



Quickstart Guide

V 9 - April 2019

Contents

1	Introduction	3
1.1	V-Dem	3
1.2	Funders	3
1.3	Disclaimer	3
1.4	Suggested Citation	3
2	V-Dem Datasets	5
2.1	Country-Year: V-Dem Core	5
2.2	Country-Year: V-Dem Full+Others	5
2.3	Country-Date: V-Dem	5
2.4	Coder-Level-Dataset	5
2.5	Note: Aggregation from Country-Date to Country-Year	5
3	General Information	6
3.1	Structure of Aggregation	6
3.2	Countries	8
3.3	Identifier Variables in the V-Dem Datasets	10
3.3.1	Country Name (country_name)	10
3.3.2	Time-Specific Country Name (histname)	10
3.3.3	V-Dem Country ID (country_id)	10
3.3.4	Country Name Abbreviation (country_text_id)	10
3.3.5	Year (year)	10
3.3.6	Historical Date (historical_date)	10
3.3.7	Start of Coding Period (codingstart)	10
3.3.8	Contemporary Start of Coding Period (codingstart_contemp)	10
3.3.9	Historical Start of Coding Period (codingstart_hist)	11
3.3.10	Gap in Coding Period Starts (gapstart)	11
3.3.11	Gap in Coding Period Ends (gapend)	11
3.3.12	End of Coding Period (codingend)	11
3.3.13	Historical End of Coding Period (codingend_contemp)	11
3.3.14	Historical End of Coding Period (codingend_hist)	11
3.3.15	V-Dem Project (project)	11
3.3.16	Historical V-Dem coding (historical)	11
3.3.17	COW Code (COWcode)	11
3.4	Variable Types	12
3.5	Variable Versions and Suffixes	12
3.6	Versions of the V-Dem indicators	14
4	V-Dem High-Level Democracy Indices	15
4.1	Electoral democracy index (D) (v2x_polyarchy)	15
4.2	Liberal democracy index (D) (v2x_libdem)	15
4.3	Participatory democracy index (D) (v2x_partipdem)	16
4.4	Deliberative democracy index (D) (v2x_delibdem)	16
4.5	Egalitarian democracy index (D) (v2x_egaldem)	16
5	Reference Documents	18
5.1	Codebook	18
5.2	Methodology	18

5.3	Country Units	18
5.4	Structure of V-Dem Indices, Components, and Indicators	18
5.5	Organization and Management	18

1 Introduction

1.1 V-Dem

Varieties of Democracy (V-Dem) produces the largest global dataset on democracy with some 27 million data points for 202 countries from 1789 to 2018. Involving over 3,000 scholars and other country experts, V-Dem measures hundreds of different attributes of democracy. V-Dem enables new ways to study the nature, causes, and consequences of democracy embracing its multiple meanings.

For questions or inquiries, do not hesitate to contact us at: contact@v-dem.net

1.2 Funders

We are very grateful for our funders' support over the years, which has made this venture possible. To learn more about our funders, please visit: <https://www.v-dem.net/en/v-dem-institute/funders/>

1.3 Disclaimer

This quickstart guide is largely an extraction of text from the codebook, but is compiled as a means for a first quick look at the V-dem dataset. For references, detailed explanations, cautionary notes, and information on all variables, please view the codebook and our other reference documents.

1.4 Suggested Citation

Nota bene: If a variable drawn from the V-Dem dataset plays an important role in your project (published or unpublished), please use all the suggested citations below:

- **V-Dem Dataset:**

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, David Altman, Michael Bernhard, M. Steven Fish, Adam Glynn, Allen Hicken, Anna Lührmann, Kyle L. Marquardt, Kelly McMann, Pamela Paxton, Daniel Pemstein, Brigitte Seim, Rachel Sigman, Svend-Erik Skaaning, Jeffrey Staton, Steven Wilson, Agnes Cornell, Lisa Gastaldi, Haakon Gjerløw, Nina Ilchenko, Joshua Krusell, Laura Maxwell, Valeriya Mechkova, Juraj Medzihorsky, Josefine Pernes, Johannes von Römer, Natalia Stepanova, Aksel Sundström, Eitan Tzelgov, Yi-ting Wang, Tore Wig, and Daniel Ziblatt. 2019. "V-Dem [Country-Year/Country-Date] Dataset v9" Varieties of Democracy (V-Dem) Project. <https://doi.org/10.23696/vdemcy19>

and:

Pemstein, Daniel, Kyle L. Marquardt, Eitan Tzelgov, Yi-ting Wang, Juraj Medzihorsky, Joshua Krusell, Farhad Miri, and Johannes von Römer. 2019. "The V-Dem Measurement Model: Latent Variable Analysis for Cross-National and Cross-Temporal Expert-Coded Data". V-Dem Working Paper No. 21. 4th edition. University of Gothenburg: Varieties of Democracy Institute.

