group outside its territory.

2.3.2.3 Start Year (min_year)

Long tag: hdata_conflict_war_min_year

Original tag: min_year

Dataset citation: Noonan & Teorell (2023)

Description:
    Year when conflict began

2.3.2.4 End Year (max_year)

Long tag: hdata_conflict_war_max_year

Original tag: max_year

Dataset citation: Noonan & Teorell (2023)

Description:
    Year when conflict ended

2.3.2.5 External Participants (external_participant_intra_state)

Long tag: hdata_conflict_war_external_participant_intra_state

Original tag: external_participant_intra_state

Dataset citation: Noonan & Teorell (2023)

Description:
    Indicator for states which participated in another country's
    intrastate war.

2.3.2.6 External Participants Side (external_participant_intra_state_side)

Long tag: hdata_conflict_war_external_participant_intra_state_side

Original tag: external_participant_intra_state_side

Dataset citation: Noonan & Teorell (2023)

Description:
    Indicator which classifies external_participant_intra_state
    as either participating in a conflict as (1) against government, (2) in support of government

2.3.2.7 Gleditsch Deaths (g_deaths)

Long tag: hdata_conflict_war_g_deaths

Original tag: g_deaths

Dataset citation: Noonan & Teorell (2023)

Description:
    Gleditsch (2004) death estimates.
    There are high levels of uncertainty and misssingness with all the death estimates. In some
    cases, only a single estimate is given in CoW when the HCD determines that multiple states
    have participated. For instance, The British-Maratha War of 1817–1818, CoW provides
    causality estimates the total number for all members of the Maratha Confederacy (2,000). In
    the HCD each individual state of the Maratha Confederacy (Nagpur, Poona, Indore and
    Gwalior) is included in the dataset. As we do not know the specific spread of causalities
    between these four states, the estimates are evenly distributed between them. As such, these
    estimates should be treated with caution. Note that in Gleditsch (2004) and CoW, interstate
    wars are broken up into two cases if a state switches side during the war. HCD only has one
    record for each war participant for all interstate wars. In HCD the casualty estimates for
    cases where a participant switched sides are the combined total deaths of the two records in
    Gleditsch (2004) or CoW.

2.3.2.8 CoW State Deaths (cow_state_deaths)

Long tag: hdata_conflict_war_cow_state_deaths

Original tag: cow_state_deaths

Dataset citation: Noonan & Teorell (2023)

Description:
    CoW estimate of state battle deaths.

2.3.2.9 CoW Non-State Deaths (cow_non_state_deaths)

Long tag: hdata_conflict_war_cow_non_state_deaths

Original tag: cow_non_state_deaths

Dataset citation: Noonan & Teorell (2023)

Description:
    CoW estimate of non-state battle deaths (where available) as Gleditsch (2004) does not differentiate between state and non-state battle deaths in g_deaths. Only available for intrastate and extrasystemic wars.

2.4 H-DATA Diplomatic Representation

Dataset tag: hdata_direp

Output Unit: H-DATA Dyad-Year, i.e., data is collected per country dyad and year. The dyad consist of two country variables, with country 2 being the country that is diplomatically represented by country 1 in a given year.

Description: This dataset builds upon the Correlates of War (COW) dataset Diplomatic Exchange, 1817-2005 (v2006.1) by expanding the country and temporal coverage of diplomatic representation during the 19th century by including states that have been excluded from the COW sample. The dataset is dyadic with each row containing two sets of identifiers, with the key variable being the level of diplomatic representation of country 2 by country 1. This dataset was used in Jan Teorell's article Rules of recognition? Explaining diplomatic representation since the Congress of Vienna published in Cooperation and Conflict in 2022.

Dataset citiation: When using this dataset, please cite the following paper:

Teorell, J. (2022), Rules of recognition? Explaining diplomatic representation since the Congress of Vienna, *Cooperation and Conflict*, `https://doi.org/10.1177/00108367221093151`

License: CC-BY-SA 4.0 International
`https://creativecommons.org/licenses/by-sa/4.0/legalcode`

More detailed information on the dataset can be found at the following web page: `https://www.su.se/english/research/research-projects/h-data/datasets-1.610144`

2.4.1 Identifier Variables

Variables or combination of variables in this section identify rows in the dataset.

2.4.1.1 Year (year)
Long tag: hdata_diprep_year
Original tag: year
Dataset citation: Teorell (2022)
Description:
    Year of observation

2.4.1.2 Country Name 1 (country_name1)
Long tag: hdata_diprep_country_name1
Original tag: country_name1
Dataset citation: Teorell (2022)
Description:
    The country name for unit 1

2.4.1.3 COW Code 1 (cow_code1)
Long tag: hdata_diprep_cow_code1
Original tag: cow_code1
Dataset citation: Teorell (2022)

Description:
   Correlates of War code for unit 1

2.4.1.4 V-Dem Code 1 (vdem_code1)

Long tag: hdata_diprep_vdem_code1

Original tag: vdem_code1

Dataset citation: Teorell (2022)

Description:
   V-Dem code for unit 1

2.4.1.5 Country Name 2 (country_name2)

Long tag: hdata_diprep_country_name2

Original tag: country_name2

Dataset citation: Teorell (2022)

Description:
   The country name for unit 2

2.4.1.6 COW Code 2 (cow_code2)

Long tag: hdata_diprep_cow_code2

Original tag: cow_code2

Dataset citation: Teorell (2022)

Description:
   Correlates of War code for unit 2

2.4.1.7 V-Dem Code 2 (vdem_code2)

Long tag: hdata_diprep_vdem_code2

Original tag: vdem_code2

Dataset citation: Teorell (2022)

Description:
   V-Dem code for unit 2

2.4.2 Diplomatic Representation

Variables in this section indicate the level of diplomatic representation by a country 1.

