

User's Guide and Codebook for the ANES 2024 Time Series Voter Validation Supplemental Data

American National Election Studies
The University of Michigan and Stanford University
June 23, 2026

Suggested Citation

ANES. 2026. User's Guide and Codebook for the ANES 2024 Time Series Voter Validation Supplemental Data. Ann Arbor, MI and Palo Alto, CA: the University of Michigan and Stanford University.

Acknowledgments

This study was funded by the National Science Foundation (www.nsf.gov) under grant SES-1835731 to the University of Michigan. Any opinions, findings and conclusions, or recommendations expressed in these materials are those of the ANES staff author(s) and do not necessarily reflect the views of the National Science Foundation.

The Principal Investigators for the 2024 ANES were Nicholas Valentino (University of Michigan), Shanto Iyengar (Stanford University), D. Sunshine Hillygus (Duke University), and Daron Shaw (University of Texas). The directors of Stanford and Michigan operations, respectively, were Matthew DeBell and David Howell.

The 2024 vote validation process was directed by Matthew DeBell. Record linkage was performed by Minako Edgar with support from Raphael Nishimura. Record linkage was performed with probabilistic linkage software (*fastlink*) developed by Ted Enamorado, Ben Fifield, and Kosuke Imai, and clerical review of probabilistic results was performed using software developed in Python by Minako Edgar, Yara Sevilla, and Matthew DeBell. Validation relied on a national voter file compiled by L2, Inc., and licensed by Stanford University Libraries.

This documentation may reprint some material from previous ANES documentation without explicit attribution.

Contact

The ANES website address is www.electionstudies.org

Any questions not answered on the ANES website or by this report may be directed to ANES staff by email at anes@electionstudies.org

Contents

Introduction and Purpose of the Dataset	4
Probabilistic Record Linkage with Fastlink.....	6
Automated Clerical Review	8
ANES Questionnaire Validation Check	9
Selecting the Best Match	11
Contents of the Dataset	12
Codebook.....	16

Introduction and Purpose of the Dataset

This is documentation for the ANES 2024 Time Series Voter Validation Supplemental Data. The data file provides data on voter turnout that were compiled from official voting records by a commercial vendor and were then matched to the ANES 2024 Time Series Study's sample by ANES personnel. The data are provided for purposes of methodological evaluation and for the analysis of voter turnout status. This documentation describes the matching procedures and resulting data.

Several commercial vendors compile voter data for sale for use in campaigns. Stanford University Libraries obtained access to a national voter file from one of these vendors, L2, Inc. ANES personnel used a probabilistic matching process supplemented by deterministic matching evaluations to link ANES respondents and sampled addresses to their corresponding records on this voter file.

Person and Address Records

This file includes voter data of two distinct types: person records and address records. Address records apply to the addresses that were sampled for the ANES 2024 Time Series Study. Person records apply to individuals who were sampled at the sampled addresses. Because individuals were not sampled at every address due to nonresponse, person records do not exist for all sample units. Person records of turnout are intended for comparison to self-reported turnout. Address records of turnout, which indicate the number of voters associated with a sampled address, may be used to analyze nonresponse, household context for sampled persons, or for other purposes.

Person records include voter turnout for the 2024, 2022, and 2020 general elections, and metadata about match quality. Address records include the number of registered voters at the sampled address who voted in the 2024, 2022, and 2020 general elections, and the numbers of registered voters, families, Democrats, independents, and Republicans at each sampled address.

Multiple Person Matches and Matching Methods

ANES respondent records were linked to voter file records using four distinct methods, often producing multiple prospective matches to different voter records.

1. *Probabilistic record linkage with fastlink.* We used the *fastlink*¹ probabilistic record linkage package for the R statistical programming language to identify prospective voter-file matches and assign each a fastlink-calculated posterior match probability score. We also identified one voter file record with the highest fastlink-calculated posterior match probability score across in-state and out-of-state searches; this approximates the record linkage method used by ANES for the 2016 Time Series Study.²
2. *Automated clerical review.* We performed an automated clerical review procedure, using software developed in Python by ANES for this purpose, that compared ANES records to each prospective voter-file match and produced match-quality indicators used to rule out matches that appeared implausible. This review result for the prospective match with the highest fastlink-calculated match score (above) approximates method used by ANES for the

¹ Ted Enamorado, Benjamin Fifield, and Kosuke Imai (2017). *fastLink: Fast Probabilistic Record Linkage with Missing Data*. Version 0.6.1. Available at <https://github.com/kosukeimai/fastLink>

² See Ted Enamorado, Benjamin Fifield, and Kosuke Imai (2018). *User's Guide and Codebook for the ANES 2016 Time Series Voter Validation Supplemental Data*. Available at https://electionstudies.org/wp-content/uploads/2018/03/anes_timeseries_2016voteval_userguidecodebook.pdf

2020 Time Series Study’s “vendor 2” validated turnout data, but used an automated clerical review procedure, in contrast to the manual coder review performed for the 2020 dataset.³

3. *ANES questionnaire validation check.* After the pre-election study, we selected up to five plausible candidate voter-file records for each eligible case and, in the post-election questionnaire, presented these records (name, gender, date of birth, and address) to participants and asked them to indicate whether any record described them, to confirm the selected record if correct, or to provide corrections if not. See the VOTEVAL section of the 2024 post-election questionnaire. If a record was corrected, we searched the voter file again using the corrected information.
4. *Selecting the best match.* After performing the three procedures above, we selected what appeared to be the best match from among what was often more than one possibility.

These procedures are described in more detail in the sections below. The dataset includes 177 variables with results from all of these procedures.

Data Quality and Recommended Usage

ANES makes no representations about the quality of the voter data from its commercial sources. Users should be aware that evaluations of the quality of such data are controversial, in part because such vendors, as commercial political campaign operatives, do not build or document their datasets for research purposes and do not follow norms of scientific transparency. Errors from several sources may affect the accuracy of voter validation data. Sources of errors in these data include data entry and transcription errors when individuals complete voter registration forms and when government agencies enter data from registration forms into their databases, data entry errors when elections officials record voter turnout, imputation errors when data vendors use statistical models to predict and impute demographic characteristics for individuals, address matching errors when ANES personnel matched ANES sampled addresses to records on vendor files, person matching errors when ANES personnel matched the names, addresses, and birth date information of ANES respondents to records on vendor files, and other data processing errors. Combinations of such errors may be common in these data, and one of the purposes of collecting the data was to evaluate their value.⁴

To use this dataset, users may merge the data file with the ANES 2024 Time Series Study by case ID (V240001) to obtain additional data about sampled persons and households. Many users may focus on the variable `val_turnout24`, which gives the best available indication of validated turnout status in the 2024 presidential election. The most conservative approach to this variable is to treat unlinked cases (coded 1 or 2) as missing data. Many analysts of validated turnout data treat unlinked cases as non-voters, but it is also possible that these are voters who could not be successfully linked to a voter file records due to the types of errors noted in the paragraph above. When choosing whether to use self-reported turnout or validated turnout, and whether to treat unlinked cases as non-voters or as missing data, users should consider the tradeoffs between false

³ For a description of the ANES 2020 procedures, see ANES (2022). User’s Guide and Codebook for the ANES 2020 Time Series Voter Validation Supplemental Data. Available at https://electionstudies.org/wp-content/uploads/2022/10/anes_timeseries_2020_userguidecodebook_VoterValidation.pdf

⁴ For discussion of the methods and accuracy of vote validation procedures, see Matthew DeBell (forthcoming), Towards Credible Vote Validation, *methods, data, analyses*; Ted Enamorado and Kosuke Imai (2019), Validating Self-Reported Turnout by Linking Public Opinion Surveys with Administrative Records. *Public Opinion Quarterly*, 83, 723-748.

positives (treating actual voters as non-voters) and false negatives (treating actual non-voters as voters) for their analysis. Self-reported turnout is likely to generate the most false positives, while validated turnout that treats unlinked cases as non-voters is likely to generate the most false negatives.

