

VEST NC 2016

State: North Carolina

Organization: VEST

Date Updated: Report Date: 04/05/2021, VEST File Date: 06/24/2019

1. Is all raw data available?

Yes

- Accessible files:
 - VEST, North Carolina 2016
 - Accessed, 03/09/2021
 - Source: VEST on the Harvard Dataverse
 - Direct link:
<https://dataverse.harvard.edu/file.xhtml?persistentId=doi:10.7910/DVN/NH5S2I/EL3UAZ&version=56.0>
 - Link to VEST 2016 Datasets:
<https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/NH5S2I&version=56.0>
 - Election Results for 11/08/2016
 - Accessed, 03/10/2021
 - Source: North Carolina State Board of Elections (NCSBE)
 - Direct link:
https://s3.amazonaws.com/dl.ncsbe.gov/ENRS/2016_11_08/results_sort_20161108.zip
 - Link to Election Data for 11/08/2016:
https://dl.ncsbe.gov/?prefix=ENRS/2016_11_08/
 - Note: The election dataset discussed in this report is titled: “results_sort_20161108.zip”. To access this file from the main NCSBE download page (<https://dl.ncsbe.gov/>), select “ENRS/”, then the election “2016_11_08/”, then “results_sort_20161108.zip”. Last updated by the NCSBE on 02/13/2017.
 - Voter Registration for 11/08/2016

- Accessed, 03/22/2021
 - Source: NCSBE
 - Direct link:
 - https://s3.amazonaws.com/dl.ncsbe.gov/ENRS/2016_11_08/voter_stats_20161108.zip
 - Link to Election Data for 11/08/2016:
 - https://dl.ncsbe.gov/?prefix=ENRS/2016_11_08/
 - Note: The voter registration file discussed in this report is titled: “voter_stats_20161108.zip”. To access this file from the main NCSBE download page (<https://dl.ncsbe.gov/>), select “ENRS”, then the election “2016_11_08”, then “voter_stats_20161108.zip”. Last updated by the NCSBE on 02/13/2017.
- Precinct Shapefile for 11/08/2016
 - Accessed, 03/09/2021
 - Source: NCSBE
 - Direct link:
 - https://s3.amazonaws.com/dl.ncsbe.gov/PrecinctMaps/SBE_PRECINCTS_20161004.zip
 - Link to state-wide precinct shapefiles:
 - <https://dl.ncsbe.gov/?prefix=PrecinctMaps/>
 - Note: The precinct shapefile discussed in this report is titled: “SBE_PRECINCTS_20161004.zip” (It is the most recent file available prior to the 11/08/2016 election). To access this file from the main NCSBE download page (<https://dl.ncsbe.gov/>), select “PrecinctMaps”, then “SBE_PRECINCTS_20161004.zip”. Last updated by the NCSBE on 10/04/2016.
- Inaccessible files:
 - N/A

2. Processing steps available?

Yes

- Description of processing steps:
 - VEST documented that they had to allocate curbside/early/provisional ballots in their documentation.txt file under North Carolina (<https://dataverse.harvard.edu/file.xhtml?fileId=4499004&version=56.0>, retrieved 03/09/2021):
 - “The following counties had curbside, provisional, or miscellaneous early votes reported at the county level: Bertie, Catawba, Durham, Halifax, Jackson, Lenoir, Martin, Mecklenburg, Pitt, Robeson, Stokes, and Surry. These were distributed to precincts based on the precinct's share of the county vote for each candidate.”
 - VEST also documented a number of precincts that needed to be split in the election results:
 - “The following precincts reported combined results. These were distributed based on voter registration shares as of Election Day (also from the NCSBE).
 - Buncombe: 05.1/56.2/68.1

- Columbus: P20A/P22A
- Craven 17/19
- Cumberland: CC17/CC31
- Edgecombe: 0101/0104
- Forsyth: 402/403
- Gaston: 39/43
- Granville: BERE/CORI/WOEL
- Mecklenburg: 013/014, 016/141, 063/064, 115/116, 150/151, 225/226
- New Hanover: W03/W21
- Robeson: 03/11A/15/38
- Scotland: 09/10
- Wake: 01-23/01-25
- Washington: LM/P3
- Wilson: PRTA/PRTO”
- They also noted that “In cases with combined precincts and countywide vote pools, the precinct votes were distributed first, then the countywide votes.”
- There were a few details that would have been helpful to have in the documentation:
 - There are multiple election files and shapefiles available - to be more clear, VEST could indicate the names of the files they used.
 - VEST did not write down the names of candidates exactly as they are recorded in the election results. These are the differences:
 - “Deborah Ross” in VEST is “Deborah K. Ross” in the election results.
 - “Donald Ray Blue” in VEST is “Donald Ray Buie” in the election results. This is a notable difference, given that they have incorrectly typed the candidate’s name.
 - “Donald Trump” in VEST is “Donald J. Trump” in the election results.
 - “Michael R. Mogan” in VEST is “Michael R. (Mike) Morgan” in the election results. This is a notable difference, given that they have incorrectly typed the candidate’s name.
 - “Robert H. Edmunds” in VEST is “Robert H. (Bob) Edmunds” in the election results.
 - Although VEST did indicate how they allocated votes, which we were able to replicate, they did not indicate their rounding methods which lead to slight (one or two) vote differences for many precincts.