2.4.2.1 Diplomatic Representation (diprep_dr)

Long tag: hdata_diprep_diprep_dr

Original tag: diprep_dr

Dataset citation: Teorell (2022)

Description:
   DIPREP_DR is the coded variable of diplomatic representation level of country_name2 by
   country_name1. The coding of DIPREP_DR follows Bayer (2006):
   (0) No evidence of diplomatic exchange
   (1) Chargé d'affaires
   (2) Minister
   (3) Ambassador
   (9) Other

2.4.3 Aggregated Variables

This section includes variables aggregated to a country year level. They are not part of the original
data. The variables indicate how many countries a country 1 represents at which level in a given

year.

2.4.3.1 Count Diplomatic Representation at Level 0 (count_diprep_level_0)

Long tag: hdata_diprep_count_diprep_level_0

Original tag: count_diprep_level_0

Dataset citation: Teorell (2022)

Description:

New variable created for the level of representation in the aggregated dataset. The variable indicates how many countries are represented by a country 1 in a given year at the diprep_dr level 0 *No evidence of diplomatic exchange.*

2.4.3.2 Count Diplomatic Representation at Level 1 (count_diprep_level_1)

Long tag: hdata_diprep_count_diprep_level_1

Original tag: count_diprep_level_1

Dataset citation: Teorell (2022)

Description:

New variable created for the level of representation in the aggregated dataset. The variable indicates how many countries are represented by a country 1 in a given year at the diprep_dr level 1 *Chargé d'affaires.*

2.4.3.3 Count Diplomatic Representation at Level 2 (count_diprep_level_2)

Long tag: hdata_diprep_count_diprep_level_2

Original tag: count_diprep_level_2

Dataset citation: Teorell (2022)

Description:

New variable created for the level of representation in the aggregated dataset. The variable indicates how many countries are represented by a country 1 in a given year at the diprep_dr level 2 *Minister.*

2.4.3.4 Count Diplomatic Representation at Level 3 (count_diprep_level_3)

Long tag: hdata_diprep_count_diprep_level_3

Original tag: count_diprep_level_3

Dataset citation: Teorell (2022)

Description:

New variable created for the level of representation in the aggregated dataset. The variable indicates how many countries are represented by a country 1 in a given year at the diprep_dr level 3 *Ambassador.*

2.4.3.5 Count Diplomatic Representation at Level 9 (count_diprep_level_9)

Long tag: hdata_diprep_count_diprep_level_9

Original tag: count_diprep_level_9

Dataset citation: Teorell (2022)

Description:

New variable created for the level of representation in the aggregated dataset. The variable indicates how many countries are represented by a country 1 in a given year at the diprep_dr level 9 *Other.*

2.5 H-DATA Foreign Minister Dataset

Dataset tag: hdata_fomin

Output Unit: H-DATA Minister-Year, i.e., data is collected per foreign minister and the date they

got into office. That means each row in the dataset can be identified by one minister in combination with a date, using the columns foreignminister and date_in. The date_in column does nor exist in the original dataset but is a concatenation of the columns fminyear, fminmonth and fminday. Several other columns, such as date_out, country_name, etc. are added to the unit table in order to aggregate and later translate to other Output Units.

Description: For their article "War, Performance, and the Survival of Foreign Ministers", Hanna Bäck, Jan Teorell, Alexander Von Hagen-Jamar and Alejandro Quiroz Flores created The Foreign Minister Dataset. The Foreign Minister Dataset consists of comparative historical data on foreign ministers' background and reasons for leaving office in the world's 13 former and current great powers from 1789 to the present. The data covers 1155 regular (non-acting) foreign ministers, as well as partial information on 173 acting foreign ministers, for the following 13 great powers: Austria (the Habsburg Empire/Austria-Hungary), Britain, China (Qing Empire/Republic/People's Republic of China), France, Italy, Japan, the Netherlands, Prussia/Germany, the Ottoman Empire/Turkey, Russia, Spain, Sweden and the United States.

Dataset citiation: When using this dataset, please cite the following paper:

Hanna Bäck, Jan Teorell, Alexander Von Hagen-Jamar, Alejandro Quiroz Flores, War, Performance, and the Survival of Foreign Ministers, *Foreign Policy Analysis*, Volume 17, Issue 2, April 2021, oraa024, `https://doi.org/10.1093/fpa/oraa024`

License: CC-BY-SA 4.0 International
`https://creativecommons.org/licenses/by-sa/4.0/legalcode`

More detailed information on the dataset can be found at the following web page: `https://www.su.se/english/research/research-projects/h-data/datasets-1.610144`

2.5.1 Identifier Variables

Identification variables in the dataset.