- **V-Dem Codebook:**

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, David Altman, Michael Bernhard, M. Steven Fish, Adam Glynn, Allen Hicken, Anna Lührmann, Kyle L. Marquardt, Kelly McMann, Pamela Paxton, Daniel Pemstein, Brigitte Seim, Rachel Sigman, Svend-Erik Skaaning, Jeffrey Staton, Agnes Cornell, Lisa Gastaldi, Haakon Gjerløw, Valeriya Mechkov, Johannes von Römer, Aksel Sundström, Eitan Tzelgov, Luca Uberti, Yi-ting Wang, Tore Wig, and Daniel Ziblatt. 2019. "V-Dem Codebook v9" Varieties of Democracy (V-Dem) Project.

- **V-Dem Methodology:**

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, Kyle

L. Marquardt, Juraj Medzihorsky, Daniel Pemstein, Josefine Pernes, Johannes von Römer, Natalia Stepanova, Eitan Tzelgov, Yi-ting Wang, and Steven Wilson. 2019. "V-Dem Methodology v9" Varieties of Democracy (V-Dem) Project.

- **V-Dem Country Coding Units:**

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, Vlad Ciobanu, and Lisa Gastaldi. 2019. "V-Dem Country Coding Units v9" Varieties of Democracy (V-Dem) Project.

- **V-Dem Organization and Management:**

Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, Lisa Gastaldi, Nina Ilchenko, Josefine Pernes, Johannes von Römer, and Natalia Stepanova. 2019. "V-Dem Organization and Management v9" Varieties of Democracy (V-Dem) Project.

2 V-Dem Datasets

2.1 Country-Year: V-Dem Core

The five high-level V-Dem Democracy Indices, 81 indices, and the indicators constituting them.

2.2 Country-Year: V-Dem Full+Others

Includes all 450+ V-Dem indicators and indices + 59 other indicators from other data sources.

2.3 Country-Date: V-Dem

Includes all 450+ V-Dem indicators and indices.

2.4 Coder-Level-Dataset

The Coder-Level-Dataset includes all multiple-choice and multiple-selection questions coded by country experts. This dataset also includes coder reliability scores from the Bayesian item response theory measurement model (IRT-model).

2.5 Note: Aggregation from Country-Date to Country-Year

C-variables, ratio/percentage variables, and High-Level/Mid-Level Democracy indices are aggregated from the country-date level to the country-year level by the day-weighted mean. Ordinal A-variables and A*-variables are aggregated by taking the last observation in the year with one exception: event-specific dichotomies or event-specific ordinal variables, which mostly concerns elections or election related data, are aggregated by max (meaning the highest observed value for a given year is retained) to reflect that an "event" of the coded type occurred within the year.

3 General Information

3.1 Structure of Aggregation

The V-Dem conceptual scheme recognizes several levels of aggregation. The table below shows the structure of aggregation for one of our high-level indices: The V-Dem Electoral Democracy Index. This index consists of five sub-components (each of these sub-components being indices themselves built from a number of indicators) that together capture Dahl's seven institutions of polyarchy: freedom of association, suffrage, clean elections, elected executive, and freedom of expression and alternative sources of information.

Please see Appendix A in our codebook (version 9) for a table with a complete hierarchy of our democracy indices, democracy component indices, democracy sub- component indices, and indicators, as well as the hierarchy of related concept indices.

Democracy Index Name	Mid-Level Democracy and Governance Index Name	Lower-Level Democracy and Governance Index Name	Indicator Name	Tag	Uniqueness Score*
<i>Electoral democracy index</i>				v2x_polyarchy	
	Additive polyarchy index			v2x_api	
	Multiplicative polyarchy index			v2x_mpi	
		Freedom of expression and alternative sources of information index		v2x_freexp_altinf	
		Government censorship effort—Media		v2mecenefm	0.2753
		Harassment of journalists		v2meharjrn	0.3169
		Media self-censorship		v2meslfcen	0.3275
		Media bias		v2mebias	0.2648
		Print/broadcast media perspectives		v2merange	0.2751
		Print/broadcast media critical		v2mecrit	0.2518
		Freedom of discussion for men		v2cldiscm	0.2696
		Freedom of discussion for women		v2cldiscw	0.2696
		Freedom of academic and cultural expression		v2clacfree	0.3403
		Freedom of association index (thick)		v2x_frassoc_thick	
		Party ban		v2psparban	0.3272
		Barriers to parties		v2psbars	0.2497
		Opposition parties autonomy		v2psoppaut	0.2511
		Elections multiparty		v2elmulpar	0.401
		CSO entry and exit		v2cseeorgs	0.3591
		CSO repression		v2csreprss	0.3984
		Share of population with suffrage		v2x_suffr	
		Percent of population with suffrage		v2elsuffrage	
		Clean elections index		v2xel_frefair	
		EMB autonomy		v2elembaut	0.3889
		EMB capacity		v2elembcap	0.4854
		Election voter registry		v2elrgstry	0.4764
		Election vote buying		v2elvotbuy	0.6131
		Election other voting irregularities		v2elirreg	0.3687
		Election government intimidation		v2elintim	0.393
		Election other electoral violence		v2elpeace	0.6538
		Election free and fair		v2elfrfair	0.3272
		Elected officials index		v2x_elecoff	
		Legislature bicameral		v2lgbicam	
		Lower chamber elected		v2lgello	
		Upper chamber elected		v2lgelecup	
		Percentage of indirectly elected legislators lower chamber		v2lginello	
		Percentage of indirectly elected legislators upper chamber		v2lginelup	
		HOS appointment in practice		v2expathhs	
		HOG appointment in practice		v2expathhg	
		HOS selection by legislature in practice		v2exaphos	
		HOG selection by legislature in practice		v2exaphogp	
		HOS appoints cabinet in practice		v2exdfcbhs	
		HOG appoints cabinet in practice		v2exdjcbhg	
		HOS dismisses ministers in practice		v2exdfdmhs	
		HOG dismisses ministers in practice		v2exdfdshg	
		HOS = HOG?		v2exhoshog	
		Chief executive appointment by upper chamber		v2exapup	
		Chief executive appointment by upper chamber implicit approval		v2exapupap	
		HOS = HOG?		v2exhoshog	
		Chief executive appointment by upper chamber		v2exapup	
		Chief executive appointment by upper chamber implicit approval		v2exapupap	