Probabilistic Record Linkage with Fastlink

ANES linked ANES 2024 sample records to the L2 national voter file using probabilistic record linkage implemented in the *fastlink* package for R. *Fastlink* estimates the likelihood that a given ANES sample record and a given voter-file record refer to the same individual by comparing multiple personally identifying information (PII) fields simultaneously and producing a posterior match probability score for each candidate pair. These probabilities are used here as an *index* of match strength and to rank prospective matches; they should not be interpreted as calibrated probabilities that a match is correct (see `val_matchprob` and `val_nn_matchprob`).

Fastlink searches with non-VOTEVAL information

Linkage using information collected outside the VOTEVAL section used three fastlink-based searches that differ in their matching variables and in how search candidates were restricted:

- *In-state person matching.* Person matching was performed within state-by-gender subsets (e.g., searching women in California, then separately searching men in California). In-state matches were based on first name, last name, house number, street name, ZIP code, and birth year. String similarity (via Jaro–Winkler) was used for first name, last name, and street name, and numeric distance was used for birth year.
- *In-state address matching.* Address matching was performed within state (without gender subsetting). Matches were based on house number, street name, and ZIP code, using string similarity for street name.
- *Out-of-state person matching.* Out-of-state person matching was performed within gender subsets and searched across voter records from multiple states. Matches were based on first name, middle name, last name, and full date of birth (birth year, month, and day).

Before matching, L2 voter-file fields were standardized for linkage (e.g., names were uppercased and ZIP codes were truncated to five digits).

Out-of-state matching was run selectively for cases where in-state matching did not yield any candidate matches, or where the best in-state person or address match was below a high-probability threshold, .9 (as indicated by the top posterior probability score).

Fastlink implementation

Fastlink compares records across multiple fields by classifying each field comparison as an agreement, partial agreement, or disagreement. For string fields listed in `stringdist.match`, it uses Jaro–Winkler similarity, treating values as agreements when similarity is at least `cut.a` and as partial agreements when similarity is at least `cut.p` (with lower similarity treated as disagreement). For numeric fields listed in `numeric.match`, it uses absolute differences, treating values as agreements when the difference is at most `cut.a.num`; in runs where partial numeric matching is enabled, differences up to `cut.p.num` are treated as partial agreements. In all runs shown here, `return.all = TRUE` retains multiple candidate matches per ANES record, and `dedupe.matches = FALSE` allows the same ANES record to match multiple voter-file records for later evaluation.

In-state person matching used the following script (with another version allowing for ZIP code to be missing):

```
fastLink(  
  dfA = df.input, dfB = df.orig,  
  varnames = c("FirstName", "LastName", "house_number", "street_name", "postalcode", "BirthYear"),  
  stringdist.match = c("FirstName", "LastName", "street_name"),  
  numeric.match = c("BirthYear"),  
  partial.match = c("FirstName", "LastName", "street_name", "BirthYear"),  
  cut.a = 0.94, cut.p = 0.85,  
  cut.a.num = 1, cut.p.num = 2.5,  
  return.all = TRUE,  
  dedupe.matches = FALSE  
)
```

Out-of-state matching script (with another version allowing for middle name to be missing):

```
fastLink(  
  dfA = df.input, dfB = df.orig,  
  varnames = c("FirstName", "MiddleName", "LastName", "BirthYear", "BirthMonth", "BirthDay"),  
  stringdist.match = c("FirstName", "MiddleName", "LastName"),  
  numeric.match = c("BirthYear", "BirthMonth", "BirthDay"),  
  cut.a = 0.94, cut.a.num = 1,  
  return.all = TRUE,  
  dedupe.matches = FALSE  
)
```

Address-based script:

```
fastLink(  
  dfA = df.input, dfB = df.orig,  
  varnames = c("house_number", "street_name", "postalcode"),  
  stringdist.match = c("street_name"),  
  partial.match = c("street_name"),  
  cut.a = 0.94, cut.p = 0.85,  
  cut.a.num = 1, cut.p.num = 2.5,  
  return.all = TRUE,  
  dedupe.matches = FALSE  
)
```

The fastlink highest probability match (FHPM)

The dataset includes a fastlink-based linkage result referred to as the fastlink highest probability match (FHPM). This is defined as the single prospective match having the highest fastlink posterior probability score across the in-state, out-of-state, and address searches, excluding any information from the VOTEVAL section. (In case of ties, preference was given to in-state matches, then address-based matches, then out-of-state matches.) In the public dataset, the FHPM corresponds to key type 12 (see the val_12_* variables, including val_12_matchprob).

Post-election re-search after respondent corrections

Fastlink was also used after the post-election validation check (described below) when a respondent corrected information for a prospective voter-file record shown in the questionnaire. In these cases, ANES performed a new fastlink search using the respondent-corrected information to identify additional possible matches and compute corresponding posterior probabilities (recorded in the relevant val_nn_matchprob variables). Because these later searches can differ from the initial linkage step (for example, due to respondent corrections that change search inputs), the FHPM is not necessarily the same record as the highest-probability match that might appear among other prospective matches identified later.

Fastlink outputs were used for ranking and selection of matches, but final linkage decisions in the released “best match” variables incorporate additional information, including automated match-quality evaluation and respondent feedback.

Automated Clerical Review⁵

In addition to probabilistic linkage scores from fastlink, ANES evaluated prospective person-level matches using an automated “clerical review” procedure implemented in a Python program developed for ANES vote validation work. This step is intended to approximate the judgments that a human reviewer would make when comparing an ANES record to a linked voter-file record and deciding whether the match is plausible, while also producing standardized metadata that can be used to evaluate match quality and analyze sensitivity to alternative linkage decisions.

For each prospective match (up to 12 per case, including the initial fastlink highest-probability match, the set of pre-election candidate matches selected for presentation in the VOTEVAL section of the questionnaire, and any later matches produced after respondent edits), the clerical review program compares identification fields on the ANES case to the corresponding fields on the candidate L2 voter-file record. (These prospective matches correspond to the match types indicated by the variables `val_01_key` through `val_12_key`.) The program then produces:

- counts of **minor** and **major** discrepancies (`val_nn_minor` and `val_nn_major`) between the ANES record and the voter-file record, and
- an ordinal **match quality concern score** (`val_nn_quality`, with lower values indicating fewer concerns and higher values indicating more concerns; -9 indicates the prospective match does not exist), and
- a dichotomous indicator of whether the prospective match is considered a **good match** (`val_nn_match`, coded 1 for good and 0 for not good; -9 indicates the prospective match does not exist).

Counting minor and major differences: The clerical review logic evaluates discrepancies in name, address, and date of birth. (Gender is not examined in clerical review because the matching procedures generating in-state and out-of-state matches are blocked on gender.) The discrepancy counts ignore formatting and other substantively equivalent representations, including capitalization, spacing, and punctuation, as well as common equivalences in suffixes (e.g., Junior and jr are treated as identical), street-type and directional variants (e.g., “E 5th Ave” and “East Fifth Avenue”), leading zeros in dates and ZIP codes, and city-name differences when ZIP codes match. A middle name written out in one record and reduced to an initial is also treated as no difference. *Minor* differences include, for example, a small typographical error (Levenshtein distance⁶ ≤ 2), a recognizable nickname for a first name (e.g. Billy for William, or Peg for Margaret), the omission of a suffix when it is present in the other record, or the omission of an apartment number when present in the other record, or a difference in exactly one DOB component (month, day, or year) when the other two match. *Major* differences include, for example, different first or last names (beyond a single typo or nickname), different (non-missing) suffixes, differences in key address elements

⁵ The clerical review procedure is described in greater detail in DeBell (forthcoming), cited above, footnote 4.

⁶ Levenshtein distance is a measure of the difference of two alphanumeric strings. It indicates the minimum number of single-character edits (insertions, deletions, or substitutions) to change one string to the other. For example, the distance between Oak and Oat is 1, and the distance between Michael and Michelle is 3.