2.5.1.1 Country code (ccode)
Long tag: hdata_fomin_ccode
Original tag: ccode
Dataset citation: Bäck et al. (2021)
Description:
    Correlates of War (COW) country code. Country codes based on Flores' dataset.

2.5.1.2 Minister Personal ID (ministerpersonalid)
Long tag: hdata_fomin_ministerpersonalid
Original tag: MinisterPersonalID
Dataset citation: Bäck et al. (2021)
Description:
    Country-code followed by 3-digit consecutive number.

2.5.1.3 Country (country_name)
Long tag: hdata_fomin_country_name
Original tag: country_name
Dataset citation: Bäck et al. (2021)
Description:
    Country name

2.5.1.4 Historical Country name (country_name2)

Long tag: hdata_fomin_country_name2

Original tag: country_name2

Dataset citation: Bäck et al. (2021)

Description:
    Historical country name

2.5.1.5 Foreign Minister (foreignminister)

Long tag: hdata_fomin_foreignminister

Original tag: foreignminister

Dataset citation: Bäck et al. (2021)

Description:
    Full name of the FM. When entering new names, all apostrophes and such are removed. Much data is imported from rulers.org, and therefore dates of office and birth/death are at times included with the names, as well as information on whether they were acting ministers and if they had recurring terms in office.

2.5.2 Time in Office

These variables capture the time the foreign minister held office.

2.5.2.1 Foreign Minister entrance day (fminday)

Long tag: hdata_fomin_fminday

Original tag: fminday

Dataset citation: Bäck et al. (2021)

Description:
    Day foreign minister entered office

2.5.2.2 Foreign Minister entrance month (fminmonth)

Long tag: hdata_fomin_fminmonth

Original tag: fminmonth

Dataset citation: Bäck et al. (2021)

Description:
    Month foreign minister entered office

2.5.2.3 Foreign Minister entrance year (fminyear)

Long tag: hdata_fomin_fminyear

Original tag: fminyear

Dataset citation: Bäck et al. (2021)

Description:
    Year foreign minister entered office

2.5.2.4 Foreign Minister exit day (fmoutday)

Long tag: hdata_fomin_fmoutday

Original tag: fmoutday

Dataset citation: Bäck et al. (2021)

Description:
    Day foreign minister exited office

2.5.2.5 Foreign Minister exit month (fmoutmonth)

Long tag: hdata_fomin_fmoutmonth

Original tag: fmoutmonth

Dataset citation: Bäck et al. (2021)

Description:
　　Month foreign minister exited office

2.5.2.6   Foreign Minister exit year (fmoutyear)

Long tag: hdata_fomin_fmoutyear

Original tag: fmoutyear

Dataset citation: Bäck et al. (2021)

Description:
　　Year foreign minister exited office

2.5.3   Position

These variables give information on the position of the foreign minister.

2.5.3.1   Name of ministry (portfolio_name)

Long tag: hdata_fomin_portfolio_name

Original tag: portfolio_name

Dataset citation: Bäck et al. (2021)

Description:
　　Free text, name of the ministry (or other) over which the FM had authority. In English where available, otherwise in original language.

2.5.3.2   Name of the position (post_name)

Long tag: hdata_fomin_post_name

Original tag: post_name

Dataset citation: Bäck et al. (2021)

Description:
　　Name of the position, e.g. Foreign Minister, Minister of Foreign Affairs, Secretary of State. In English where available, otherwise in original language

2.5.3.3   Acting (acting)

Long tag: hdata_fomin_acting

Original tag: acting

Dataset citation: Bäck et al. (2021)

Description:
　　Interim and such appointments are coded as acting. All current entries are based on Flores and Truhart.
　　1. No
　　2. Yes
　　3. Started as acting and became proper

2.5.3.4   Term number (term_no)

Long tag: hdata_fomin_term_no

Original tag: Term_no

Dataset citation: Bäck et al. (2021)

Description:
　　The number of the term the minister is serving. 1 for the first term, 2 for the second, 3 for the third etc.

2.5.3.5 Repeated terms (repeated_terms)

Long tag: hdata_fomin_repeated_terms

Original tag: repeated_terms

Dataset citation: Bäck et al. (2021)

Description:

Based on sources Flores and Truhart, which are unreliable on this matter, so for those ministers where full data is not entered and this data checked, caution is suggested.
1. No — only held this office one
2. Yes — held office two times or more.