*Unmodeled Variance. Uniqueness is the variance that is 'unique' to the variable and not shared with other variables.

3.2 Countries

The following table contains all country units (and their year coverage) that are included in the V-Dem Dataset. Some countries are coded prior to independence, and some have gaps in their coding periods. For a more detailed description of the country units and their year coverage please consult the V-Dem *Country Coding Units* document.

Name	ID	Coverage	Name	ID	Coverage
Afghanistan	36	1789–2018	Egypt	13	1789–2018
Albania	12	1912–2018	El Salvador	22	1838–2018
Algeria	103	1900–2018	Equatorial Guinea	160	1900–2018
Angola	104	1900–2018	Eritrea	115	1900–2018
Argentina	37	1789–2018	Estonia	161	1918–2018
Armenia	105	1990–2018	Ethiopia	38	1789–2018
Australia	67	1789–2018	Fiji	162	1900–2018
Austria	144	1789–2018	Finland	163	1809–2018
Azerbaijan	106	1990–2018	France	76	1789–2018
Baden	349	1789–1871	Gabon	116	1910–2018
Bahrain	146	1900–2018	Georgia	118	1990–2018
Bangladesh	24	1971–2018	German Democratic Republic	137	1949–1990
Barbados	147	1900–2018	Germany	77	1789–2018
Bavaria	350	1789–1871	Ghana	7	1902–2018
Belarus	107	1990–2018	Greece	164	1822–2018
Belgium	148	1789–2018	Guatemala	78	1789–2018
Benin	52	1900–2018	Guinea	63	1900–2018
Bhutan	53	1900–2018	Guinea-Bissau	119	1900–2018
Bolivia	25	1825–2018	Guyana	166	1900–2018
Bosnia and Herzegovina	150	1992–2018	Haiti	26	1789–2018
Botswana	68	1900–2018	Hamburg	362	1789–1867
Brazil	19	1789–2018	Hanover	357	1789–1866
Brunswick	363	1789–1867	Hesse-Darmstadt	359	1789–1866
Bulgaria	152	1878–2018	Hesse-Kassel	358	1789–1866
Burkina Faso	54	1919–2018	Honduras	27	1838–2018
Burma/Myanmar	10	1789–2018	Hong Kong	167	1900–2018
Burundi	69	1916–2018	Hungary	210	1789–2018
Cambodia	55	1900–2018	Iceland	168	1900–2018
Cameroon	108	1961–2018	India	39	1789–2018
Canada	66	1841–2018	Indonesia	56	1800–2018
Cape Verde	70	1900–2018	Iran	79	1789–2018
Central African Republic	71	1920–2018	Iraq	80	1920–2018
Chad	109	1920–2018	Ireland	81	1919–2018
Chile	72	1789–2018	Israel	169	1948–2018
China	110	1789–2018	Italy	82	1861–2018
Colombia	15	1789–2018	Ivory Coast	64	1900–2018
Comoros	153	1900–2018	Jamaica	120	1900–2018
Costa Rica	73	1838–2018	Japan	9	1789–2018
Croatia	154	1941–2018	Jordan	83	1922–2018
Cuba	155	1789–2018	Kazakhstan	121	1990–2018
Cyprus	156	1900–2018	Kenya	40	1900–2018
Czech Republic	157	1918–2018	Kosovo	43	1999–2018
Democratic Republic of the Congo	111	1900–2018	Kuwait	171	1789–2018
Denmark	158	1789–2018	Kyrgyzstan	122	1990–2018
Djibouti	113	1900–2018	Laos	123	1900–2018
Dominican Republic	114	1789–2018	Latvia	84	1920–2018
Ecuador	75	1830–2018	Lebanon	44	1918–2018

