VEST MO 2016

State: Missouri
Organization: Voting and Election Science Team (VEST)
Summary of Races included: Presidential, U.S. Senate, Attorney General, Governor, Lieutenant Governor, Treasurer, Secretary of State
Date File Updated: 02/03/2021
Date Report Updated: 05/19/2021
RDH Validation Code (Github): https://github.com/nonpartisan-redistricting-datahub/pdv-mo

RDH Criteria	Explanation

Is all raw data available?
 	
No
	<u>Accessible files:</u>
	● VEST MO 16 data file
	  ○ Accessed: 05/06/21, Source: VEST
	  ○ https://dataverse.harvard.edu/file.xhtml?fileId=4366198&version=56.0
	● VEST MO 16 documentation file
	  ○ Accessed: 05/06/21, Source: VEST
	  ○ https://dataverse.harvard.edu/file.xhtml?fileId=4499004&version=56.0
	● Precinct-Level Election Results
	  ○ Accessed: 05/06/21, Source: Open Elections
	  ○ https://github.com/openelections/openelections-data-mo/tree/master/2016
	● U.S. Census Bureau 2020 Redistricting Data Program Phase 2 release
	  ○ Accessed: 03/03/21, Source: Census
	  ○ https://www.census.gov/geo/partnerships/pvs/partnership19v2/st29_mo.html
	● 2010 Census VTD release
	  ○ Accessed: 03/03/21, Source: Census
	  ○ https://www.census.gov/cgi-bin/geo/shapefiles/index.php?year=2010&layergroup=Voting+Districts
	  ○ Note: Monroe County is the only county from this source.
	● 2020 Census VTD release
	  ○ Accessed: 03/03/21, Source: Census

- https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.html
- Note: Platte County is the only county from this source.
- Camden County Precincts (tif)
  - Accessed: 03/09/21, Source: County GIS Platform
  - https://camdengis.integritygis.com/H5/Index.html?viewer=camden
  - Note: Unable to easily load this file due to the format.
- Cooper County Precincts (tif)
  - Accessed: 03/09/21, Source: County GIS Platform
  - https://coopergis.integritygis.com/H5/Index.html?viewer=cooper
  - Note: Unable to easily load this file due to the format.
- Marion County Precincts (tif)
  - Accessed: 03/09/21, Source: County GIS Platform
  - https://mariongis.integritygis.com/H5/Index.html?viewer=marion_public
  - Note: Unable to easily load this file due to the format.
- Lafayette County Precincts (tif)
  - Accessed: 03/09/21, Source: County GIS Platform
  - https://lafayettegis.integritygis.com/H5/Index.html?viewer=lafayette
  - Note: Unable to easily load this file due to the format.
- Laclede County Precincts (tif)
  - Accessed: 03/09/21, Source: County GIS Platform
  - https://lacledegis.integritygis.com/H5/Index.html?viewer=laclede
  - Note: Unable to easily load this file due to the format.
- Bates County Precincts (tif)
  - Accessed: 03/09/21, Source: County GIS Platform
  - https://batesgis.integritygis.com/H5/Index.html?viewer=bates
  - Note: Unable to easily load this file due to the format.
- Audrain County Precincts (tif)
  - Accessed: 03/03/21, Source: County GIS Platform
  - https://audraingis.integritygis.com/H5/Index.html?viewer=audrain)
  - Note: Unable to easily load this file due to the format.
- File: Jasper County Precincts (shapefile)
  - Accessed: 03/09/21, Source: County GIS Platform
  - https://jaspercountymogisintiatives-jcmo.hub.arcgis.com/datasets/voting-precincts-2020)
  - Note: Able to load this file.

Inaccessible files:
- VEST used a variety of different sources for shapefiles. For some of the shapefiles that were sourced from county governments, we were unable to locate them. These counties were: Caldwell, Callaway, Cape Girardeau,

Cedar, Franklin, Greene, Jackson, Jefferson, Lawrence, McDonald, Miller, Nodaway, Osage, Ozark, Pemiscot, Pike, Randolph, Scott, Ste. Genevieve, Texas, Warren, Washington, Worth, Wright. For these counties, we used the census redistricting phase 2 files in their place to attempt to validate the data.

Note #1: Of the shapefiles sourced from counties, which we were able to locate (Camden, Cooper, Marion, Lafayette, Laclede, Jasper, Bates, Audrain), only Jasper came in shapefile format. The others came in tif formats that were unable to load into the notebook. For these counties, we also used the census redistricting phase 2 files in their place to attempt to validate the data.

Note #2: As part of their shapefile modifications, VEST uses a number of local boundary files as well as a voter file. Because VEST does not list the specific files used and any information about where to find them, we did not attempt to locate these files and reproduce these steps.