2.5.3.6 Head of government (headofgovernment)

Long tag: hdata_fomin_headofgovernment

Original tag: HeadofGovernment

Dataset citation: Bäck et al. (2021)

Description:

Is the foreign minister also head of government (prime minister/state chancellor/president/king or similar).
1. No
2. Yes

2.5.3.7 Manner out (manner_out)

Long tag: hdata_fomin_manner_out

Original tag: Manner_out

Dataset citation: Bäck et al. (2021)

Description:

How did the FM leave the post as FM? Removal from office is coded as 'regular' when the FM is removed in accordance with explicit rules or established conventions of his or her particular country. Examples include voluntary retirement, dismissal by head of government ('forced resignation') and defeat in elections. Removal from office is coded as 'irregular' when the FM was removed by domestic forces in contravention of explicit rules and established conventions. Irregular removal from office is overwhelmingly the result of the threat or use of force as exemplified in coups, (popular) revolts and assassinations (regardless of whether the motivation behind the assassination was explicitly political or not). Foreign removal includes only cases where a foreign state directly removes a foreign leader, for example through invasion or kidnapping. In particular, cases where someone is seen as orchestrating a removal of a leader through a coup carried out by domestic forces are not coded as foreign removal, but as an irregular procedure.
1. By regular procedures
2. Death by natural causes
3. Retired due to ill health
4. Suicide
5. By irregular procedures
6. By deposition by a foreign state

2.5.3.8 Reason out (reason_out)

Long tag: hdata_fomin_reason_out

Original tag: Reason_out

Dataset citation: Bäck et al. (2021)

Description:

When exiting by regular procedures (when manner_out is coded 1, otherwise this variable should be coded 9999), why did the FM leave the post as FM?

1. End of government/party leaves government:

11. End of constitutionally mandated term/elections

12. Other reason for end of government (e.g. new PM/change in party composition)

13. Party of minister leaves government

2. Forced resignation (removed by head of government or similar) primarily due to:

21. Political scandal

22. Policy disagreement between minister and premier/PM

23. Policy disagreement between minister and monarch/president

24. Policy disagreement between minister and own party/other minister

25. Personal/departmental error or low personal performance

26. Move to other post within cabinet (individual move of the FM to another post)

27. Move to another post within cabinet in the general context of a reshuffle

28. Loss of eligibility for the post

29. Other reason

3. Unforced resignation primarily due to:

31. Move to alternative (high) position outside of cabinet

32. Other reason (e.g. voluntary retirement, not due to ill health)

2.5.4 Personal Background

Variables providing information on the personal background of the foreign minister.

2.5.4.1 Sex (sex)

Long tag: hdata_fomin_sex
Original tag: sex
Dataset citation: Bäck et al. (2021)
Description:
    Sex:
    1. Male
    2. Female

2.5.4.2 Birth Year (birthyear)

Long tag: hdata_fomin_birthyear
Original tag: birthyear
Dataset citation: Bäck et al. (2021)
Description:
    Birth year

2.5.4.3 Death Year (deadyear)

Long tag: hdata_fomin_deadyear

Original tag: deadyear

Dataset citation: Bäck et al. (2021)

Description:
Death year

2.5.4.4   Birthplace (birthplace)

Long tag: hdata_fomin_birthplace

Original tag: birthplace

Dataset citation: Bäck et al. (2021)

Description:
Free-text: City, comma country. When taken from Truhart sometimes only city is stated and can be completed with contemporary country later

2.5.4.5   Marital status (marital)

Long tag: hdata_fomin_marital

Original tag: marital

Dataset citation: Bäck et al. (2021)

Description:
Marital status at time of entering office (each time)
1. Single
2. Married
3. Divorced
4. Widowed
5. Remarried

2.5.4.6   Children (children)

Long tag: hdata_fomin_children

Original tag: children

Dataset citation: Bäck et al. (2021)

Description:
1. No
2. Yes

2.5.4.7   Number of Children (nr_children)

Long tag: hdata_fomin_nr_children

Original tag: nr_children

Dataset citation: Bäck et al. (2021)

Description:
The number of children when entering office. If children=1, enter 9999.

2.5.4.8   Family: noble (fambgnoble)

Long tag: hdata_fomin_fambgnoble

Original tag: fambgnoble

Dataset citation: Bäck et al. (2021)

Description:
Was the family (father) noble/royal/similar:
1. No
2. Yes

2.5.4.9   Family: profession (fambgprof)

Long tag: hdata_fomin_fambgprof

Original tag: fambgprof

Dataset citation: Bäck et al. (2021)

Description:
   The family (father's) main profession (during FM's childhood)
   1. Farming (incl. landowner)
   2. Business (trader, merchant, crafts)
   3. Worker (blue collar employees)
   4. Clerical, religious
   5. Military
   6. Politics/state (employed and elected, within the country)
   7. Other white collar employee
   8. International organisations
   9. Other

2.5.5 Educational Background

Variables providing information on the educational background of the foreign minister.

