Indonesian Election Archive (IEA): Codebook

June 2025

Contents

1 Project description

This codebook describes the structure of the files and the common variable identifiers in the Indonesia Election Archive.

It is the hope of the authors of this project that the archive become a community project. Thus, users are encouraged to share additional cleaned election data that they may have. With the permission of the contributor we will post any cleaned and compatible data that has been used in publication along with the article/book from where this data originates.

2 Bibliographic citation

For the archive (V1) please cite:

Agustyati, K. N., Pratama, H. M., Birnir, J. K., Overos, H., Okthariza, N., Kholidin, I., Ramadhanil, F., & Savirani, A. (2025). Indonesia Election Archive (IEA). Retrieved from osf.io/hjqzd

The data described in this first version (V1) are available on the Open Science Framework (OSF).

These data and future versions of the data as described in the supplementary appendix accompanying the above paper will also be available through:
rumahpemilu.org
johannabirnir.com

3 Credits

For Version 1, Jóhanna K. Birnir, Noory Okthariza, Iqbal Kholidin, Khoirunnisa Agustyati and Heroik Pratama collected and verified the data. Jóhanna K. Birnir, Henry Overos and Noory Okthariza systematized and analyzed the data. Jóhanna K. Birnir and Henry Overos wrote the text with substantive input from Khoirunnisa Agustyati and Heroik Pratama. Jóhanna K. Birnir created the digital archive. Fadli Ramadhanil helped design the Perludem portal of which the election archive is a part. Amalinda Savirani helped design the project. All authors reviewed the manuscript.

4 Data Procedures

The data were all collected from a variety of Indonesian government sources in the public domain. Specifically, in Indonesia, Law 14, 2008 on public infor-

mation disclosure[1] establishes the importance of transparency of government provided information. To further cement the public provision of government information, in 2012, the Indonesian government launched the Open Government Indonesia (OGI) initiative to address issues at the intersection of civil society and government, particularly electoral administration and voter information (ogi.bappenas.go.id). The OGI secretariat, including the KPU, collaborates with civil society organizations - including Perludem - to promote transparency, participation, innovation, accountability, and inclusiveness in governance and public services, including election governance.

Great care has been taken in curating and cleaning these data for the greatest accuracy possible. Once collected, the data were checked for errors both internally and through comparison against other sources. The datasets were then transformed into a common format and merged into data files that contain respectively legal information and election information, for analysis. Users are encouraged to contact the corresponding author Dr. Birnir at jkbirnir@umd.edu with any suggestions and/or concerns.

5 Variables

To maximize the compatibility with other available and similarly organized sources for Indonesian election data, the variable list and descriptions are (with permission)[2] modeled after the Constituency Level Elections Archive (CLEA)[3]. Because CLEA focuses on the lower house of national legislatures only, and includes elections across the world, whereas the Indonesian Election Archive has as an objective to provide election data at multiple levels of elected representation, but for Indonesia only, variables and variable descriptions have been modified, introduced and deleted as necessary.

5.1 Variable list

- Missing data and variables that do not apply to a given election are denoted with NA.

- PDS Publication data source.

- VN Version number

- ID Election Identifier

- YR Election Year

- MN Election Month

- PROV Province Name

---

[1]Available at: https://peraturan.bpk.go.id/Details/39047/uu-no-14-tahun-2008
[2]Personal email Allen Hicken 5-30-25
[3]For more on CLEA see: https://electiondataarchive.org/

- DIST_N District Name

- DIST_C District Code

- PTY_N Party Name

- PTY_C Party Code

- CAN_N Candidate Name

- CAN_G Candidate Gender

- RV Registered Voters

- VV Valid Votes

- IVV Invalid Votes

- CV Candidate Votes

- TPV Total Party Votes

- SEAT Seats Won

5.2   Variable Descriptions

Missing data and variables that do not apply to a given election are denoted with NA.

PDS Publication data source.
This variable accounts for where the data were sourced and are described. It is our hope that the community of scholars who study Indonesian elections adopt this data collection as a communal project. This variable denotes the scholarly publication where successive iterations of the data are introduced. Each publication will be assigned a number in the order that the data are added and listed in this codebook and on the source website. Version 1 is the article introducing the data.