(street number, street name, unit number, state, or ZIP), multiple DOB differences, large DOB differences exceeding 15 years, and cumulative differences where discrepancies reinforce each other, such as the combination of a different name and street number or a suffix difference combined with a DOB gap exceeding 15 years.

Considering alternate names addresses: Survey respondents sometimes report their name differently at different times, or report being registered to vote at a different address than the one where they were sampled. We compared matches to alternate names and addresses when they were available and used the least discrepant match available. The automated clerical review evaluated both the primary and alternate values and used the name that minimized the discrepancy score for the name and the address that minimized the discrepancy score for the address.

Judging matches as good or not: The match quality concern scores (`val_quality` for the best match and `val_nn_quality` for the prospective matches) are calculated from the number of minor and major differences: if $\text{minor} < 3$, $\text{quality} = \text{major}$, with quality values capped at 3; else if $\text{minor} \geq 3$, $\text{quality} = \text{major} + 1$, with quality values capped at 3. The dichotomous match indicator is based on the quality score, where values of 0 or 1 for (concerns about) quality are treated as a good match and values of 2 or 3 (concerns about) quality are treated as not good. Cases were not selected as best matches if there were serious concerns about quality, so `val_quality` does not have values of 2 or 3, but `val_nn_quality` variables do.

These clerical-review outputs serve two distinct purposes in the dataset. First, they provide users with diagnostic information about the plausibility of each prospective match, including matches that were ultimately not selected as the best match. Second, they inform the selection of the single “best matched” voter-file record reported in the summary turnout variables (`val_turnout24`, `val_turnout22`, `val_turnout20`) and summary match metadata (`val_match`, `val_quality`, `val_minor`, `val_major`, `val_matchprob`). In the public file, the “best match” summary variables correspond to whichever prospective match was selected as best (indicated by `val_keymatch`), and the associated clerical review results for that selected prospective match are propagated to `val_match`, `val_quality`, `val_minor`, and `val_major`.

Because the clerical review is algorithmic, it can be applied consistently across all cases and all match types, but it is not a guarantee of correctness. Users analyzing turnout or linkage performance may wish to incorporate these clerical review measures to assess robustness (for example, by restricting analyses to cases with low concern scores or few discrepancies, or by comparing results across alternative prospective matches).

ANES Questionnaire Validation Check

After the pre-election linkage work, ANES incorporated an additional validation step into the post-election questionnaire (the VOTEVAL section) that allowed survey participants to help adjudicate their own potential voter-file matches. This step served two purposes: (1) to evaluate how often algorithmically generated linkages appear correct when reviewed by respondents themselves, and (2) to improve linkage quality in cases where the best available match required corrections (e.g., name or address spelling corrections) or where multiple plausible matches existed.

For sampled persons with sufficient information and at least one plausible prospective match, and meeting other qualifying criteria,⁷ ANES presented up to five candidate voter-file records to the respondent. These candidates were selected using pre-election fastlink-based linkage results generated by in-state person matching, out-of-state person matching, and an address-based matching search⁸ (based on house number, street name, and ZIP code), and then filtered and selected using a set of deterministic selection rules (“VOTEVAL rules”) designed to identify a small set of plausible possibilities to present to survey participants. In outline, the VOTEVAL rules:

- *Filter candidates by fastlink probability*, dropping very low-probability pairs (posterior probability < 0.1).
- *Treat in-state and out-of-state candidates differently*, applying stricter retention criteria for out-of-state candidates (e.g., requiring posterior probability > 0.8 and strong agreement on date of birth and/or name, with a limited allowance for small date-of-birth discrepancies).
- Where many out-of-state candidates are tied, *apply a small tie-break adjustment* based on exact agreement on name components (including middle initial and suffix when available) and closeness of date of birth; if ties remain too numerous to adjudicate, out-of-state candidates are dropped for that case.
- *Deduplicate* candidates that appear multiple times across sources (in-state, out-of-state, and address-based searches), using a consistent priority ordering so that each underlying voter-file record appears only once among the presented candidates.
- *Select up to five total candidates* by retaining the highest-probability candidates within source categories (i.e., the two remaining top in-state candidates, the remaining top out-of-state candidate, and the top address-based candidate where available) and then filling remaining slots with the highest remaining posterior probabilities.
- *Apply additional filtering* to address-based candidate matches (generated from an address-only fastlink search), including dropping candidates when first line of the address (i.e., the house number, name, and unit, if any, such as 123 Main St Apt 4) differs and posterior probability is below a very high threshold of .999, and retaining only the highest-posterior address-based candidate(s) when posterior scores are differentiated.

These selection rules were designed to identify a small set of plausible possibilities rather than simply listing cases with the top five posterior probabilities. As a result, the candidates shown in the VOTEVAL questionnaire module are not guaranteed to include the single record with the highest fastlink posterior probability from the separate FHPM linkage step, though this was often the case, and the FHPM is not necessarily the highest-probability record among all candidates shown to the respondent (although it often is).

During the VOTEVAL module, respondents were shown the full name, gender, date of birth, and address of up to five prospective voter-file records identified as resembling the respondent or linked to the sampled address. Respondents were asked to (a) select which candidate record appeared to be theirs, if any, and then (b) either confirm that the selected record was correct or indicate that it was not entirely correct and provide corrections. The public dataset captures these outcomes using:

⁷ VOTEVAL was not asked if the participant was in the GSS sample or completed the FTF survey by video or phone.

⁸ The address-based matching search used here produces candidate person records for respondent review and is distinct from the address-level household aggregate variables (val_hh_*) included elsewhere in the dataset.

- `val_confirmed`, indicating that a respondent selected one of the VOTEVAL candidates and confirmed it as correct with no edits;
- `val_corrected`, indicating that a respondent selected a candidate but reported corrections; and
- `val_selected_FHPM`, indicating whether the respondent-selected record was the same voter-file record as the initial fastlink highest probability match.

When a respondent provided corrections in VOTEVAL, ANES used the corrected information to perform an additional voter-file search to identify up to five new prospective matches based on the respondent-edited details. These post-edit prospective matches are included among the “considered matches” in the dataset (alongside pre-election prospective matches) and are evaluated with the same match metadata (including fastlink probability scores and automated clerical review outputs). In combination, these questionnaire-based checks and (when applicable) post-edit searches provide a respondent-informed linkage pathway that can be compared to purely algorithmic linkage approaches.

Selecting the Best Match

Voter validation efforts often produced more than one plausible voter-file record for a given ANES case. The public dataset therefore includes (a) turnout indicators and match metadata for *up to 12 prospective matches* (the `val_01_*` through `val_12_*` series), and (b) a single set of summary turnout and match variables (`val_turnout24`, `val_turnout22`, `val_turnout20`, `val_match`, `val_quality`, `val_minor`, `val_major`, `val_matchprob`) corresponding to the *best matched* voter-file record selected by ANES. The variable `val_keymatch` indicates which prospective match (key 1–12) was selected as the best match; a value of 0 indicates that no match was selected.

Best-match selection followed a hierarchy of preference that prioritized respondent-confirmed information when available and otherwise relied on algorithmic linkage results as evaluated by automated clerical review. Specifically, ANES selected the first available match in the list ordered below; except for respondent-confirmed matches, candidates were required to be judged a good match by clerical review to be selected as the best match.

1. *Respondent-confirmed match (no edits)*. If, in the post-election VOTEVAL section, the respondent selected one of the candidate records and confirmed it as correct with no edits (`val_confirmed`), that record was taken as the best match (key type 1). This respondent-confirmed match does not require additional agreement from clerical review to be selected as best.
2. *Matches from a new search after respondent edits*. If the respondent selected a candidate record but provided corrections (`val_corrected`), ANES used the corrected information to run an additional fastlink-based search and identify up to five new prospective matches. These post-edit prospective matches were considered next, in order (key types 2–6 in the public file’s key scheme).
3. *Pre-election VOTEVAL candidate matches*. If no respondent-confirmed match was available (for instance, because the respondent did not complete the post-election questionnaire) and either no corrections were provided or the post-edit search yielded no acceptable match, ANES next considered the up-to-five pre-election VOTEVAL candidate matches (key types 7–11).