2.5.5.1 Education (education)

Long tag: hdata_fomin_education

Original tag: education

Dataset citation: Bäck et al. (2021)

Description:
   Level of education (the highest level of graduated studies). If education was provided by a
   tutor/governess in private, code 1.
   1. Primary only
   2. Primary & Secondary only
   3. Higher education non-university
   4. University/college
   5. Post-graduate

2.5.5.2 Education: Field of education (edufield)

Long tag: hdata_fomin_edufield

Original tag: edufield

Dataset citation: Bäck et al. (2021)

Description:
   Field of education (if more than one field, use the two first digits to indicate the first field of
   higher education and the two last to indicate the other field of higher education, e.g. 0405 if
   mathematics and chemistry, 0708 if social sciences and law)
   01. Agronomy
   02. Economics, business, management
   03. Engineering
   04. Mathematics, computer science
   05. Biology, chemistry, physics
   06. Humanities
   07. Social sciences
   08. Law
   09. Medicine
   10. Military
   11. Theology
   12. Other

2.5.5.3 Education: School (school_name)

Long tag: hdata_fomin_school_name

Original tag: school_name

Dataset citation: Bäck et al. (2021)

Description:

Free text name of school from which they have their highest degree, including city/country.

2.5.5.4 Education: School (multiple) (sec_school_name)

Long tag: hdata_fomin_sec_school_name

Original tag: sec_school_name

Dataset citation: Bäck et al. (2021)

Description:

In case of multiple schools (e.g. undergrad degree at one and PhD at another) enter the second highest degree school here.

2.5.6 Occupational Background

Variables providing information on the occupational background of the foreign minister.

2.5.6.1 Occupation: Farmer (farmer)

Long tag: hdata_fomin_farmer

Original tag: Farmer

Dataset citation: Bäck et al. (2021)

Description:

Did the FM have experience as a farmer/landowner before appointment as FM?
1. No
2. Yes

2.5.6.2 Occupation: Businessman (business)

Long tag: hdata_fomin_business

Original tag: Business

Dataset citation: Bäck et al. (2021)

Description:

Did the FM have experience as a trader/merchant or other businessman before appointment as FM?
1. No
2. Yes

2.5.6.3 Occupation: Worker (worker)

Long tag: hdata_fomin_worker

Original tag: Worker

Dataset citation: Bäck et al. (2021)

Description:

Did the FM have experience as a blue-collar employee before appointment as FM?
1. No
2. Yes

2.5.6.4 Occupation: Cleric (cleric)

Long tag: hdata_fomin_cleric

Original tag: Cleric

Dataset citation: Bäck et al. (2021)

Description:

Did the FM have experience of a clerical/religious job before appointment as FM?

1. No
2. Yes

2.5.6.5 Occupation: Politics (politics)

Long tag: hdata_fomin_politics

Original tag: Politics

Dataset citation: Bäck et al. (2021)

Description:

Did the FM have experience from domestic politics, as employed or elected, before becoming FM? Not that either employment/work as minister, civil servant or as a legislative deputy counts.
1. No
2. Yes

2.5.6.6 Occupation: Law (law)

Long tag: hdata_fomin_law

Original tag: Law

Dataset citation: Bäck et al. (2021)

Description:

Did the FM have experience of working as a lawyer or other judicial work before appointment as FM?
1. No
2. Yes

2.5.6.7 Occupation: White collar (whitecollar)

Long tag: hdata_fomin_whitecollar

Original tag: Whitecollar

Dataset citation: Bäck et al. (2021)

Description:

Did the FM have experience as a white-collar employee before appointment as FM?
1. No
2. Yes

2.5.6.8 Occupation: International Organization (int_org)

Long tag: hdata_fomin_int_org

Original tag: Int_org

Dataset citation: Bäck et al. (2021)

Description:

Did the FM have experience of working within an international organisation before appointment as FM?
1. No
2. Yes

2.5.6.9 Occupation: International Organization name (internationalorganizationname)

Long tag: hdata_fomin_internationalorganizationname

Original tag: InternationalOrganizationname

Dataset citation: Bäck et al. (2021)

Description:

If the FM had experience of working within an international organization before appointment, free text the name(s) of the organization(s) (if several, separate with commas).

2.5.7 Diplomatic Background

Variables providing information on to which extent the foreign minister has a diplomatic background.

2.5.7.1 Diplomat (diplomat)

Long tag: hdata_fomin_diplomat

Original tag: Diplomat

Dataset citation: Bäck et al. (2021)

Description:
    Did the FM hold a diplomatic office before becoming FM?
    1. No
    2. Yes

2.5.7.2 Diplomat: Post (dip_post)

Long tag: hdata_fomin_dip_post

Original tag: Dip_post

Dataset citation: Bäck et al. (2021)

Description:
    If the FM had held a diplomatic office, free text the post name.

2.5.7.3 Diplomat: Country (dip_country)

Long tag: hdata_fomin_dip_country

Original tag: Dip_country

Dataset citation: Bäck et al. (2021)

Description:
    If the FM had held a diplomatic office, free text the country/ies that these offices were held in (separate with commas). If dip_post =1, enter 9999.

2.5.8 Military Background

Variables providing information on to which extent the foreign minister has a military background.

