Multi-County Vote Analysis

Each day, we manually download cumulative votes to date (early voting roster/activity) from the county in one or more csv files for the subject election. Each county has its own different process, content and format of providing accessibility to its files.

The content of what is in those early voting roster files varies significantly from county to county. There is consistency in INconsistency between the voter roll files and the voting activity files from within each county. Many of the votes do not have a consistent type of voter id assigned that creates a definitive relationship between the registered voters in the voter roll and the ballots cast in the voting activity files. Names provided by a county in the voter roll file are formatted differently than what is included in the voting activity files. Same for addresses. Names and addresses are NOT standardized as they are when displayed on a voter registration card. In short, it seems the counties, individually, go to extra efforts to ensure that matching the information in the voter activity files and the voter roll files is difficult for the public, without a benefit to the voter or the public.

Counties do not provide unique identifiers of a vote, e.g., ballot number or pseudo-assigned ballot number. (Voter rolls contain some number of multiple voters also.) Simply, there is NO reliable relational integrity in the data provided by the counties. What the counties do provide for early voting roster/activity is a "header" of sorts that minimally contains a voter name, voter id number (theoretically the state-assigned one, but this is questionable based on the data) and some indicator of the election (Bexar county provides the least info here – may not comply with Texas Election Code Section 87.121[1]). Data provided by Dallas and Harris county csv files are detailed in Appendix A.

---

1    Texas Election Code Section 87.121 Early voting rosters
(a) The early voting clerk shall maintain for each election a roster listing each person who votes an early voting ballot by personal appearance and a roster listing each person to whom an early voting ballot to be voted by mail is sent.

(b) For each person listed, the applicable roster must include:

(1) the person's name, address, and voter registration number;
(2) an identification of the person's county election precinct of registration; and
(3) the date of voting or the date the ballot was mailed to the person, as applicable.
(c) Each roster shall be updated daily.

(d) Each roster may be maintained in any form approved by the secretary of state.

(e) The clerk shall preserve each roster after the election for the period for preserving the precinct election records.

(f) Information on the roster for a person to whom an early voting mail ballot has been sent is not available for public inspection, except to the voter seeking to verify that the information pertaining to the voter is accurate, until the first business day after election day.

(g) Information on the roster for a person who votes an early voting ballot by personal appearance shall be made available for public inspection not later than the beginning of regular business hours on the day after the date the information is entered on the roster under Subsection (c).

(h) Information on the roster for a person who votes an early voting ballot by mail shall be made available for public inspection not later than the day following the day the early voting clerk receives a ballot voted by mail.

How does one define a uniquely cast vote?  When is a voting record final?  Simple questions, not so simple answers.  For analysis purposes all the data provided by the county for each row in the voter roll files and early voting roster csv files is effectively concatenated together to determine its historical uniqueness.

For analysis purposes there are two ways a vote is "purged".
1. <u>Purged and Re-instated with Changes/New Ballot</u>:  If anything in the csv file data row is changed when reported in a subsequent cumulative voter roll or early voting roster provided by the county (a "re-instated" ballot/new ballot) it is presumed for analysis purposes that the originally reported vote row is discarded/purged.  We have no way of knowing if the new row reported by the county is a change to the previously cast ballot or a new ballot.  Valid reasons for anything re-instating/changing for a cast ballot are unknown at this time.  Sample changes/diferences in new ballots versus the related purged ballots (for Dallas County) are concentrated in the following categories:
   a) **voter name**
   b) voter address
   c) **date ballot requested**
   d) **ballot type**
   e) **ballot mailed date**
   f) **ballot returned date**
2. <u>Purged Only</u>:  Any row that is no longer reported in subsequent cumulative files.  Valid reasons for votes to be purged include unused mail-in ballots surrendered by a voter so that they can vote in person.  No specific reason or explanation are provided by the county for votes no longer included in subsequent county-provided files.

Typical county disclaimer for csv files containing daily voting activity:

**Disclaimer**: The early voting election rosters are unofficial list of voters who vote an early voting ballot by personal appearance and voters whom an early voting ballot to be voted by mail is received. Each roster is updated daily.