4. *Fastlink highest probability match (FHPM)*. If none of the above produced an acceptable match, ANES considered the initial fastlink highest probability match from the FHPM linkage step (key type 12).

This hierarchy is designed to (a) make maximal use of respondent adjudication when it exists, while (b) retaining an algorithmic “best match” for cases without respondent confirmation or successful post-edit linkage. Because prospective matches can sometimes refer to the same underlying voter-file record, the dataset also provides indicators of duplicate matches across keys (the `val_*vs*` variables) and summary counts of the number of matches considered (`val_totalmatches`) and the number of unique voter-file records among them (`val_nummatch`).

Contents of the Dataset

The dataset has 12,955 cases, corresponding to the ANES sample, excluding cases from the General Social Survey’s GSS-ANES sample and a small number of additional cases for which critical identification data were missing.

The dataset contains 177 variables, as follows.

version. ANES 2024 Voter Validation dataset version. This variable identifies the version of the vote validation dataset release. The first release version is 20260619 (produced June 19, 2026).

V240001. ANES 2024 case ID. This case identification number corresponds to the case IDs on the ANES 2024 Time Series Study and Methodology data files.

val_turnout24: Voting status, 2024 presidential election, for the best matched person record. Most data users seeking validated voter turnout for the 2024 presidential election will want to use this variable.

- 1 Linkage not attempted due to missing PII (usually unit nonresponse)
- 2 Linkage attempted, no record found
- 3 Linked, not voting
- 4 Linked, voting

val_turnout22: Voting status, 2022 general election, for the best matched person record. Same codes as `val_turnout24`.

val_turnout20: Voting status, 2020 presidential election, for the best matched person record. Same codes as `val_turnout24`.

val_match: Match status from clerical review for the best-matched person record that was considered good. The code 1 indicates a person match was found for the respondent, while 0 indicates linkage was attempted but no satisfactory match was found, and -1 indicates no person match was attempted due to insufficient information (usually for cases of unit nonresponse). The same cases coded 2 for `val_turnout24`, `val_turnout22`, and `val_turnout20` are coded 0 for `val_match`.

val_quality: Match quality concern score from clerical review, for the best matched person record that was considered good. Codes:

- 1 No satisfactory match was found.
- 0 Lowest concern (highest quality match)
- 1 Some concern (satisfactory match with some discrepancy)

val_minor: Count of minor discrepancies between the ANES record and the best-matched voter file record for cases where a person match was successful. A code of -9 indicates no successful match (either because linkage was not attempted or because no satisfactory record was found when linkage was attempted; these are indicated by a code of -1 on val_quality), and other codes are the integer count of minor discrepancies. See Automated Clerical Review section, above, for definitions of minor and major discrepancies.

val_major: Count of major discrepancies between the ANES record and the best-matched voter file record for cases where a person match was successful. Cases with more than 1 major discrepancy were considered unsuccessful matches and are coded only as “no successful match” in this variable, so the largest value shown is 1. When counts were higher, they are shown in the val_nn_major variables described later.

- 9 No successful match
- 0 No major discrepancies
- 1 One major discrepancy

val_matchprob: The posterior probability calculated by fastlink to index the likelihood of a successful match. This is coded -9 for non-matches and takes a value between 0 and 1 for matches. This is not a calibrated probability; multiple scores with sums exceeding 1 are often generated for possible matches for the same case, and these probabilities differ substantially from the probability that the match is considered correct. See DeBell (forthcoming) for discussion.

val_keymatch: Values from 1 through 12 indicate which of the potential matches indicated by the variables val_01_key through val_12_key, below, was selected as the best match and reported for the variables above. A 0 indicates no match was selected.

val_01_turnout24 through val_12_turnout24: Turnout status in the 2024 general election for the prospective match with key value 1 through 12 as indicated in the variable name. Codes are -1 (no match of this type), 3 (not voting), or 4 (voting). See val_01_key through val_12_key (below) for record types indicated by the numbers (key values).

val_01_turnout22 through val_12_turnout22: As above, but for the 2022 general election.

val_01_turnout20 through val_12_turnout20: As above, but for the 2020 general election.

val_01_key through val_12_key: These 12 variables respectively indicate whether there was a prospective match of each of 12 possible match types, as follows. Key 12 is the fastlink highest probability match (FHPM): the single candidate with the highest fastlink posterior probability score across the searches excluding VOTEVAL information. Key 7 through 11 are the prospective matches selected for display in the VOTEVAL section of the post-election questionnaire (with their clerical review results indicated in val_07_match through val_11_match and val_07_quality through val_11_quality) (and correspond, respectively, to VOTEVAL persons 1 through 5 in the VOTEVAL section of the post-election questionnaire). Key 1 reflects respondent selection of one of the matches 7 through 11, and key 2 through key 6 reflect the results of a follow-up search based on

respondent corrections (to one possible match) provided during the post-election questionnaire. Codes for these variables are 0 to indicate no prospective match of this type, or the number of the variable to indicate a match exists (e.g., val_02_key is coded 0 or 2, val_05_key is coded 0 or 5, etc.).

- 01 Prospective match status: R-confirmed best match with no edits
- 02 Prospective match status: Second search with R-corrected information, person 1
- 03 Prospective match status: Second search with R-corrected information, person 2
- 04 Prospective match status: Second search with R-corrected information, person 3
- 05 Prospective match status: Second search with R-corrected information, person 4
- 06 Prospective match status: Second search with R-corrected information, person 5
- 07 Prospective match status: Questionnaire match candidate, VOTEVAL person 1
- 08 Prospective match status: Questionnaire match candidate, VOTEVAL person 2
- 09 Prospective match status: Questionnaire match candidate, VOTEVAL person 3
- 10 Prospective match status: Questionnaire match candidate, VOTEVAL person 4
- 11 Prospective match status: Questionnaire match candidate, VOTEVAL person 5
- 12 Prospective match status: Fastlink highest probability match (FHPM)

val_fhpm: Indicates if the best match is also the fastlink highest probability match (FHPM). Also see val_selected_FHPM, below.

val_nummatch: Indicates the number of prospective matches (among key 1 through 12) that are unique, that is, that have different voter IDs on the voter file.

val_totalmatches: Indicates the number of prospective matches among key 1 through 12. This is a count of key values present and includes duplicates. It takes values from 0 through 11. The range is 0 to 11 (not 12) because key 01 (respondent-confirmed with no edits) and keys 02 through 06 (second search after respondent corrections) are mutually exclusive; a case can have key 01, or one or more of keys 02 through 06, but not both.

val_corrected: R corrected matching record at VOTEVAL. This indicates the respondent edited a match in the post-election questionnaire. Codes:

- 1 Not applicable, VOTEVAL not administered
- 0 No correction at VOTEVAL (incl. item nonresponse and confirmation without edits)
- 1 VOTEVAL person 1 (key 7) was corrected
- 2 VOTEVAL person 2 (key 8) was corrected
- 3 VOTEVAL person 3 (key 9) was corrected
- 4 VOTEVAL person 4 (key 10) was corrected
- 5 VOTEVAL person 5 (key 11) was corrected

val_confirmed: R confirmed matching record at VOTEVAL. Codes 1-5 indicate the respondent selected the indicated record at VOTEVAL_LIST and confirmed that it was correct with no edits. Codes:

- 1 Not applicable, VOTEVAL not administered
- 0 No confirmation at VOTEVAL (includes item nonresponse and corrections)
- 1 VOTEVAL person 1 (key 7) was correct
- 2 VOTEVAL person 2 (key 8) was correct
- 3 VOTEVAL person 3 (key 9) was correct
- 4 VOTEVAL person 4 (key 10) was correct
- 5 VOTEVAL person 5 (key 11) was correct

val_selected_FHPM: R selected FHPM at VOTEVAL. Indicates if the case the R selected at VOTEVAL is the fastlink highest probability match. Codes:

- 1 Not applicable, VOTEVAL not administered
- 0 No, case selected at VOTEVAL is not FHPM
- 1 Yes, VOTEVAL person 1 (key 7) is FHPM
- 2 Yes, VOTEVAL person 2 (key 8) is FHPM
- 3 Yes, VOTEVAL person 3 (key 9) is FHPM
- 4 Yes, VOTEVAL person 4 (key 10) is FHPM
- 5 Yes, VOTEVAL person 5 (key 11) is FHPM

val_selected_best: At VOTEVAL, R selected the voter record that was ultimately designated the best match after any further record searches and clerical review. Codes:

- 1 Not applicable, VOTEVAL not administered
- 0 No, case selected at VOTEVAL is not best match
- 1 Yes, VOTEVAL person 1 (key 7) is best match
- 2 Yes, VOTEVAL person 2 (key 8) is best match
- 3 Yes, VOTEVAL person 3 (key 9) is best match
- 4 Yes, VOTEVAL person 4 (key 10) is best match
- 5 Yes, VOTEVAL person 5 (key 11) is best match

val_01_match through **val_12_match:** Match status of each prospective match as determined by clerical review. Codes -9 (no prospective match), 0 (not good) and 1 (good). Variables 01 through 12 correspond to key values above.

val_01_quality through **val_12_quality:** Match quality score from algorithmic clerical review. -9 indicates no prospective match of this type, and codes from 0 (best) to 3 (worst) indicate match quality/concerns.

val_01_minor through **val_12_minor:** count of “minor” differences between the ANES record and the prospective match indicated by the number 01 through 12.

val_01_major through **val_12_major:** count of “major” differences between the ANES record and the prospective match indicated by the number 01 through 12.

val_01_matchprob through **val_12_matchprob:** fastlink posterior probability assigned to the prospective match indicated by the number 01 through 12.

val_hh_addmatch: Address/household match status (1=yes, 0=no)

val_hh_number: Number of registered voters (i.e., persons on the L2 voter file) in the matched household. -9 inapplicable. High values for this variable and the other *hh* variables below are likely to constitute errors in the voter database resulting from failures to distinguish among households at a particular address (such as apartment buildings where all units at the address are aggregated). Values for this and the next 7 variables (through *val_hh_independents*) higher than 3 (or 5 for *val_hh_number* only) are topcoded to limit participant disclosure risk.

val_hh_families: Number of families in the matched household.

val_hh_turnout24: Number of household voters who voted in the 2024 presidential election.

val_hh_turnout22: Number of household voters who voted in the 2022 general election.

val_hh_turnout20: Number of household voters who voted in the 2020 presidential election.

val_hh_democrats: Number of registered Democrats in the matched household (modeled).

val_hh_republicans: Number of registered Republicans in the matched household (modeled).

val_hh_independents: Number of registered independents in the matched household (modeled).

val_hh_matchtype: Household match type:

- 0 No address match
- 1 Address match only (without a person match)
- 2 Address match with a person match

val_01vs07 through **val_11vs12:** Indicates if the cases indicated by the two numbers corresponding to *val_nn_key* are duplicates. That is, for example, if *val_01vs07* has the value of 1 (yes), this indicates that the voter data *val_01_turnout* and *val_07_turnout* refer to the same voter file case. Duplication indicator variables are not included when duplicates are precluded by the data generation process. For example, key values 02 through 06 were created to be mutually exclusive so there is no duplication indicator for, for example, 02vs03.

Codebook

```
-----  
version  
-----  
Dataset version
```

Type: String (str51)

Unique values: 1 Missing "": 0/12,955

Tabulation: Freq. Value
 12,955 "ANES 2024 Voter Validation dataset
 version 20260623"

Warning: Variable has embedded blanks.

```
-----  
V240001  
-----  
Case ID
```

Type: Numeric (double)

Range: [140001,399997] Units: 1
Unique values: 12,955 Missing .: 0/12,955

Mean: 240678
Std. dev.: 77860.2

Percentiles: 10% 25% 50% 75% 90%

141308 203489 222666 317835 367009

val_turnout24 Voter turnout 2024 presidential

Type: Numeric (double)
Label: val_turnout24

Range: [1,4] Units: 1
Unique values: 4 Missing : 0/12,955

Tabulation:	Freq.	Numeric	Label
	7,679	1	Linkage not attempted
	672	2	Linkage attempted, no record found
	850	3	Linked, not voting
	3,754	4	Linked, voting

val_turnout22 Voter turnout 2022 general

Type: Numeric (double)
Label: val_turnout22

Range: [1,4] Units: 1
Unique values: 4 Missing : 0/12,955

Tabulation:	Freq.	Numeric	Label
	7,679	1	Linkage not attempted
	672	2	Linkage attempted, no record found
	1,464	3	Linked, not voting
	3,140	4	Linked, voting

val_turnout20 Voter turnout 2020 presidential

Type: Numeric (double)
Label: val_turnout20

Range: [1,4] Units: 1
Unique values: 4 Missing : 0/12,955

Tabulation:	Freq.	Numeric	Label
	7,679	1	Linkage not attempted
	672	2	Linkage attempted, no record found
	914	3	Linked, not voting
	3,690	4	Linked, voting

val_match Person match status

Type: Numeric (double)
Label: val_match

Range: [-1,1] Units: 1
Unique values: 3 Missing : 0/12,955

Tabulation:	Freq.	Numeric	Label
	7,679	-1	Linkage not attempted
	672	0	Linkage attempted, no match found
	4,604	1	Match found

val_quality Person match quality (0=best)

Type: Numeric (double)
Label: val_quality

Range: [-1,1] Units: 1
Unique values: 3 Missing .: 0/12,955

Tabulation:	Freq.	Numeric	Label
	8,351	-1	No satisfactory match found
	3,740	0	Lowest concern (highest quality match)
	864	1	Some concern (satisfactory match with some discrepancy)

val_minor Person match, count of minor discrepancies

Type: Numeric (double)
Label: val_minor, but 5 nonmissing values are not labeled

Range: [-9,4] Units: 1
Unique values: 6 Missing .: 0/12,955

Tabulation:	Freq.	Numeric	Label
	8,351	-9	No successful match
	1,606	0	
	2,280	1	
	656	2	
	60	3	
	2	4	

val_major Person match, count of major discrepancies

Type: Numeric (double)
Label: val_major, but 2 nonmissing values are not labeled

Range: [-9,1] Units: 1
Unique values: 3 Missing .: 0/12,955

Tabulation:	Freq.	Numeric	Label
	8,351	-9	No successful match
	3,802	0	
	802	1	

val_matchprob Fastlink match quality score (posterior probability)

Type: Numeric (double)
Label: val_matchprob, but 1003 nonmissing values are not labeled

Range: [-9,1] Units: 1.000e-10
Unique values: 1,004 Missing .: 0/12,955

Examples: -9 No successful match
-9 No successful match
-9 No successful match
1

val_keymatch Which match from val_01_key thru val_12_key selected as best

val_05_turnout22 2022 turnout status for prospective match key=5

Type: Numeric (double)
Label: val_05_turnout22
Range: [-1,4] Units: 1
Unique values: 3 Missing .: 0/12,955
Tabulation: Freq. Numeric Label
12,905 -1 No match of this type
22 3 Not voting
28 4 Voting

val_05_turnout20 2020 turnout status for prospective match key=5

Type: Numeric (double)
Label: val_05_turnout20
Range: [-1,4] Units: 1
Unique values: 3 Missing .: 0/12,955
Tabulation: Freq. Numeric Label
12,905 -1 No match of this type
20 3 Not voting
30 4 Voting

val_06_turnout24 2024 turnout status for prospective match key=6

Type: Numeric (double)
Label: val_06_turnout24
Range: [-1,4] Units: 1
Unique values: 3 Missing .: 0/12,955
Tabulation: Freq. Numeric Label
12,944 -1 No match of this type
3 3 Not voting
8 4 Voting