2.5.8.1 Military: Education (mili_edu)

Long tag: hdata_fomin_mili_edu

Original tag: Mili_edu

Dataset citation: Bäck et al. (2021)

Description:
    Did the FM have military education before (each) appointment as FM?
    1. No
    2. Yes

2.5.8.2 Military: Service (mili_service)

Long tag: hdata_fomin_mili_service

Original tag: Mili_service

Dataset citation: Bäck et al. (2021)

Description:
    Did the FM serve in the military before (each) appointment as FM?
    1. No
    2. Yes

2.5.8.3 Military: Rank (mili_rank)

Long tag: hdata_fomin_mili_rank

Original tag: Mili_rank

Dataset citation: Bäck et al. (2021)

Description:

Free text the highest awarded military rank before (each) appointment as FM? If mili_service=1, enter 9999.

2.5.8.4 Military: Combat (mili_combat)

Long tag: hdata_fomin_mili_combat

Original tag: Mili_combat

Dataset citation: Bäck et al. (2021)

Description:

Did the FM have combat experience before (each) appointment as FM? Participating in combat is defined as deployment as part of a national military in what would generally be considered a war zone (absent evidence of non- participation), deployment/general participation in a battle, or affirmative evidence of direct combat. Note that this does not require affirmative evidence of a given person firing a weapon.
1. No
2. Yes

2.5.8.5 Military: Combat (War) (combat_war)

Long tag: hdata_fomin_combat_war

Original tag: Combat_war

Dataset citation: Bäck et al. (2021)

Description:

Free text the war that FM was in combat in, if multiple use commas. If mili_combat=1, enter 9999.

2.5.8.6 Military: Injury (mili_injury)

Long tag: hdata_fomin_mili_injury

Original tag: Mili_injury

Dataset citation: Bäck et al. (2021)

Description:

If the FM had combat experience, was the FM injured? If V45=1, enter 9999.
1. No
2. Yes

2.5.8.7 Military: Rebel (rebel)

Long tag: hdata_fomin_rebel

Original tag: Rebel

Dataset citation: Bäck et al. (2021)

Description:

Did the FM have experience of rebel organisations before (each) appointment? Rebel is defined as those that participate in actions designed to fight against the state or overturn the state. Thus, someone that participates in a coup counts as a rebel even if the coup is non-violent, because coup plotters take on the risk of violence (and the risks if they are caught) in attempting a coup. Someone is also a rebel if they are fighting against the ruling government in a territory, even if they believe that ruling government is illegitimate. Thus, the category of rebel includes both future autocratic leaders such as Mao Zedong as well as future democratic leaders such as Charles de Gaulle.
1. No
2. Yes

2.5.9 Political Background

Variables providing information on the political background of the foreign minister.

2.5.9.1 Politics: Party (party)

Long tag: hdata_fomin_party

Original tag: party

Dataset citation: Bäck et al. (2021)

Description:
  Was the FM acting as member of a party?
  1. No
  2. Yes

2.5.9.2 Politics: Party Name (party_name)

Long tag: hdata_fomin_party_name

Original tag: party_name

Dataset citation: Bäck et al. (2021)

Description:
  Free text the name of the political party. In English where available, otherwise in original language. If party=1, enter 9999.

2.5.9.3 Politics: Party Lead (party_lead)

Long tag: hdata_fomin_party_lead

Original tag: party_lead

Dataset citation: Bäck et al. (2021)

Description:
  Did the FM hold a leading position or special mandate on a national level within a political party?
  1. No
  2. Yes

2.5.9.4 Politics: Parliament (parliament_bg)

Long tag: hdata_fomin_parliament_bg

Original tag: parliament_bg

Dataset citation: Bäck et al. (2021)

Description:
  Did the FM have any experience from national legislature or elected bodies prior to appointment?
  1. No
  2. Yes

2.5.9.5 Politics: Former minister (former_minister)

Long tag: hdata_fomin_former_minister

Original tag: Former_minister

Dataset citation: Bäck et al. (2021)

Description:
  Was the FM a cabinet minister before appointment as FM?
  1. No
  2. Yes

2.5.9.6 Politics: Former minister post (former_min_post)

Long tag: hdata_fomin_former_min_post

Original tag: Former_min_post

Dataset citation: Bäck et al. (2021)

Description:
   If the FM was minister before appointment, free text the name of the post. If several, free text the most recent one. If parliament_bg=1, enter 9999.

2.5.9.7   Politics: Former government office (former_gov_office)

Long tag: hdata_fomin_former_gov_office

Original tag: Former_gov_office

Dataset citation: Bäck et al. (2021)

Description:
   Did the FM hold a high government office (not in cabinet) before appointment?
   1. No
   2. Yes

2.5.9.8   Politics: Former government office name (former_gov_post)

Long tag: hdata_fomin_former_gov_post

Original tag: Former_gov_post

Dataset citation: Bäck et al. (2021)

Description:
   If the FM held a high government office before appointment, free text the name of the office. If several, free text the most recent one. If former_minister=1, enter 9999.