3. Able to replicate joining election data and shapefiles?

Yes

- The election results file required a notable amount of processing in order to join it to the shapefile.
 - The data was broken into subsets by candidates and contests recorded in the file.

- Two precincts were renamed because of typos in the raw dataset (Mecklenburg 225;226 was recorded as 115;116 in one field, but correctly as 225;226 in the other, and Gaston 39;43 was recorded as 38;43).
- A unique ID field was made across all files (VEST, election results, precinct shapefile, voter registration file) using the county name (all capitalized) underscore, then the precinct code (e.g. BUNCOMBE_58.1)
- Since the write-in votes are recorded as candidate “Write-In” and are not unique, a combined contest name and candidate field was created (e.g. US PRESIDENT Write-In, NC ATTORNEY GENERAL Buck Newton)
- The data was pivoted on the unique ID and null values were assigned 0.
- The data was subsetted by fields that were designated as “votes” (there were other fields that were carried over in the pivot that are not needed)
- The columns were renamed to fit VESTs nomenclature, and then reordered to match VEST.
- The voter registration file was grouped by the unique ID field that was made to give one number of total voters registered, instead of broken out by age categories. There is a typo in the election results that is fixed (“ROBESON_03;5;11A;38;PROV” to “ROBESON_03;15;11A;38;PROV”)
- A list of precincts that needed to be split was extracted (based on VEST documentation). This list was used in a function that was written to reassign votes to the precincts. The voter registration is provided for the split precincts. It is summed together for the respective matching election results (e.g. ROBESON_03, ROBESON_15, ROBESON_11A, and ROBESON_38 voter registration totals are summed to get a voter registration number for ROBESON_03;15;11A;38;PROV). The ratio of the split precincts (such as ROBESON_03) to the summed total voter registrants as the precinct is represented in the election data (e.g. ROBESON_03;15;11A;38;PROV) was used to reallocate the precincts share of the results. These new results were added to the data frame and the joined results in the original dataset were removed. Votes were rounded to the nearest integer.
- The list of counties provided by VEST was used to reallocate all early/provisional votes to all other precincts within their respective counties based on a candidates performance. A largest remainder method was used to round votes.
- Join 1:
 - After completing the steps above, a join attempt was made between the election results and the shapefile, based on the unique ID that was created.
 - There were a few precincts that did not join correctly because their precinct names differed slightly between the two files, as the election results included more information such as “MISC” following the precinct name. To remedy this, the unique ID for all the election results that did not join correctly were split by the semi-colon, and everything prior to the semi-colon was saved, the rest removed.
- Join 2:

- With the changes made to the election results for the 7 precincts that required modification, a second join was attempted in which the election results and the shapefile successfully joined 100% (2,704) of the precincts.

4. Able to replicate joining demographic data to block-level shapefiles?

N/A

- There was no demographic data in the VEST NC 2016 shapefile

5. Able to replicate joining boundary data?

N/A

- There was no additional boundary information in the VEST NC 2016 shapefile

6. Successfully validated election results?

No, not all.

- Election results:
 - Candidate totals:
 - All candidate totals state-wide matched within two votes. These slight discrepancies are likely due to rounding differences from assigning absentee votes and splitting precincts.
 - Precinct by precinct:
 - All candidates in all precincts, except four precincts (Mecklenburg 115, 116, 225, and 226) validated to be within a max of two votes (likely due to rounding differences).
 - Four precincts had some candidate values that matched between the RDH created file and the VEST file, however there were some large discrepancies.
 - The four precincts were Mecklenburg 225, Mecklenburg 226, Mecklenburg 115, and Mecklenburg 116.
 - The sum of the four precincts appear to be the same for all candidates, so any discrepancy is internal to these four precincts.
- Geographies:
 - The file produced by RDH and the VEST file both have 2,704 precincts (and thus, 2,704 precinct geometries). 2,700 of these precincts matched, and four differed slightly in area. However, it rounded down to 0.00% of the state's area. The difference is likely due to our verification being too precise. See below for the plotted precincts that did not match in geometry (using geom_almost_equals function in geopandas library). The first are the VEST precincts, and the second are the RDH precincts.