val_06_turnout22 2022 turnout status for prospective match key=6

Type: Numeric (double)
Label: val_06_turnout22
Range: [-1,4] Units: 1
Unique values: 3 Missing .: 0/12,955
Tabulation: Freq. Numeric Label
12,944 -1 No match of this type
4 3 Not voting
7 4 Voting

val_06_turnout20 2020 turnout status for prospective match key=6

Type: Numeric (double)
Label: val_06_turnout20
Range: [-1,4] Units: 1
Unique values: 3 Missing .: 0/12,955

```

Tabulation: Freq.   Numeric   Label
            12,944      -1    No match of this type
            2          3    Not voting
            9          4    Voting
-----
val_07_turnout24                                     2024 turnout status for prospective match key=7
-----

Type: Numeric (double)
Label: val_07_turnout24

Range: [-1,4]                                     Units: 1
Unique values: 3                                  Missing .: 0/12,955

Tabulation: Freq.   Numeric   Label
            7,900      -1    No match of this type
            1,090      3    Not voting
            3,965      4    Voting
-----
val_07_turnout22                                     2022 turnout status for prospective match key=7
-----

Type: Numeric (double)
Label: val_07_turnout22

Range: [-1,4]                                     Units: 1
Unique values: 3                                  Missing .: 0/12,955

Tabulation: Freq.   Numeric   Label
            7,900      -1    No match of this type
            1,710      3    Not voting
            3,345      4    Voting
-----
val_07_turnout20                                     2020 turnout status for prospective match key=7
-----

Type: Numeric (double)
Label: val_07_turnout20

Range: [-1,4]                                     Units: 1
Unique values: 3                                  Missing .: 0/12,955

Tabulation: Freq.   Numeric   Label
            7,900      -1    No match of this type
            1,044      3    Not voting
            4,011      4    Voting
-----
val_08_turnout24                                     2024 turnout status for prospective match key=8
-----

Type: Numeric (double)
Label: val_08_turnout24

Range: [-1,4]                                     Units: 1
Unique values: 3                                  Missing .: 0/12,955

Tabulation: Freq.   Numeric   Label
            9,125      -1    No match of this type
            1,104      3    Not voting
            2,726      4    Voting
-----
val_08_turnout22                                     2022 turnout status for prospective match key=8
-----

```


1,308 4 Voting

val_10_turnout24 2024 turnout status for prospective match key=10

Type: Numeric (double)
Label: val_10_turnout24

Range: [-1,4] Units: 1
Unique values: 3 Missing .: 0/12,955

Tabulation:	Freq.	Numeric	Label
	11,953	-1	No match of this type
	371	3	Not voting
	631	4	Voting

val_10_turnout22 2022 turnout status for prospective match key=10

Type: Numeric (double)
Label: val_10_turnout22

Range: [-1,4] Units: 1
Unique values: 3 Missing .: 0/12,955

Tabulation:	Freq.	Numeric	Label
	11,953	-1	No match of this type
	535	3	Not voting
	467	4	Voting

val_10_turnout20 2020 turnout status for prospective match key=10

Type: Numeric (double)
Label: val_10_turnout20

Range: [-1,4] Units: 1
Unique values: 3 Missing .: 0/12,955

Tabulation:	Freq.	Numeric	Label
	11,953	-1	No match of this type
	381	3	Not voting
	621	4	Voting

val_11_turnout24 2024 turnout status for prospective match key=11

Type: Numeric (double)
Label: val_11_turnout24

Range: [-1,4] Units: 1
Unique values: 3 Missing .: 0/12,955

Tabulation:	Freq.	Numeric	Label
	12,512	-1	No match of this type
	169	3	Not voting
	274	4	Voting

val_11_turnout22 2022 turnout status for prospective match key=11

Type: Numeric (double)
Label: val_11_turnout22

val_01_key Prospective match status: R-confirmed best match with no edits

Type: Numeric (double)
Label: val_01_key
Range: [0,1] Units: 1
Unique values: 2 Missing .: 0/12,955
Tabulation: Freq. Numeric Label
9,880 0 No match of this type
3,075 1 The respondent-confirmed best
match with no edits

val_02_key Prospective match status: Second search with R-corrected info, person 1

Type: Numeric (double)
Label: val_02_key
Range: [0,2] Units: 1
Unique values: 2 Missing .: 0/12,955
Tabulation: Freq. Numeric Label
12,526 0 No match of this type
429 2 Second search with R-corrected
info, person 1

val_03_key Prospective match status: Second search with R-corrected info, person 2

Type: Numeric (double)
Label: val_03_key
Range: [0,3] Units: 1
Unique values: 2 Missing .: 0/12,955
Tabulation: Freq. Numeric Label
12,692 0 No match of this type
263 3 Second search with R-corrected
info, person 2

val_04_key Prospective match status: Second search with R-corrected info, person 3

Type: Numeric (double)
Label: val_04_key
Range: [0,4] Units: 1
Unique values: 2 Missing .: 0/12,955
Tabulation: Freq. Numeric Label
12,847 0 No match of this type
108 4 Second search with R-corrected
info, person 3

val_05_key Prospective match status: Second search with R-corrected info, person 4

Type: Numeric (double)
Label: val_05_key
Range: [0,5] Units: 1
Unique values: 2 Missing .: 0/12,955

```

Tabulation: Freq.  Numeric  Label
             12,905      0  No match of this type
             50         5  Second search with R-corrected
                           info, person 4
-----
val_06_key      Prospective match status: Second search with R-corrected info, person 5
-----

Type: Numeric (double)
Label: val_06_key

Range: [0,6]          Units: 1
Unique values: 2      Missing .: 0/12,955

Tabulation: Freq.  Numeric  Label
             12,944      0  No match of this type
             11         6  Second search with R-corrected
                           info, person 5
-----
val_07_key      Prospective match status: Questionnaire match candidate, VOTEVAL person 1
-----

Type: Numeric (double)
Label: val_07_key

Range: [0,7]          Units: 1
Unique values: 2      Missing .: 0/12,955

Tabulation: Freq.  Numeric  Label
             7,900      0  No match of this type
             5,055      7  Questionnaire match candidate,
                           VOTEVAL person 1
-----
val_08_key      Prospective match status: Questionnaire match candidate, VOTEVAL person 2
-----

Type: Numeric (double)
Label: val_08_key

Range: [0,8]          Units: 1
Unique values: 2      Missing .: 0/12,955

Tabulation: Freq.  Numeric  Label
             9,125      0  No match of this type
             3,830      8  Questionnaire match candidate,
                           VOTEVAL person 2
-----
val_09_key      Prospective match status: Questionnaire match candidate, VOTEVAL person 3
-----

Type: Numeric (double)
Label: val_09_key

Range: [0,9]          Units: 1
Unique values: 2      Missing .: 0/12,955

Tabulation: Freq.  Numeric  Label
             10,984      0  No match of this type
             1,971      9  Questionnaire match candidate,
                           VOTEVAL person 3
-----
val_10_key      Prospective match status: Questionnaire match candidate, VOTEVAL person 4
-----

```



```

17      3 VOTEVAL person 3 (key 09) was
        correct
5       4 VOTEVAL person 4 (key 10) was
        correct

```

```
-----
val_selected_FHPM                                R selected fastlink highest prob match at VOTEVAL
-----
```

```

Type: Numeric (double)
Label: val_selected_FHPM

```

```

Range: [-1,5]                                Units: 1
Unique values: 7                             Missing .: 0/12,955

```

```

Tabulation: Freq.  Numeric  Label
8,986      -1  Not applicable, VOTEVAL not
           administered
           763    0  No, case selected at VOTEVAL is
           not FHPM
3,163      1  Yes, VOTEVAL person 1 (key 07)
           is FHPM
           36    2  Yes, VOTEVAL person 2 (key 08)
           is FHPM
           3     3  Yes, VOTEVAL person 3 (key 09)
           is FHPM
           3     4  Yes, VOTEVAL person 4 (key 10)
           is FHPM
           1     5  Yes, VOTEVAL person 5 (key 11)
           is FHPM