**Please Note**: The reports provided on this page are considered Unofficial until after Election Day reconciliation of all voters to include voters that cast a ballot by mail, early in person, and/or Election Day. Once the election has been canvassed and reconciled a final official list of voters will be posted.

The voter id columns provided by the county in the daily early voting roster files are named differently from those found in the Voter Roll files. The column names for the loading and reporting tables are identical, but the indexes are different within the general categories of voter roll and voting activity. Name and street address data has been parsed into the base components. This could be improved somewhat by using a USPS standardized address API and/or geocoding.

Harris County Data
Daily Voting Activity files are from: https://harrisvotes.com/ElectionRosters
Voter Roll files are from: https://www.hctax.net/Voter/PrecinctData

Dallas County Data
Daily Voting Activity files are from: https://www.dallascountyvotes.org/early-voting-reports/?election_code=1120-00#EarlyVoting
Voter Roll files are not available online

The general process so far is:
1. Create Voter Roll loading table for loading data into the database from csv files and insert original county data, plus augmented data. No data is intentionally discarded in this process.
2. Create Voter Roll reporting table based upon data from loading table. Duplicate registered voters (all data across all columns are identical) are removed (about 20,630 for Harris County in the current run, smaller number for Dallas County)
3. Create Early Voting activity table for loading daily voting activity from csv files and insert original county data, plus augmented data. No data is intentionally discarded in this process, including known cumulative data.
4. Create Early Voting activity table for reporting purposes based upon data from loading table.

# Appendix A

Dallas and Harris County CSV File Column Names

| Dallas | Harris |
|---|---|
| Precinct<br>PrecinctSub<br>StateIDNumber<br>VoterName<br>BallotType<br>DateBallotRequestedChar<br>DateBallotMailedChar<br>DateBallotReturnedChar<br>ResidenceAddress<br>ResidenceCity<br>ResidenceZip<br>PartyCode<br>ElectionCode<br>VoterCityCode<br>VoterCityName<br>VoterISDCode<br>VoterISDName<br>InPersonVotingLocation<br>VoterUSCongressCode<br>VoterUSCongressName<br>VoterUSStSenateCode<br>VoterUSStSenateName<br>VoterStRepCode<br>VoterStRepName<br>VoterCommissionerCode<br>VoterCommissionerName<br>VoterCitySingleMemberCode<br>VoterCitySingleMemberName<br>VoterISDSingleMemberCode<br>VoterISDSingleMemberName<br>VoterWaterDistCode<br>VoterWaterDistName<br>VoterDCCCDCode<br>VoterDCCCDName<br>VoterFloodControlCode<br>VoterFloodControlName<br>VoterCountyJPCode<br>VoterCountyJPName<br>VoterSBOECode<br>VoterSBOEName | ElectionCode<br>IDNumber<br>StateVoterID<br>Party<br>VoterName<br>VoterAddress<br>VoterCity<br>VoterZip<br>Precinct<br>EarlyVotingSite<br>ActivityDateChar<br>Dist02<br>Dist03<br>Dist04<br>Dist05<br>Dist06<br>Dist07<br>Dist08<br>Dist09<br>Dist10<br>Dist11<br>Dist12<br>Dist13<br>Dist14<br>Dist15<br>Dist16<br>Dist17<br>Dist18<br>Dist19<br>Dist20<br>Dist21<br>Dist22<br>Dist23<br>Dist24<br>Dist25<br>Dist26<br>Dist27<br>Dist28<br>Dist29<br>Dist30<br>Dist31<br>Dist32<br>Dist33 |

# Appendix B – Miscellaneous

What does the term 'Suspense' mean on a voter's record of status? Pursuant to Sec.15.111, 63.0011 of the Election Code, the term 'suspense' is used to designate those voters that the Dallas County Voter Registrar does not have a current address. A voter with this designation will be removed from the suspense list once that person has advised us of his or her new address. This new address is important, not only so that our records will be accurate, but also so that the voter is voting in the elections where he or she resides.

If you are a suspense voter and would like to update your address you can now update it through the Secretary of State portal. Please click below to be directed to that page.

https://txapps.texas.gov/tolapp/sos/SOSACManager