2.5.9.9   Politics: Interest organizations (org_connections)

Long tag: hdata_fomin_org_connections

Original tag: org_connections

Dataset citation: Bäck et al. (2021)

Description:
   Did the FM hold office within any interest organisations other than political parties (such as trade unions or business organisations) prior to appointment?
   1. No
   2. Yes

2.5.9.10   Politics: Interest organizations name (organisation)

Long tag: hdata_fomin_organisation

Original tag: organisation

Dataset citation: Bäck et al. (2021)

Description:
   With which organisation was the FM affiliated? If there are several connections, enter the most important (visible or long-lasting) one. Free text the name or purpose of organization. If org_connections=1, enter 9999.

2.5.10   Dataset Version

The version of the dataset.

2.5.10.1   Dataset Version (version)

Long tag: hdata_fomin_version

Original tag: version

Dataset citation: Bäck et al. (2021)

Description:
   Current version

2.6   H-DATA Information Capacity Dataset

Dataset tag: hdata_infocap

Output Unit: H-DATA Country-Year, i.e., data is collected per country and year. That means each row in the dataset can be identified by one country in combination with a year, using the columns cname and year. Alternatively, each row in the dataset can be identified by a combination of ccodecow and year.

Description: Thomas Brambor, Agustín Goenaga, Johannes Lindvall and Jan Teorell created The Information Capacity Dataset for their article "The Lay of the Land: Information Capacity and the Modern State". The Information Capacity Dataset offers numerical data on five institutions and policies that modern states use to collect information about their populations and territories: (1) the regular implementation of a reliable census, (2) the regular release of statistical yearbooks, the operation of (3) civil and (4) population registers, and (5) the establishment of a government agency tasked with processing statistical information. The dataset also includes an overall index of "information capacity" for 85 polities from 1750 to 2015.

Dataset citation: When using this data, please cite the following paper:

Brambor, Thomas, Agustín Goenaga, Johannes Lindvall and Jan Teorell. 2019. "The Lay of the Land: Information Capacity and the State." Forthcoming in *Comparative Political Studies*.

License: CC-BY-SA 4.0 International
`https://creativecommons.org/licenses/by-sa/4.0/legalcode`

More detailed information on the dataset can be found at the following web page:
`https://www.su.se/english/research/research-projects/h-data/datasets-1.610144`

2.6.1   Identifier Variables

Identification variables in the dataset.

2.6.1.1   Country name (cname)
Long tag: hdata_infocap_cname
Original tag: cname
Dataset citation: Brambor et al. (2020)
Description:
    Country name

2.6.1.2   Year (year)
Long tag: hdata_infocap_year
Original tag: year
Dataset citation: Brambor et al. (2020)
Description:
    Year

2.6.1.3   Correlates of War Project Country Code (ccodecow)
Long tag: hdata_infocap_ccodecow
Original tag: ccodecow
Dataset citation: Brambor et al. (2020)
Description:
    Numeric code for the country from the Correlates of War Project

2.6.1.4 V-Dem Project Country Code (vdemcode)

Long tag: hdata_infocap_vdemcode

Original tag: VDemcode

Dataset citation: Brambor et al. (2020)

Description:
  Numeric code for the country from the Varieties of Democracy Project.

2.6.2 Institutions and Policies

Variables in this section give information on institutions used to collect information about their populations and territories. These include yearbooks, census, civil and population registers, as well as the processing of statistical information by a government agency.

2.6.2.1 Statistical Agency Exists (statagency)

Long tag: hdata_infocap_statagency

Original tag: statagency

Dataset citation: Brambor et al. (2020)

Description:
  Dummy variable that takes a value 1 if the country has a statistical agency operating that year, and 0 otherwise. A statistical agency is an official government organization that is tasked with gathering numerical information on a variety of topics. We assign the value 0 to this variable when these tasks were absorbed by several government offices that per-formed other functions besides collecting and maintaining numerical information.

2.6.2.2 Census Taken (census)

Long tag: hdata_infocap_census

Original tag: census

Dataset citation: Brambor et al. (2020)

Description:
  Dummy variable that takes the value 1 if the country ran a census that year and 0 other-wise. A census is a population count that meets three requirements: (1) universality (it seeks to cover the entire population); (2) defined territory (it counts the population within a well-defined territory); and (3) simultaneity (each person is counted within a discrete time window).

2.6.2.3 First Modern Census (Ind.) (first_moderncensus)

Long tag: hdata_infocap_first_moderncensus

Original tag: first_moderncensus

Dataset citation: Brambor et al. (2020)

Description:
  Dummy variable that takes the value 1 the year a country first implemented a modern census (it is missing otherwise). A modern census meets all three requirements from the census variable plus a fourth one — (4) individual enumeration, which means that the census produces tabular data sets with disaggregated information for each individual.

2.6.2.4 Register-Based Census (register_based_census)

Long tag: hdata_infocap_register_based_census

Original tag: register_based_census

Dataset citation: Brambor et al. (2020)

Description:
  Dummy variable that takes the value 1 if the country runs a census that is based on a continuously updated population register (as opposed to an enumeration carried out dur-ing a discrete period of time) and 0 otherwise.

