

# STANCE

State-Making and the Origins of Global Order  
in the Long Nineteenth Century and Beyond

Information Capacity Dataset Codebook

Codebooks, 2019  
STANCE, Lund University



**LUND**  
UNIVERSITY

STANCE is a six-year research program at the Department of Political Science at Lund University, Sweden. The program, consisting of several separate but connected research projects, aims to answer the question of how state-making and the international system co-evolved in the long 19<sup>th</sup> century (1789-1914) and beyond. The program is constructed around three research themes: (1) How did the different dimensions of state-making evolve? What actors and organized interests supported or put up resistance to these processes?; (2) How were these dimensions of state-making affected by geopolitical competition, warfare and the diffusion of novel political technologies?; and (3) What were the consequences for the international system, both with respect to the type of state that emerged and what entities were granted membership in the state system? The program aims to bridge the gaps between comparative politics and IR, as well as those between the study of political thought and positive empirical political science. The research has been made possible by the Bank of Sweden Tercentenary Foundation (Riksbankens Jubileumsfond). Visit the research program's website at [www.stanceatlund.org](http://www.stanceatlund.org)

Please address comments and/or queries for information to:

*Email address:* info@stanceatlund.org

*Mailing address:* STANCE

Department of Political Science

Lund University

Box 52, SE-221 00 Lund, Sweden

STANCE datasets are available at [www.stanceatlund.org](http://www.stanceatlund.org)

COPYRIGHT © 2019 by authors. All rights reserved.

# Information Capacity Dataset Codebook

---

**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*.

**Country Sample:** The dataset covers the sample of countries that is covered by the Historical Varieties of Democracy project, with the exception of Luxembourg, Oman, Kuwait, Singapore and Zanzibar, which were added to Historical V-Dem after the completion of our data collection on information capacity, and with the addition of the Democratic Republic of Vietnam since 1975, for a total of 86 polities. The following countries and territories are included:

Afghanistan, Argentina, Australia, Austria, Baden, Bavaria, Belgium, Bolivia, Brazil, Brunswick, Bukhara/Uzbekistan, Bulgaria, Canada, Chile, China, Colombia, Costa Rica, Cuba, Denmark, Dominican Republic, Ecuador, Egypt, El Salvador, Ethiopia, Finland, France, Prussia/Germany, Greece, Guatemala, Haiti, Hamburg, Hanover, Hesse Electoral, Hesse Grand Ducal, Honduras, Hungary, India, Indonesia, Iran, Italy, Japan, Korea/South Korea, Liberia, Madagascar, Mecklenburg-Schwerin, Mexico, Modena, Montenegro, Morocco, Myanmar, Nassau, Nejd/Saudi Arabia, Nepal, Netherlands, New Zealand, Nicaragua, Norway, Oldenburg, Papal States, Paraguay, Parma, Peru, Poland, Portugal, Romania, Russia, Saxe-Weimar-Eisenach, Saxony, Serbia/Yugoslavia, Spain, Sweden, Switzerland, Thailand, Tripolitania/Libya, Tunisia, Turkey, Tuscany, Two Sicilies, United Kingdom, United States, Uruguay, Venezuela, South Vietnam/Vietnam, Wuerttemberg, and Yemen.

The main rule is that these polities are coded, regardless of whether they are formally sovereign or not, as soon after 1789 as they can be designated “effective governance units” (e.g., having some centralized power and a territory that at least resembles the country’s territory after

independence). This implies that some colonies, suzerainties and other non-fully sovereign units are included. The following ten countries are included since 1750, due to the fact that they held a census or established a civil register already prior to 1789: Cuba, Denmark (as part of the Kingdom of Denmark-Norway), (the Captaincy-General of) Guatemala, Norway (as part of the Kingdom of Denmark-Norway), Piedmont-Sardinia, Portugal, Spain, Sweden, United Kingdom, and United States.

**Time Period:** The years 1750–2015 are included.

### **Variables:**

***cname:*** Country name

***year:*** Year

***ccodecow:*** Numeric code for the country from the Correlates of War Project

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

***statagency:*** 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 performed other functions besides collecting and maintaining numerical information.

***census:*** Dummy variable that takes the value 1 if the country ran a census that year and 0 otherwise. 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).

***first\_moderncensus:*** 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.

*register\_based\_census:* 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 during a discrete period of time) and 0 otherwise.

*yearbook\_published:* 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 cultural affairs.

*yearbook\_coverage:* 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.

*civreg:* 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 registers 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).

*popreg:* 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 detailed and permanently updated records of every person in the territory.

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

***censusgraded\_ability***: A graded index of the ability of the country to carry out census-es consistently with different 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 implement-ed 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.

***ybcov\_ability***: 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 yearbook 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.

***infcap\_irt***: 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*).

***infcap\_pca***: An aggregate index of information capacity based on a Principal Components Analysis (PCA) model, using the same component variables as *infcap\_irt*.