```

```
-----
val_selected_best                                R selected best match at VOTEVAL
-----
```

```

Type: Numeric (double)
Label: val_selected_best

```

```

Range: [-1,5]                                Units: 1
Unique values: 7                             Missing .: 0/12,955

```

```

Tabulation: Freq.  Numeric  Label
9,432      -1  Not applicable, VOTEVAL not
           administered
           126    0  No, case selected at VOTEVAL is
           not best match
3,278      1  Yes, VOTEVAL person 1 (key 07)
           is best match
           102   2  Yes, VOTEVAL person 2 (key 08)
           is best match
           13    3  Yes, VOTEVAL person 3 (key 09)
           is best match
           3     4  Yes, VOTEVAL person 4 (key 10)
           is best match
           1     5  Yes, VOTEVAL person 5 (key 11)
           is best match

```

```
-----
val_01_match                                    Match status of prospective match identified by key 01
-----
```

```

Type: Numeric (double)
Label: val_01_match

```

```

Range: [-9,1]                                Units: 1
Unique values: 3                             Missing .: 0/12,955

```

```

Tabulation: Freq.  Numeric  Label
9,880      -9  No prospective match of this

```

		type
82	0	Not good
2,993	1	Good

val_01_quality Match quality score for match identified by key 01

Type: Numeric (double)
Label: val_01_quality

Range: [-9,3] Units: 1
Unique values: 5 Missing .: 0/12,955

Tabulation:	Freq.	Numeric	Label
	9,880	-9	No prospective match of this type
	2,606	0	Best
	387	1	Likely match, not best
	38	2	Unlikely match, not worst
	44	3	Worst

val_01_minor Minor difference count b/w ANES and record key 01

Type: Numeric (double)
Label: val_01_minor, but 5 nonmissing values are not labeled

Range: [-9,4] Units: 1
Unique values: 6 Missing .: 0/12,955

Tabulation:	Freq.	Numeric	Label
	9,880	-9	No prospective match of this type
	1,087	0	
	1,511	1	
	426	2	
	50	3	
	1	4	

val_01_major Major difference count b/w ANES and record key 01

Type: Numeric (double)
Label: val_01_major, but 7 nonmissing values are not labeled

Range: [-9,6] Units: 1
Unique values: 8 Missing .: 0/12,955

Tabulation:	Freq.	Numeric	Label
	9,880	-9	No prospective match of this type
	2,651	0	
	347	1	
	33	2	
	21	3	
	15	4	
	4	5	
	4	6	

val_01_matchprob Fastlink posterior prob for key 01

Type: Numeric (double)
Range: [-9,1] Units: 1.000e-10

Unique values: 568 Missing .: 0/12,955

Mean: -6.62756
Std. dev.: 4.25279

Percentiles:	10%	25%	50%	75%	90%
	-9	-9	-9	-9	1

val_02_match Match status of prospective match identified by key 02

Type: Numeric (double)
Label: val_02_match

Range: [-9,1] Units: 1
Unique values: 3 Missing .: 0/12,955

Tabulation:	Freq.	Numeric	Label
	12,526	-9	No prospective match of this type
	42	0	Not good
	387	1	Good

val_02_quality Match quality score for match identified by key 02

Type: Numeric (double)
Label: val_02_quality

Range: [-9,3] Units: 1
Unique values: 5 Missing .: 0/12,955

Tabulation:	Freq.	Numeric	Label
	12,526	-9	No prospective match of this type
	227	0	Best
	160	1	Likely match, not best
	25	2	Unlikely match, not worst
	17	3	Worst

val_02_minor Minor difference count b/w ANES and record key 02

Type: Numeric (double)
Label: val_02_minor, but 4 nonmissing values are not labeled

Range: [-9,3] Units: 1
Unique values: 5 Missing .: 0/12,955

Tabulation:	Freq.	Numeric	Label
	12,526	-9	No prospective match of this type
	222	0	
	196	1	
	9	2	
	2	3	

val_02_major Major difference count b/w ANES and record key 02

Type: Numeric (double)
Label: val_02_major, but 7 nonmissing values are not labeled

Range: [-9,6] Units: 1
Unique values: 8 Missing .: 0/12,955

Tabulation:	Freq.	Numeric	Label
	12,526	-9	No prospective match of this type
	227	0	
	160	1	
	27	2	
	6	3	
	5	4	
	3	5	
	1	6	

val_02_matchprob Fastlink posterior prob for key 02

Type: Numeric (double)
Range: [-9,1] Units: 1.000e-09
Unique values: 198 Missing .: 0/12,955
Mean: -8.66917
Std. dev.: 1.78774
Percentiles: 10% 25% 50% 75% 90%
 -9 -9 -9 -9 -9

val_03_match Match status of prospective match identified by key 03

Type: Numeric (double)
Label: val_03_match
Range: [-9,1] Units: 1
Unique values: 3 Missing .: 0/12,955
Tabulation: Freq. Numeric Label
 12,692 -9 No prospective match of this type
 248 0 Not good
 15 1 Good

val_03_quality Match quality score for match identified by key 03

Type: Numeric (double)
Label: val_03_quality
Range: [-9,3] Units: 1
Unique values: 5 Missing .: 0/12,955
Tabulation: Freq. Numeric Label
 12,692 -9 No prospective match of this type
 3 0 Best
 12 1 Likely match, not best
 12 2 Unlikely match, not worst
 236 3 Worst

val_03_minor Minor difference count b/w ANES and record key 03

Type: Numeric (double)
Label: val_03_minor, but 4 nonmissing values are not labeled
Range: [-9,3] Units: 1


```

1      4
5      5
2      6
2      7

```

val_06_matchprob Fastlink posterior prob for key 06

```

Type: Numeric (double)
Range: [-9, .99999563]           Units: 1.000e-08
Unique values: 9                 Missing  .: 0/12,955

```

```

Tabulation: Freq. Value
12,944      -9
1           .49547108
1           .9842693
1           .99879134
1           .999175
1           .99992972
2           .99999509
3           .9999955
1           .99999563

```

val_07_match Match status of prospective match identified by key 07

```

Type: Numeric (double)
Label: val_07_match
Range: [-9,1]                   Units: 1
Unique values: 3                 Missing  .: 0/12,955

```

```

Tabulation: Freq. Numeric Label
7,900      -9 No prospective match of this
              type
743         0 Not good
4,312       1 Good

```

val_07_quality Match quality score for match identified by key 07

```

Type: Numeric (double)
Label: val_07_quality
Range: [-9,3]                   Units: 1
Unique values: 5                 Missing  .: 0/12,955

```

```

Tabulation: Freq. Numeric Label
7,900      -9 No prospective match of this
              type
3,572       0 Best
740         1 Likely match, not best
163         2 Unlikely match, not worst
580         3 Worst

```

val_07_minor Minor difference count b/w ANES and record key 07

```

Type: Numeric (double)
Label: val_07_minor, but 5 nonmissing values are not labeled
Range: [-9,4]                   Units: 1
Unique values: 6                 Missing  .: 0/12,955

```

Tabulation:	Freq.	Numeric	Label
	7,900	-9	No prospective match of this type
	1,584	0	
	2,648	1	
	756	2	
	66	3	
	1	4	

val_07_major Major difference count b/w ANES and record key 07

Type: Numeric (double)
Label: val_07_major, but 8 nonmissing values are not labeled

Range: [-9,7] Units: 1
Unique values: 9 Missing .: 0/12,955

Tabulation:	Freq.	Numeric	Label
	7,900	-9	No prospective match of this type
	3,623	0	
	698	1	
	156	2	
	148	3	
	223	4	
	147	5	
	54	6	
	6	7	

val_07_matchprob Fastlink posterior prob for key 07

Type: Numeric (double)

