**Block Group Voting**

**Problem and Opportunity**

In the United States, voting is largely a private matter. A registered voter is given a randomized ballot form or machine to prevent linkage between their voting choices and their identity. This disconnect supports confidence in the election process, but it provides obstacles to an election's analysis. A common solution is to field exit polls, interviewing voters immediately after leaving their polling location. This method is rife with bias, however, and functionally limited in direct demographics data collected.

For the 2020 general election, though, most states published their election results for each voting location. These publications were additionally supported by the geographical areas assigned to each location, the voting precincts. As a result, geographic processing can now be applied to project precinct election results onto Census block groups. While precinct have few demographic traits directly, their geographies have characteristics that make them projectable onto U.S. Census geographies. Both state voting precincts and U.S. Census block groups:

* are exclusive, and do not overlap
* are adjacent, fully covering their corresponding state and potentially county
* have roughly the same size in area, population and voter presence

Analytically, a projection of local demographics does not allow conclusions about voters themselves. However, the dataset does allow statements related to the geographies that yield voting behavior. One could say, for example, that an area dominated by a particular voting pattern would have mean traits of age, race, income or household structure.

The dataset that results from this programming provides voting results allocated by Census block groups. The block group identifier can be joined to Census Decennial and American Community Survey demographic estimates.

* The dataset is hosted by Harvard's dataverse as <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/NKNWBX>
* The publication may be cited and shared using the Digitial Object Identification System <https://www.doi.org/> as 10.7910/DVN/NKNWBX
* The software and documentation are publicly available on <https://github.com/OpenEnvironments/blockgroupvoting>

**Data Sources**

The state election results and geographies have been compiled by [Voting and Election Science team](https://dataverse.harvard.edu/dataverse/electionscience) on Harvard's dataverse. State voting precincts lie within state and county boundaries.

The Census Bureau, on the other hand, publishes its estimates across a variety of geographic definitions including a hierarchy of states, counties, census tracts and block groups. [Their definitions can be found here](https://www.census.gov/programs-surveys/geography/about/glossary.html#par_textimage_4). The geometric shapefiles for each block group [are available here](https://www2.census.gov/geo/tiger/TIGER2021/BG/).

The lowest level of this geography changes often and can obsolesce before the next census survey (Decennial or American Community Survey programs). The second to lowest census level, block groups, have the benefit of both granularity and stability however. The 2020 Decennial survey details US demographics into 217,740 block groups with between a few hundred and a few thousand people.

**Dataset Structure**

The dataset's columns include:

|  | **Column** | **Definition** |
| --- | --- | --- |
|  | BLOCKGROUP\_GEOID | 12 digit primary key. Census GEOID of the block group row. This code concatenates: |
|  |  | 2 digit state |
|  |  | 3 digit county within state |
|  |  | 6 digit Census Tract identifier |
|  |  | 1 digit Census Block Group identifier within tract |
|  | STATE | State abbreviation, redundent with 2 digit state FIPS code above |
|  | REP | Votes for Republican party candidate for president |
|  | DEM | Votes for Democratic party candidate for president |
|  | LIB | Votes for Libertarian party candidate for president |
|  | OTH | Votes for presidential candidates other than Republican, Democratic or Libertarian |
|  | area | square kilometers of area associated with this block group |
|  | gap | total area of the block group, net of area attributed to voting precincts |
|  | precincts | Number of voting precincts that intersect this block group |