Name	ID	Coverage	Name	ID	Coverage
Lesotho	85	1900–2018	Saudi Arabia	197	1789–2018
Liberia	86	1821–2018	Saxe-Weimar-Eisenach	365	1809–1867
Libya	124	1789–2018	Saxony	353	1789–1867
Lithuania	173	1918–2018	Senegal	31	1904–2018
Luxembourg	174	1815–2018	Serbia	198	1804–2018
Macedonia	176	1991–2018	Seychelles	199	1903–2018
Madagascar	125	1817–2018	Sierra Leone	95	1900–2018
Malawi	87	1900–2018	Singapore	200	1867–2018
Malaysia	177	1900–2018	Slovakia	201	1939–2018
Maldives	88	1900–2018	Slovenia	202	1989–2018
Mali	28	1900–2018	Solomon Islands	203	1900–2018
Malta	178	1900–2018	Somalia	130	1900–2018
Mauritania	65	1904–2018	Somaliland	139	1900–2018
Mauritius	180	1900–2018	South Africa	8	1900–2018
Mecklenburg-Schwerin	360	1789–1867	South Korea	42	1789–2018
Mexico	3	1789–2018	South Sudan	32	2011–2018
Modena	351	1789–1859	South Yemen	23	1900–1990
Moldova	126	1990–2018	Spain	96	1789–2018
Mongolia	89	1911–2018	Sri Lanka	131	1900–2018
Montenegro	183	1789–2018	Sudan	33	1900–2018
Morocco	90	1789–2018	Suriname	4	1900–2018
Mozambique	57	1900–2018	Swaziland	132	1900–2018
Namibia	127	1900–2018	Sweden	5	1789–2018
Nassau	366	1806–1866	Switzerland	6	1789–2018
Nepal	58	1789–2018	Syria	97	1918–2018
Netherlands	91	1789–2018	Taiwan	48	1900–2018
New Zealand	185	1841–2018	Tajikistan	133	1990–2018
Nicaragua	59	1838–2018	Tanzania	47	1914–2018
Niger	60	1922–2018	Thailand	49	1789–2018
Nigeria	45	1914–2018	The Gambia	117	1900–2018
North Korea	41	1945–2018	Timor-Leste	74	1900–2018
Norway	186	1789–2018	Togo	134	1916–2018
Oldenburg	364	1789–1867	Trinidad and Tobago	135	1900–2018
Oman	187	1789–2018	Tunisia	98	1789–2018
Pakistan	29	1947–2018	Turkey	99	1789–2018
Palestine/British Mandate	209	1918–1948	Turkmenistan	136	1990–2018
Palestine/Gaza	138	1948–2018	Tuscany	354	1789–1861
Palestine/West Bank	128	1948–2018	Two Sicilies	356	1789–1860
Panama	92	1903–2018	Uganda	50	1900–2018
Papal States	361	1789–1870	Ukraine	100	1990–2018
Papua New Guinea	93	1900–2018	United Arab Emirates	207	1971–2018
Paraguay	189	1811–2018	United Kingdom	101	1789–2018
Parma	352	1789–1859	United States of America	20	1789–2018
Peru	30	1789–2018	Uruguay	102	1825–2018
Philippines	46	1900–2018	Uzbekistan	140	1789–2018
Piedmont-Sardinia	373	1789–1861	Vanuatu	206	1906–2018
Poland	17	1789–2018	Venezuela	51	1789–2018
Portugal	21	1789–2018	Vietnam	34	1945–2018
Qatar	94	1900–2018	Württemberg	355	1789–1871
Republic of the Congo	112	1903–2018	Yemen	14	1789–2018
Republic of Vietnam	35	1802–1975	Zambia	61	1911–2018
Romania	190	1789–2018	Zanzibar	236	1856–2018
Russia	11	1789–2018	Zimbabwe	62	1900–2018
Rwanda	129	1916–2018	.	.	.
Sao Tome and Principe	196	1900–2018	Total number of countries	202	

3.3 Identifier Variables in the V-Dem Datasets

3.3.1 Country Name (`country_name`)

Name of coded country. A V-Dem country is a political unit enjoying at least some degree of functional and/or formal sovereignty. For more details on country units consult the V-Dem *Country Coding Units* document.

Response: Text.

3.3.2 Time-Specific Country Name (`histname`)

Time-specific name of coded country. Many countries go by different names in different time-periods, for example due to name changes, changes in territory, colonization, occupation, or independence. This variable contains a brief description of the identity of each polity that comprises a country's history. This variable is based on the V-Dem *Country Coding Units* document.

Response: Text.

3.3.3 V-Dem Country ID (`country_id`)

Unique country ID designated for each country. A list of countries and their corresponding IDs used in the V-Dem dataset can be found in the country table in the codebook, as well as in the V-Dem *Country Coding Units* document.

Response: Numeric.

3.3.4 Country Name Abbreviation (`country_text_id`)

Abbreviated country names.

Response: Text.

3.3.5 Year (`year`)

V-Dem year coded annually from 1789–2018. This variable is included in the V-Dem Country Year as well as Country Date datasets.

Response: Date.