> 1: Agustyati, K. N., Pratama, H. M., Birnir, J. K., Overos, H., Okthariza, N., Kholidin, I., Ramadhanil, F., & Savirani, A. (2025). Indonesia Election Archive (IEA). Retrieved from osf.io/hjqzd

VN Version Number. While the data presented here is a standalone contribution - these data are also a part of what we hope will be a continuing project of Indonesian electoral data. Each time the data are updated the version number will be updated and changes to the data between versions will be recorded in the notes.

ID Election Identifier This variable uniquely identifies each election in the archive.
In V1 these are:

- **PRES_VP** Presidency and Vice Presidency
- **DPD** Dewan Perwakilan Daerah or the upper house Regional Representative Council
- **DPR** Dewan Perwakilan Rakyat or the lower house People's Representative Council

YR Election Year
Year of election. In V1 2014, 2019, 2024.

MN Election Month
Month of election, if available.
January
February
March
April
May
June
July
August
September
October
November
December

Geographic identifier

- **PROV** Province
  First level sub-national administrative unit. This variable is always given because lower level administrative units are nested in upper levels and because in V1 some results are reported at this level.
- **Dist_N** District Name
  Second level sub-national administrative unit. Some provinces have only a single district. Other provinces have multiple districts. This variable is only given if votes are reported at the district level.
- **Dist_C** District Code
  A number assigned to a given district within a province. Some provinces have only a single district. Other provinces have multiple districts. The number is only unique within the province but many districts share numbers across provinces.This variable is only given if votes are reported at the district level.

PTY_N Party Name
Name of a party or electoral alliance. This information is only recorded for DPR elections in V1.

DPD candidates run as independents.

No party affiliations are reported for presidential candidate pairs.

PTY_C Party Code
A unique numeric code is assigned to each party that runs a candidate in any given election. These codes are not unique across election years. This information is only recorded for DPR elections in V1.

DPD candidates run as independents.

No party affiliations are reported for presidential candidate pairs.

CAN_N Candidate Name
Name of a candidate, if available. In Presidential elections the President and Vice Presidential candidates run in pairs. Therefore, the candidate names in elections for the Presidency account for the names of each of the presidential and vice presidential candidate pairs.

CAN_G Candidate Gender
A binary variable designating the candidate as either Male or Female.

RV Number of Registered Voters
The number of registered voters at the administrative level indicated by the geographic identifier of the document (in V1 PROV or DIST).

There are several different accounts for registered voters. The first is DPT (daftar pemilih tetap/permanent voter list). The KPU receives a potential voter list from the from civil registry office and verifies it through a census before every election). The DPTb (daftar pemilih tambahan/added voter list) is a list of the voter who have changed address (and therefore voting station) between elections. The DPK (daftar pemilih khusus/special voter list), in turn, is a list of new voters who were not registered on the initial permanent list but come to the voting station with their id and are registered there on voting day. Thus, the number of registered voters that we use is a sum of all three categories.

In Indonesia only the office of the presidency and some governorships are subject to runoffs if no candidate receives the requisite share of votes. In this version of the data (V1) we only include numbers for the last round of voting.

VV Valid Votes
The total number of valid votes at the administrative level indicated by the geographic identifier of the document (in V1 PROV or DIST).

In Indonesia only the office of the presidency and some governorships are subject to runoffs if no candidate receives the requisite share of votes. In this version of the data (V1) we only include numbers for the last round of voting.

IVV Invalid Votes
The total number of invalid and spoilt votes at the administrative level

indicated by the geographic identifier of the document (in V1 PROV or DIST).

In Indonesia only the office of the presidency and some governorships are subject to runoffs if no candidate receives the requisite share of votes. In this version of the data (V1) we only include numbers for the last round of voting.

CV Candidate Votes
The total number of votes received by the candidate at the administrative level indicated by the geographic identifier of the document (in V1 PROV or DIST).

In Indonesia only the office of the presidency and some governorships are subject to runoffs if no candidate receives the requisite share of votes. In this version of the data (V1) we only include numbers for the last round of voting.

TPV Total Party Votes
The total number of votes received by the party at the administrative level indicated by the geographic identifier of the document (in V1 PROV or DIST). Note that in the elections reported in V1 (2014, 2019 and 2024) Indonesia employed an open list system in elections for the DPD and the DPR. The party vote reported in this variable is the total vote that was allocated to the party and any candidate running for that party.

In elections to the DPD candidates run independently. Therefore, total party votes are coded NA for this election.

SEAT Seats Won
In V1 the number of seats won by a party for either the DPD or the DPR.

6   Notes

- The province of Kalimantan Utara was created in 2012 but elections results were not counted separately from there until 2019.