**Assumptions, Notes and Concerns:**

* Votes are attributed based upon the proportion of the precinct's area that intersects the corresponding block group. Alternative methods are left to the analyst's initiative.
* 50 states and the District of Columbia are in scope as those U.S. possessions voting in the general election for the U.S. Presidency.
* Three states did not report their results at the precinct level: South Dakota, Kentucky and West Virginia. A dummy block group is added for each of these states to maintain national totals. These states represent 2.1% of all votes cast.
* Counties are commonly coded using FIPS codes. However, each election result file may have the county field named differently. Also, three states do not share county definitions - Delaware, Massachusetts, Alaska and the District of Columbia.
* Block groups may be used to capture geographies that do not have population like bodies of water. As a result, block groups without intersection voting precincts are not uncommon.
* In the U.S., elections are administered at a state level with the Federal Elections Commission compiling state totals against the Electoral College weights. The states have liberty, though, to define and change their own voting precincts <https://en.wikipedia.org/wiki/Electoral_precinct>.
* The Census Bureau practices "data suppression", filtering some block groups from demographic publication because they do not meet a population threshold. This practice is done to maintain statistical reliability in the estimates and to prevent accidental disclosure of individual respondents. As a result,
  + the shape files for state block groups may have additional block groups not available in demographic estimates.
  + ignoring the suppressed block groups will cause statistical bias for these smallest geographies
* As written, this projection takes more than 6 days to complete on a familiar Intel-64 based laptop. Its performance would benefit from:
  + Running states in parallel rather than serially
  + Looking for intersecting precincts within the shared county rather than state level
* Allocation details causes challenges in efforts to tie totals to state and national summaries. By allocating each of 233,866 detailed block groups based on area, many double precision proportions area applied to original integer counts. The allocations themselves then are not integers and may not sum exactly to the reported state election counts.

**RECOGNITION**

Special thanks to the meticulous efforts of:

The Voting and Election Science Team (University of Florida, Wichita State University) (<https://dataverse.harvard.edu/dataverse/electionscience>)

@data{DVN/K7760H\_2020, author = {Voting and Election Science Team}, publisher = {Harvard Dataverse}, title = {{2020 Precinct-Level Election Results}}, year = {2020}, version = {V29}, doi = {10.7910/DVN/K7760H}, url = {[https://doi.org/10.7910/DVN/K7760H}](https://doi.org/10.7910/DVN/K7760H%7D) }

MIT's Election Data and Science Lab MEDSL (<https://dataverse.harvard.edu/dataverse/medsl>

“U.S. Census TIGER/Line Files for Block Groups 2021.” Index of /Geo/Tiger/TIGER2021/BG, 22 Sept. 2021, <https://www2.census.gov/geo/tiger/TIGER2021/BG/>.

**License**

This code is available subject to the [MIT Open Source License](https://choosealicense.com/licenses/mit/)