3.3.6 Historical Date (`historical_date`)

This variable is included in the V-Dem Country Date dataset. The default date is December 31st, as in 2018-12-31, referring to the time span from 01-01 to 12-31 in a respective year. Additionally, specific changes, such as the appointment of a Head of State, are coded on the specific date within a certain year. Thus, a code can change within a year, and will be reflected in the 12-31 date.

Response: Date.

3.3.7 Start of Coding Period (`codingstart`)

V-Dem country coding starts in 1789, or from when a country first enjoyed at least some degree of functional and/or formal sovereignty. For detailed information, please see the V-Dem *Country Coding Units* document.

Response: Date.

3.3.8 Contemporary Start of Coding Period (`codingstart_contemp`)

This variable indicates the coding start for the countries coded by Contemporary V-Dem.

Response: Date.

3.3.9 Historical Start of Coding Period (codingstart_hist)

This variable indicates the coding start for the countries coded by Historical V-Dem.

Response: Date.

3.3.10 Gap in Coding Period Starts (gapstart)

Time periods when a country does not fulfill V-Dem's coding period criteria are not coded. The date that indicates the gap start is the last date coded before the gap. For more details about V-Dem country coding periods, please see the V-Dem *Country Coding Units* document.

Response: Date.

3.3.11 Gap in Coding Period Ends (gapend)

The periods of when a country does not fulfill V-Dem's coding period macriteria are not coded. The date that indicates the gap end is the first date coded after the gap. For more details about V-Dem country coding periods, please see the V-Dem *Country Coding Units* document.

Response: Date.

3.3.12 End of Coding Period (codingend)

V-Dem country coding ends in 2018, or from when a country formally stopped enjoying at least some degree of functional and/or formal sovereignty. For detailed information, please see the V-Dem *Country Coding Units* document.

Response: Date.

3.3.13 Historical End of Coding Period (codingend_contemp)

This variable indicates when the coding ends for countries coded by the Contemporary V-Dem project.

Response: Date.

3.3.14 Historical End of Coding Period (codingend_hist)

This variable indicates when the coding ends for countries coded by the Historical V-Dem Project.

Response: Date.

3.3.15 V-Dem Project (project)

This variable indicates which V-Dem project coded that country-year: Contemporary V-Dem (0), Historical V-Dem (1), or both (2).

3.3.16 Historical V-Dem coding (historical)

This variable indicates if the Historical V-Dem project coded a country at any time: No (0), Yes (1).

3.3.17 COW Code (COWcode)

COW country codes according to the Correlates of War Project (2016).

Response: Numeric.

3.4 Variable Types

The V-Dem Codebook divides variables into the following variable types:

- **Type A*: Variables coded by Research Assistants**
This data is based on extant sources and is factual in nature. Country Experts indicate their confidence for this pre-coded data.
- **Type A: Variables coded by Project Managers and Research Assistants**
This data is based on extant sources and is factual in nature.
- **Type B: Variables coded by Country Coordinators or Research Assistants**
The coder is typically a graduate student or recent graduate from the country in question. These variables are factual in nature.
- **Type C: Variables coded by Country Experts**
A Country Expert is typically a scholar or professional with deep knowledge of a country and of a particular political institution. Furthermore, the expert is usually a citizen or resident of the country. Multiple experts (usually 5 or more) code each variable. More information about the Country Experts can be found in the *V-Dem Methodology* document.
- **Type A,C: Variables coded by Country Experts and crosschecked by Research Assistants**
- **Type D: Indices**
Variables composed of type A, B, or C variables. This data may be accomplished by adding a denominator (*e.g.*, per capita), by creating a cumulative scale (Total number of...), or by aggregating larger concepts (*e.g.*, components or indices of democracy).
- **Type E: Non-V-Dem variables**
If we import a variable from another source without doing any original coding, except for perhaps imputing missing data, it is not considered a V-Dem product. These variables are found in the section of the Codebook labeled "Background Factors" or "Other Democracy Indices and Indicators." If, however, we gather data from a number of sources and combine them in a more than purely mechanical fashion (requiring some judgment on our part), we regard this as a V-Dem product and classify it as type A, B, or C. All "E" variables are included in the codebook except those drawn from sources that have more than 30 variables such as Archigos (Goemans et al.), BDM (Bueno de Mesquita et al.), Henisz/POLCON (2000, 2002), Miller (Democratic Pieces), Performance of Nations (Kugler and Tammen), PEI (Norris et al.), PIPE (Przeworski et al.), QoG (Quality of Government). For these variables, we ask users to consult separate codebooks, as listed above.

3.5 Variable Versions and Suffixes

The V-Dem Dataset contains several versions of the variables coded by country experts (type C variables):

- **Model Estimates**
"Model Estimates" — Measurement Model Output:
This version has no special suffix (*e.g.* v2elmulpar). This version of the variables provides country-year (country-date in the alternative dataset) point estimates from the V-Dem measurement model (see Pemstein et al. 2019). The measurement model aggregates the ratings provided by multiple country experts and, taking disagreement and measurement error into account, produces a probability distribution over country-year scores on a standardized interval scale (see the *V-Dem Methodology* document). The point estimates are the median values of these distributions for each country-year. The scale of a measurement model variable is similar to a normal ("Z") score (*e.g.* typically between -5 and 5, with 0 approximately representing the mean for all country-years in the sample) though it does not necessarily follow a normal distribution. For most purposes, these are the preferred versions of the variables for time series regression and other estimation strategies.