Range: [-9,1] Units: 1.000e-10
Unique values: 1,154 Missing .: 0/12,955

Mean: -5.10755
Std. dev.: 4.86674

Percentiles:	10%	25%	50%	75%	90%
	-9	-9	-9	1	1

val_08_match Match status of prospective match identified by key 08

Type: Numeric (double)
Label: val_08_match

Range: [-9,1] Units: 1
Unique values: 3 Missing .: 0/12,955

Tabulation:	Freq.	Numeric	Label
	9,125	-9	No prospective match of this type
	3,611	0	Not good
	219	1	Good

val_08_quality Match quality score for match identified by key 08

Type: Numeric (double)
Label: val_08_quality

Tabulation:	Freq.	Numeric	Label
	10,984	-9	No prospective match of this type
	1,936	0	Not good
	35	1	Good

val_09_quality Match quality score for match identified by key 09

Type: Numeric (double)
Label: val_09_quality

Range: [-9,3] Units: 1
Unique values: 5 Missing .: 0/12,955

Tabulation:	Freq.	Numeric	Label
	10,984	-9	No prospective match of this type
	9	0	Best
	26	1	Likely match, not best
	90	2	Unlikely match, not worst
	1,846	3	Worst

val_09_minor Minor difference count b/w ANES and record key 09

Type: Numeric (double)
Label: val_09_minor, but 5 nonmissing values are not labeled

Range: [-9,4] Units: 1
Unique values: 6 Missing .: 0/12,955

Tabulation:	Freq.	Numeric	Label
	10,984	-9	No prospective match of this type
	492	0	
	1,172	1	
	278	2	
	28	3	
	1	4	

val_09_major Major difference count b/w ANES and record key 09

Type: Numeric (double)
Label: val_09_major, but 9 nonmissing values are not labeled

Range: [-9,8] Units: 1
Unique values: 10 Missing .: 0/12,955

Examples: -9 No prospective match of this type
-9 No prospective match of this type
-9 No prospective match of this type
-9 No prospective match of this type

val_09_matchprob Fastlink posterior prob for key 09

Type: Numeric (double)

Range: [-9,1] Units: 1.000e-11
Unique values: 683 Missing .: 0/12,955

Mean: -7.50475
Std. dev.: 3.53169

-9 No prospective match of this type

val_10_matchprob Fastlink posterior prob for key 10

Type: Numeric (double)
Range: [-9,.99999988] Units: 1.000e-10
Unique values: 497 Missing .: 0/12,955
Mean: -8.24757
Std. dev.: 2.60049
Percentiles: 10% 25% 50% 75% 90%
-9 -9 -9 -9 -9

val_11_match Match status of prospective match identified by key 11

Type: Numeric (double)
Label: val_11_match
Range: [-9,1] Units: 1
Unique values: 3 Missing .: 0/12,955
Tabulation: Freq. Numeric Label
12,512 -9 No prospective match of this
type
441 0 Not good
2 1 Good

val_11_quality Match quality score for match identified by key 11

Type: Numeric (double)
Label: val_11_quality
Range: [-9,3] Units: 1
Unique values: 4 Missing .: 0/12,955
Tabulation: Freq. Numeric Label
12,512 -9 No prospective match of this
type
2 1 Likely match, not best
16 2 Unlikely match, not worst
425 3 Worst

val_11_minor Minor difference count b/w ANES and record key 11

Type: Numeric (double)
Label: val_11_minor, but 4 nonmissing values are not labeled
Range: [-9,3] Units: 1
Unique values: 5 Missing .: 0/12,955
Tabulation: Freq. Numeric Label
12,512 -9 No prospective match of this
type
103 0
247 1
89 2
4 3

val_12_minor Minor difference count b/w ANES and record key 12

Type: Numeric (double)
Label: val_12_minor, but 5 nonmissing values are not labeled

Range: [-9,4] Units: 1
Unique values: 6 Missing .: 0/12,955

Tabulation:	Freq.	Numeric	Label
	7,679	-9	No prospective match of this type
	1,611	0	
	2,689	1	
	878	2	
	96	3	
	2	4	

val_12_major Major difference count b/w ANES and record key 12

Type: Numeric (double)
Label: val_12_major, but 9 nonmissing values are not labeled

Range: [-9,8] Units: 1
Unique values: 10 Missing .: 0/12,955

Examples: -9 No prospective match of this type
-9 No prospective match of this type
0
0

val_12_matchprob Fastlink posterior prob for key 12

Type: Numeric (double)

Range: [-9,1] Units: 1.000e-10
Unique values: 790 Missing .: 0/12,955

Mean: -4.93361
Std. dev.: 4.90628

Percentiles:	10%	25%	50%	75%	90%
	-9	-9	-9	1	1

val_hh_addmatch Address/household match status

Type: Numeric (double)
Label: val_hh_addmatch

Range: [0,1] Units: 1
Unique values: 2 Missing .: 0/12,955

Tabulation:	Freq.	Numeric	Label
	3,264	0	No
	9,691	1	Yes

val_hh_number Number of voters in household

Type: Numeric (double)

val_04vs10 Duplicate status of matches in key 04 vs key 10

Type: Numeric (double)
Label: val_04vs10

Range: [-1,1] Units: 1
Unique values: 3 Missing : 0/12,955

Tabulation:	Freq.	Numeric	Label
	12,914	-1	No pairing (so not a duplicate either)
	34	0	Not duplicate
	7	1	Duplicate

val_04vs11 Duplicate status of matches in key 04 vs key 11

Type: Numeric (double)
Label: val_04vs11

Range: [-1,1] Units: 1
Unique values: 3 Missing : 0/12,955

Tabulation:	Freq.	Numeric	Label
	12,940	-1	No pairing (so not a duplicate either)
	14	0	Not duplicate
	1	1	Duplicate

val_04vs12 Duplicate status of matches in key 04 vs key 12

Type: Numeric (double)
Label: val_04vs12

Range: [-1,1] Units: 1
Unique values: 3 Missing : 0/12,955

Tabulation:	Freq.	Numeric	Label
	12,847	-1	No pairing (so not a duplicate either)
	106	0	Not duplicate
	2	1	Duplicate

val_05vs07 Duplicate status of matches in key 05 vs key 07

Type: Numeric (double)
Label: val_05vs07

Range: [-1,1] Units: 1
Unique values: 3 Missing : 0/12,955

Tabulation:	Freq.	Numeric	Label
	12,905	-1	No pairing (so not a duplicate either)
	48	0	Not duplicate
	2	1	Duplicate

val_05vs08 Duplicate status of matches in key 05 vs key 08

Type: Numeric (double)

val_06vs11 Duplicate status of matches in key 06 vs key 11

Type: Numeric (double)
Label: val_06vs11

Range: [-1,1] Units: 1
Unique values: 3 Missing .: 0/12,955

Tabulation:	Freq.	Numeric	Label
	12,951	-1	No pairing (so not a duplicate either)
	3	0	Not duplicate
	1	1	Duplicate

val_06vs12 Duplicate status of matches in key 06 vs key 12

Type: Numeric (double)
Label: val_06vs12

Range: [-1,1] Units: 1
Unique values: 3 Missing .: 0/12,955

Tabulation:	Freq.	Numeric	Label
	12,944	-1	No pairing (so not a duplicate either)
	10	0	Not duplicate
	1	1	Duplicate

val_07vs12 Duplicate status of matches in key 07 vs key 12

Type: Numeric (double)
Label: val_07vs12

Range: [-1,1] Units: 1
Unique values: 3 Missing .: 0/12,955

Tabulation:	Freq.	Numeric	Label
	7,900	-1	No pairing (so not a duplicate either)
	777	0	Not duplicate
	4,278	1	Duplicate

val_08vs12 Duplicate status of matches in key 08 vs key 12

Type: Numeric (double)
Label: val_08vs12

Range: [-1,1] Units: 1
Unique values: 3 Missing .: 0/12,955

Tabulation:	Freq.	Numeric	Label
	9,125	-1	No pairing (so not a duplicate either)
	3,618	0	Not duplicate
	212	1	Duplicate

val_09vs12 Duplicate status of matches in key 09 vs key 12