**Summary Statistics**

| **State** | **Republican** | **Democrat** | **Libertarian** | **Other** | **Precincts** | **Block Groups** |
| --- | --- | --- | --- | --- | --- | --- |
| AL | 1,441,170 | 849,624 | 25,176 | 7,312 | 1,972 | 3,925 |
| AK | 189,951 | 153,778 | 8,897 | 4,943 | 441 | 504 |
| AZ | 1,661,686 | 1,672,143 | 51,465 | 0 | 1,489 | 4,773 |
| AR | 760,647 | 423,932 | 13,133 | 21,357 | 2,591 | 2,294 |
| CA | 6,006,428 | 11,110,493 | 187,907 | 192,232 | 20,799 | 25,607 |
| CO | 1,364,607 | 1,804,352 | 52,460 | 35,561 | 3,215 | 4,058 |
| CT | 714,717 | 1,080,831 | 20,230 | 8,079 | 741 | 2,716 |
| DE | 200,603 | 296,268 | 5,000 | 2,139 | 434 | 706 |
| DC | 18,586 | 317,323 | 2,036 | 6,411 | 144 | 571 |
| FL | 5,668,731 | 5,297,045 | 70,324 | 54,769 | 6,010 | 13,388 |
| GA | 2,461,837 | 2,474,507 | 62,138 | 0 | 2,679 | 7,446 |
| HI | 196,864 | 366,130 | 5,539 | 5,936 | 262 | 1,083 |
| ID | 554,119 | 287,021 | 16,404 | 9,737 | 935 | 1,284 |
| IL | 2,446,891 | 3,471,915 | 66,544 | 48,088 | 10,083 | 9,898 |
| IN | 1,729,857 | 1,242,498 | 58,901 | 0 | 5,166 | 5,290 |
| IA | 897,672 | 759,061 | 19,637 | 14,501 | 1,661 | 2,703 |
| KS | 771,406 | 570,323 | 30,574 | 0 | 4,070 | 2,461 |
| LA | 1,255,776 | 856,034 | 21,645 | 14,607 | 3,753 | 4,294 |
| ME | 360,767 | 435,070 | 14,120 | 9,412 | 573 | 1,184 |
| MD | 976,414 | 1,985,023 | 33,488 | 42,106 | 2,043 | 4,079 |
| MA | 1,167,202 | 2,382,202 | 47,013 | 34,985 | 2,173 | 5,116 |
| MI | 2,649,859 | 2,804,036 | 60,406 | 23,907 | 4,756 | 8,386 |
| MN | 1,484,065 | 1,717,077 | 34,976 | 41,053 | 4,110 | 4,706 |
| MS | 756,764 | 539,398 | 8,026 | 9,571 | 1,764 | 2,445 |
| MO | 1,718,736 | 1,253,014 | 41,205 | 12,202 | 3,733 | 5,031 |
| MT | 343,602 | 244,786 | 15,252 | 0 | 666 | 900 |
| NE | 556,846 | 374,583 | 20,283 | 0 | 1,386 | 1,648 |
| NV | 669,890 | 703,486 | 14,783 | 17,217 | 2,094 | 1,963 |
| NH | 365,660 | 424,937 | 13,236 | 0 | 321 | 997 |
| NJ | 1,883,274 | 2,608,335 | 31,677 | 26,067 | 726 | 6,599 |
| NM | 401,894 | 501,614 | 12,585 | 7,872 | 1,917 | 1,614 |
| NY | 3,251,997 | 5,244,886 | 60,383 | 74,987 | 15,376 | 16,070 |
| NC | 2,758,775 | 2,684,292 | 48,678 | 33,059 | 2,662 | 7,111 |
| ND | 235,751 | 115,042 | 9,371 | 1,860 | 422 | 632 |
| OH | 3,154,834 | 2,679,165 | 67,569 | 18,812 | 8,941 | 9,472 |
| OK | 1,020,280 | 503,890 | 24,731 | 11,798 | 1,948 | 3,374 |
| OR | 958,448 | 1,340,383 | 41,582 | 33,908 | 1,331 | 2,970 |
| PA | 3,378,442 | 3,460,475 | 79,432 | 0 | 9,150 | 10,173 |
| RI | 199,922 | 307,486 | 5,053 | 5,296 | 423 | 792 |
| SC | 1,385,103 | 1,091,541 | 27,916 | 8,769 | 2,263 | 3,408 |
| TN | 1,852,475 | 1,143,711 | 29,877 | 26,926 | 1,962 | 4,562 |
| TX | 5,890,570 | 5,259,281 | 126,269 | 44,299 | 9,014 | 18,638 |
| UT | 865,140 | 560,282 | 38,447 | 42,773 | 2,424 | 2,020 |
| VT | 112,704 | 242,820 | 3,608 | 8,296 | 284 | 552 |
| VA | 1,962,430 | 2,413,568 | 64,761 | 19,765 | 2,477 | 5,963 |
| WA | 1,584,651 | 2,369,612 | 80,500 | 52,868 | 7,464 | 5,311 |
| WI | 1,610,184 | 1,630,866 | 38,491 | 18,500 | 7,090 | 4,692 |
| WY | 193,559 | 73,491 | 5,768 | 3,947 | 481 | 457 |
| Total | 72,091,786 | 80,127,630 | 1,817,496 | 1,055,927.00 | 166,419 | 233,866 |