"Model Estimates Measure of Uncertainty" — Measurement Model Highest Posterior Density (HPD) Intervals:

This version has the suffixes: "codelow" and "codehigh" (e.g. v2elmulpar_codelow and v2elmulpar_codehigh). These two kinds of variables ["code low" and "code high"] demarcate the interval in which the measurement model places 68 percent of the probability mass for each country-year score, which is approximately equivalent to one standard deviation upper and lower bounds. If the underlying posterior distribution is skewed, the HPDs reflect this with unequal distances between the point estimate and the high and low estimates. We also provide a standard calculation for standard deviation which is marked with the suffix "sd" (e.g., v2elmulpar_sd). The SD might be used to compute the standard frequentist confidence intervals.

- **Original Scale (*_osp)**

"Original Scale" — Linearized Original Scale Posterior Prediction:

This version has the suffix "_osp," (e.g. v2elmulpar_osp). In this version of the variables, we have linearly translated the measurement model point estimates back to the original ordinal scale of each variable (e.g. 0–4 for v2elmulpar_osp) as an interval measure. The decimals in the _osp version roughly indicate the distance between the point estimate from the linearized measurement model posterior prediction and the threshold for reaching the next level on the original ordinal scale. Thus, a _osp value of 1.25 indicates that the median measurement model posterior predicted value was closer to the ordinal value of 1 than 2 on the original scale. Technically, it calculates the sum of the posterior probabilities that the estimate is in a particular category: If a particular country-year-variable has a probability of 90% to be in category "4", a 10% probability of being in category "3", and 0% probability of being in categories "2", "1", and "0", the result is a value of 3.9 ($4 \cdot 0.9 + 3 \cdot 0.1 = 3.6 + 0.3$). Since there is no conventional theoretical justification for linearly mapping ordinal posterior predictions onto an interval scale, these scores should primarily be used for heuristic purposes. Using the "Ordinal Scale" estimates—or incorporating the properties of ordinal probit models into the estimation procedure—is thus preferable to using the _osp estimates in statistical analyses. However, since the _osp version maps onto the coding criteria found in the V-Dem Codebook, and is strongly correlated with the Measurement Model output (typically at .98 or higher), some users may find the _osp version useful in estimating quantities such as marginal effects with a clear substantive interpretation. If a user uses _osp data in statistical analyses it is imperative that she confirm that the results are compatible with estimations using Measurement Model output.

"Original Scale Measure of Uncertainty" — Linearized Original Scale HPD Intervals:

This version has the suffixes - "codelow" and "codehigh" (e.g. v2elmulpar_osp_codelow and v2elmulpar_osp_codehigh). We estimate these quantities in a similar manner as the Measurement Model Highest Posterior Density Intervals. These two variables ["code low" and "code high"] demarcate the interval in which the measurement model places 70 percent of the probability mass for each country-year score, which is approximately equivalent to one standard deviation upper and lower bounds. If the underlying posterior distribution is skewed, the HPDs reflect this with unequal distances between the point estimate and the high and low estimates. We also provide a standard calculation for standard deviation which is marked with the suffix "sd" (e.g., v2elmulpar_sd). The SD might be used to compute the standard frequentist confidence intervals.

- **Ordinal Scale (*_ord)**

"Ordinal Scale" — Measurement Model Estimates of Original Scale Value:

This version has the suffix "_ord" (e.g. v2elmulpar_ord). This method translates the measurement model estimates back to the original ordinal scale of a variable (as represented in the Codebook) after taking coder disagreement and measurement error into account. More precisely, it represents the most likely ordinal value on the original codebook scale into which a country-year would fall, given the average coder's usage of that scale. More specifically, we assign each country-year a value that corresponds to its integerized median ordinal highest posterior probability category over Measurement Model output.

"Ordinal Scale Measure of Uncertainty" — Original Scale Value HPD Intervals:

This version has the suffixes - "codelow" and "codehigh" (e.g. v2elmulpar_ord_codelow and v2elmulpar_ord_codehigh). We estimate these values in a similar manner as the Measurement Model Highest Posterior Density Intervals. These two variables ["code low" and "code high"]

demarcate the interval in which the measurement model places 70 percent of the probability mass for each country-year score, which is approximately equivalent to one standard deviation upper and lower bounds. If the underlying posterior distribution is skewed, the HPDs reflect this with unequal distances between the point estimate and the high and low estimates. We also provide a standard calculation for standard deviation which is marked with the suffix "sd" (*e.g.* v2elmulpar_sd). The SD might be used to compute the standard frequentist confidence intervals.