Processing steps available?
 	
Yes
	Description of processing steps:
	● VEST's full documentation can be accessed at the link above. The documentation we used was accessed on May 6th.
	● VEST processing steps:
	  ○ Election results:
	    ■ Listed the source for the results (Open Elections) and mentioned that absentee, provisional, and mail ballot were reported in almost every county and allocated by the share of precinct-level reported vote.
	    ■ Provided a list of counties where precinct mergers were made to match reporting units.
	  ○ Shapefiles:
	    ■ Described the four different sources for shapefiles: (U.S. Census Bureau's 2020 Redistricting Data Program Phase 2 release, 2010 Census VTD release, 2020 Census VTD release, and maps or shapefiles sourced from the respective county governments).
	    ■ Provided a list of counties where precinct mergers were made to match reporting units.
	    ■ Provided a list of additional modifications made to match reporting units.
	Information not in their processing steps:

- What the names are for absentee, provisional and mail ballot reported votes. Although we were able to do this on our own, because of the variety of naming conventions, this was not trivial.
- Specifics on how they rounded when allocating the totals, as we allocated votes using precinct-level reported votes, but had some slight differences, most likely due to different degrees of precision before rounding.
- Name changes to precincts in their final file, and to join with shapefiles.
- A list of which election precincts or shapefiles were merged together to produce their final file.
- Explanations for vote differences in Butler and Linn County.

Able to replicate joining election data and shapefiles? No	No, given the documentation they provided we would not be able to follow their steps to complete the join. However, we were able to join 3180 out of 3273 precincts that are in VEST's file. Of the precincts that did not join: • 93 of them contained just election results (51 of these in the counties where we used a different shapefile source) • 159 of them contained just shapefile data (89 of these in the counties where we used a different shapefile source) In order to perform these joins, significant shapefile name changes and mergers occurred, which all can be viewed in our public notebook.
Able to replicate joining demographic data to block-level shapefiles? N/A	There is no demographic data on the file.
Able to replicate joining boundary data?	There is no boundary data on the file.

N/A	

| **Successfully validated election results?**

Options:
No | <u>Election results:</u>
● For the election results, the race totals were equal, as well as the race totals for every county.
● In comparing the vote totals precinct-by-precinct, there were 3273 total precincts.
  ○ 835 of these precincts had election result differences
  ○ 2438 of these precincts were exactly the same
● However, the vast majority of these differences were 1 vote, which likely occurred due to the rounding that took place when absentee ballots were allocated. Excluding these, there are:
  ○ 35 precincts with election result differences greater than 1 vote.
● There are two groups for these 35 precincts:
  ○ G18ATGDHEN votes in Butler County. In 023Broseley, VEST has 302 ATGDHEN votes, whereas the source file has 0. In examining the Open Elections data, it appears as though 292 ATGDHEN Broseley votes were put in a "WRITE-IN" precinct, which we included in the countywide vote allocation. As such, all ATGDHEN values in Butler County are slightly off as we had more absentee votes than VEST, although the totals were the same.
  ○ Something similar to this occurs in LTGLHED votes in Linn county (115), as there are 12 votes for a particular candidate that were given to a "WRITE-IN" precinct and then were allocated countywide. In addition, some of VEST's votes are slightly different for this race, although it isn't as clear why.
● The Python notebook prints out the particular race and precinct where any different values occur.

<u>Geographies:</u>
● As mentioned above, we were not always able to use the same shapefiles that VEST used in their file. In addition, we were not able to replicate many of VEST's shapefile modifications, because they used files we could not easily access including a voter file and various local files. As such, we did not expect the geographies to match exactly.
● The counties were we tried to use different sources were: Audrain, Bates, Caldwell, Callaway, Camden, Cape Girardeau, Cedar, Cooper, Franklin, Greene, Jackson, Jefferson, Laclede, Lafayette, Lawrence, Marion, |

McDonald, Miller, Nodaway, Osage, Ozark, Pemiscot, Pike, Randolph, Scott, Ste. Genevieve, Texas, Warren, Washington, Worth, Wright.

- There were 3,180 precincts that we were able to match:
  - 106 precincts w/ a difference of 0 km^2
  - 2675 precincts w/ a difference between 0 and .1 km^2
  - 185 precincts w/ a difference between .1 and .5 km^2
  - 67 precincts w/ a difference between .5 and 1 km^2
  - 45 precincts w/ a difference between 1 and 2 km^2
  - 40 precincts w/ a difference between 2 and 5 km^2
  - 62 precincts w/ a difference greater than 5 km^2
- Of the 147 shapefiles with a diff > 1km^2 75 of them are in the counties where we used a different source.