2.6.2.5 Yearbook Published (yearbook_published)

Long tag: hdata_infocap_yearbook_published

Original tag: yearbook_published

Dataset citation: Brambor et al. (2020)

Description:

Dummy variable that takes the value 1 if the country published a statistical yearbook that year and 0 otherwise. A statistical yearbook is an annual publication of the state that contains numerical data about the country regarding several of the following categories: physical environment, demography, economic affairs, political affairs, and social and cul-tural affairs.

2.6.2.6 Year covered in Yearbook (yearbook_coverage)

Long tag: hdata_infocap_yearbook_coverage

Original tag: yearbook_coverage

Dataset citation: Brambor et al. (2020)

Description:

Dummy variable that takes the value 1 if the country published a statistical yearbook with information for that year (typically, but not necessarily, in the yearbook published the following year) and 0 otherwise.

2.6.2.7 Birth Register is Being Maintained (civreg)

Long tag: hdata_infocap_civreg

Original tag: civreg

Dataset citation: Brambor et al. (2020)

Description:

Dummy variable that takes the value 1 if the country maintained a birth register that year and 0 otherwise. A civil register is a government agency that records key events in the life trajectories of individuals (such as births, marriages, and deaths). We include parish regis-ters when they were administered by official state churches (as in the Nordic countries) but not when the Church was not part of the state (as in the Catholic world).

2.6.2.8 Central Population Register is Being Maintained (popreg)

Long tag: hdata_infocap_popreg

Original tag: popreg

Dataset citation: Brambor et al. (2020)

Description:

Dummy variable that takes the value 1 if the country maintained a population register that year and 0 otherwise. Population registers are centralized state agencies that keep de-tailed and permanently updated records of every person in the territory.

2.6.2.9 Statistical Agency (Ind.) (statag_intro)

Long tag: hdata_infocap_statag_intro

Original tag: statag_intro

Dataset citation: Brambor et al. (2020)

Description:

This variable takes the value 1 for every year after the first adoption of a statistical agen-cy, even if it was later abolished (which only occurred in France, Germany and Denmark) and 0 for all previous years.

2.6.2.10 Graded Census Ability (Depreciated) (censusgraded_ability)

Long tag: hdata_infocap_censusgraded_ability

Original tag: censusgraded_ability

Dataset citation: Brambor et al. (2020)

Description:

A graded index of the ability of the country to carry out censuses consistently with dif-ferent degrees of difficulty. The index takes the value 30 if the country ran a census based on a continuously updated population register (based on register_based_census), 20 if it ran a modern (enumeration-based) census (censuses implemented after first_moderncensus takes the value 1 but that were not based on population registers), and 10 if it ran a non-modern census. The index is depreciated by 1 point for every consecutive year without a census, starting 10 years after the previous census. Therefore, the index takes a value of 0 after 20 years without any census, after 30 years from the last modern census, and after 40 years from the last register-based census.

2.6.2.11 Yearbook Coverage Ability (Depreciated) (ybcov_ability)

Long tag: hdata_infocap_ybcov_ability

Original tag: ybcov_ability

Dataset citation: Brambor et al. (2020)

Description:

A graded index of the ability of the country to publish a statistical yearbook over a ten-year window. The index takes the value 10 when the country published a statistical year-book covering that year (based on yearbook_coverage) and is depreciated by 1 point for each consecutive year without a statistical yearbook. Therefore, the index takes the value 0 after 10 years without statistical yearbooks published.

2.6.2.12 Predicted Information Capacity from IRT (infcap_irt)

Long tag: hdata_infocap_infcap_irt

Original tag: infcap_irt

Dataset citation: Brambor et al. (2020)

Description:

An aggregate index of information capacity. It is based on a hybrid two-parameter and graded Item Response Model (IRT) that is based on five component indicators — when the country first established a statistical agency (statag_intro), whether the country had in place a civil register (civreg) and a population register (popreg), and the graded indexes of census ability (censusgraded_ability) and yearbook ability (ybcov_ability).

2.6.2.13 Predicted Information Capacity from PCA (infcap_pca)

Long tag: hdata_infocap_infcap_pca

Original tag: infcap_pca

Dataset citation: Brambor et al. (2020)

Description:

An aggregate index of information capacity based on a Principal Components Analysis (PCA) model, using the same component variables as infcap_irt.

3 Bibliography

Brambor, T., Goenaga, A., Lindvall, J. & Teorell, J. (2020), 'The lay of the land: Information capacity and the state', *Comparative Political Studies* **53**(2), 175–213.

Bäck, H., Teorell, J., Von Hagen-Jamar, A. & Quiroz Flores, A. (2021), 'War, performance, and the survival of foreign ministers', *Foreign Policy Analysis* **17**(2).

Noonan, J. & Teorell, J. (2023), 'Historical Conflict Dataset (HCD) Dataset v1'.
URL: *https://www.su.se/english/research/research-projects/h-data*

Teorell, J. (2022), 'Rules of recognition? explaining diplomatic representation since the congress of vienna', *Cooperation and Conflict* **58**(2).

Teorell, J., Hellström, J. & Noonan, J. (2023), 'Historical Cabinets (H-CAB) Dataset v1'.
URL: *https://www.su.se/english/research/research-projects/h-data*