- **Number of Coders per Country, Variable and Year/Date (*_nr)**

The number of V-Dem Country Experts (regular coders, bridge- and lateral coders) who provided data on country, variable and year. V-Dem's methodology is based on the assumption that we have a minimum of five Country Experts for every single country-variable-year. Sometimes, however, we end up with fewer than five Country Experts. From v7 of the Country-Year, and the Country-Date type datasets, we provide all data we have for full transparency. By providing the number of Country Experts for each country-variable-year/date, we suggest that users primarily base analyses on observations based on five or more coders. We strongly advise against using observations based on three or fewer coders. This concerns all C type variables.
Response: Numeric.

3.6 Versions of the V-Dem indicators

Suffix	Scale	Description	Recommended Use
None	Interval	Original output of the V-Dem measurement model	Regression analysis
_osp	Interval	Linearized transformation of the measurement model output on the original scale	Substantive interpretation of graphs and data
_ord	Ordinal	Most likely ordinal value taking uncertainty estimates into account	Substantive interpretation of graphs and data
_codelow	Interval	One Standard deviation below the point estimate	Evaluating differences over time within units
_codehigh	Interval	One Standard deviation above the point estimate	Evaluating differences over time within units
_sd	Interval	Standard deviation of the interval estimate	Creating confidence intervals based on user needs

4 V-Dem High-Level Democracy Indices

The entries for all V-Dem variables in our dataset are available in our codebook.

4.1 Electoral democracy index (D) (v2x_polyarchy)

Project Manager(s): Jan Teorell

Question: To what extent is the ideal of electoral democracy in its fullest sense achieved?

Clarification: The electoral principle of democracy seeks to embody the core value of making rulers responsive to citizens, achieved through electoral competition for the electorate's approval under circumstances when suffrage is extensive; political and civil society organizations can operate freely; elections are clean and not marred by fraud or systematic irregularities; and elections affect the composition of the chief executive of the country. In between elections, there is freedom of expression and an independent media capable of presenting alternative views on matters of political relevance. In the V-Dem conceptual scheme, electoral democracy is understood as an essential element of any other conception of representative democracy — liberal, participatory, deliberative, egalitarian, or some other.

Scale: Interval, from low to high (0-1).

Source(s): v2x_freexp_altinf v2x_frassoc_thick v2x_suffr v2xel_frefair v2x_elecoff

Data release: 6-9. Release 1-5 used a different, preliminary aggregation formula.

Aggregation: The index is formed by taking the average of, on the one hand, the weighted average of the indices measuring freedom of association thick (v2x_frassoc_thick), clean elections (v2xel_frefair), freedom of expression (v2x_freexp_altinf), elected officials (v2x_elecoff), and suffrage (v2x_suffr) and, on the other, the five-way multiplicative interaction between those indices. This is half way between a straight average and strict multiplication, meaning the average of the two. It is thus a compromise between the two most well known aggregation formulas in the literature, both allowing partial "compensation" in one sub-component for lack of polyarchy in the others, but also punishing countries not strong in one sub-component according to the "weakest link" argument. The aggregation is done at the level of Dahl's sub-components with the one exception of the non-electoral component. The index is aggregated using this formula:

$$\begin{aligned}
 v2x_polyarchy &= .5 * MPI + .5 * API \\
 &= .5 * (v2x_elecoff * v2xel_frefair * v2x_frassoc_thick * \\
 &\quad v2x_suffr * v2x_freexp_altinf) \\
 &\quad + .5 * ((1/8) * v2x_elecoff + (1/4) * v2xel_frefair \\
 &\quad + (1/4) * v2x_frassoc_thick + (1/8) * v2x_suffr \\
 &\quad + (1/4) * v2x_freexp_altinf)
 \end{aligned}$$

Citation: Teorell *et al.* (2016, *V-Dem Working Paper Series* 2016:25); *V-Dem Codebook* (see suggested citation at the top of this document).

4.2 Liberal democracy index (D) (v2x_libdem)

Project Manager(s): Jan Teorell

Question: To what extent is the ideal of liberal democracy achieved?

Clarification: The liberal principle of democracy emphasizes the importance of protecting individual and minority rights against the tyranny of the state and the tyranny of the majority. The liberal model takes a "negative" view of political power insofar as it judges the quality of democracy by the limits placed on government. This is achieved by constitutionally protected civil liberties, strong rule of law, an independent judiciary, and effective checks and balances that, together, limit the exercise of executive power. To make this a measure of liberal democracy, the index also takes the level of electoral democracy into account.

Scale: Interval, from low to high (0-1).

Source(s): v2x_liberal v2x_polyarchy

Data release: 4-9. Release 1, 2, and 3 used a different, preliminary aggregation formula.

Aggregation: The index is aggregated using this formula:

$$v2x_libdem = .25 * v2x_polyarchy^{1.585} + .25 * v2x_liberal + .5 * v2x_polyarchy^{1.585} * v2x_liberal$$

Citation: Coppedge *et al.* (2015, *V-Dem Working Paper Series* 2015:6); *V-Dem Codebook* (see suggested citation at the top of this document).

4.3 Participatory democracy index (D) (v2x_partipdem)

Project Manager(s): Jan Teorell

Question: To what extent is the ideal of participatory democracy achieved?

Clarification: The participatory principle of democracy emphasizes active participation by citizens in all political processes, electoral and non-electoral. It is motivated by uneasiness about a bedrock practice of electoral democracy: delegating authority to representatives. Thus, direct rule by citizens is preferred, wherever practicable. This model of democracy thus takes suffrage for granted, emphasizing engagement in civil society organizations, direct democracy, and subnational elected bodies. To make it a measure of participatory democracy, the index also takes the level of electoral democracy into account.

Scale: Interval, from low to high (0-1).

Source(s): v2x_polyarchy v2x_partip

Data release: 4-9. Release 1-3 used a different, preliminary aggregation formula.

Aggregation: The index is aggregated using this formula:

$$v2x_partipdem = .25 * v2x_polyarchy^{1.585} + .25 * v2x_partip + .5 * v2x_polyarchy^{1.585} * v2x_partip$$

Citation: Coppedge *et al.* (2015, *V-Dem Working Paper Series* 2015:6); *V-Dem Codebook* (see suggested citation at the top of this document).

4.4 Deliberative democracy index (D) (v2x_delibdem)

Project Manager(s): Jan Teorell

Question: To what extent is the ideal of deliberative democracy achieved?

Clarification: The deliberative principle of democracy focuses on the process by which decisions are reached in a polity. A deliberative process is one in which public reasoning focused on the common good motivates political decisions—as contrasted with emotional appeals, solidary attachments, parochial interests, or coercion. According to this principle, democracy requires more than an aggregation of existing preferences. There should also be respectful dialogue at all levels—from preference formation to final decision—among informed and competent participants who are open to persuasion. To make it a measure of not only the deliberative principle but also of democracy, the index also takes the level of electoral democracy into account.

Scale: Interval, from low to high (0-1).

Source(s): v2xdl_delib v2x_polyarchy

Data release: 4-9. Release 1-3 used a different, preliminary aggregation formula.

Aggregation: The index is aggregated using this formula:

$$v2x_delibdem = .25 * v2x_polyarchy^{1.585} + .25 * v2x_delib + .5 * v2x_polyarchy^{1.585} * v2x_delib$$

Citation: Coppedge *et al.* (2015, *V-Dem Working Paper Series* 2015:6); *V-Dem Codebook* (see suggested citation at the top of this document).

4.5 Egalitarian democracy index (D) (v2x_egaldem)

Project Manager(s): Rachel Sigman, Staffan Lindberg

Question: To what extent is the ideal of egalitarian democracy achieved?

Clarification: The egalitarian principle of democracy holds that material and immaterial inequalities inhibit the exercise of formal rights and liberties, and diminish the ability of citizens from all social groups to participate. Egalitarian democracy is achieved when 1 rights and freedoms of individuals are protected equally across all social groups; and 2 resources are distributed equally across all social groups; 3 groups and individuals enjoy equal access to power. To make it a measure of egalitarian democracy, the index also takes the level of electoral democracy into account.

Scale: Interval, from low to high (0-1).

Source(s): v2x_egal v2x_polyarchy

Data release: 5-9. Release 1-4 used a different, preliminary aggregation formula.

Aggregation: The index is aggregated using this formula:

$$v2x_egaldem = .25 * v2x_polyarchy^{1.585} + .25 * v2x_egal + .5 * v2x_polyarchy^{1.585} * v2x_egal$$

Citation: Sigman *et al.* (2015, *V-Dem Working Paper Series* 2015:22); Coppedge *et al.* 2015, *V-Dem Working Paper Series* 2015:6; *V-Dem Codebook* (see suggested citation at the top of this document).

5 Reference Documents

5.1 Codebook

Please note that all variables that V-Dem is compiling are included in the Codebook.

Available at: *V-Dem Codebook v9*

5.2 Methodology

Part I sets forth the V-Dem conceptual scheme. Part II discusses the process of data collection. Part III describes the measurement model along with efforts to identify and correct errors.

Available at: *V-Dem Methodology document v9*

5.3 Country Units

This document lists (a) every country coding unit in the V-Dem database, (b) the years for which we have collect data or plan to collect data (in parentheses next to the entry); (c) the polities that comprise each country's 20th century history (even if falling outside the time-period that we wish to code); and (d) the borders of each country (wherever this might be unclear).

Available at: *V-Dem Country Coding Units document v9*

5.4 Structure of V-Dem Indices, Components, and Indicators

An overview of V-Dem democracy indices, democracy component indices, democracy subcomponent indices, and indicators, as well as the hierarchy of related concept indices.

Available at: *V-Dem Structure of V-Dem Indices, Components, and Indicators document v9*

5.5 Organization and Management

This document includes a brief overview of the Varieties of Democracy (V-Dem) project – the team, the web site, outreach to the international community, funding, progress to date, and sustainability.

Available at: *V-Dem Organization and Management document v9*