Data Export

Colorado Risk Limiting Audit Tool

2017

Table of Contents

Separate from the RLA central server and client server programs that support the Department of State and the Counties in carrying out the Risk Limiting Audit, there is a command, called rla\_export, allowing export of data from the central server and the underlying database.

The command is part of a python package, whose technical description can be found in a README.rst file in the python site-packages directory tree wherever the package is installed. The most current online version (which may or may not match the version you have installed) is available in the [public code repository](https://github.com/FreeAndFair/ColoradoRLA/tree/auditcenter/server/eclipse-project/script/rla_export).

The README.rst file gives instructions for installing the python package, and describes various run-time options. For a catalog of the exports produced by the command, see below.

## Database Exports

The rla\_export command exports many of the files necessary for independent verification of the RLA, whether by candidates, parties, other organizations. Specific exports are detailed below.

### Minimum Data Required to Allow Public to Reproduce Audit Calculations

To allow independent verification of the RLA, the Colorado Department of State must provide to the public all the export files listed in this section, along with:

* all CVR files
* all ballot manifest files
* opportunity to observe the random seed selection
* opportunity to observe the activities of the County Audit Boards
* announced tabulated results for contests selected for audit
* the risk limit
* the error inflation factor

Providing the above to the public, along with the database exports listed below will fulfill the minimum requirements of a publicly-verifiable audit.

The data files in this section are generated based on sql query files. These are always produced in two formats: json and csv. The basename of each resulting file is the same as the basename of the query file. Thus, given the query file seed.sql, the files seed.json and seed.csv will be produced.

#### m\_selected\_contest\_audit\_details\_by\_cvr.sql

For each contest under audit, and for each cast vote record presented to the Audit Board for verification that contains the given contest, the RLA system's record of the Audit Board's review of the physical ballot for that contest

|  |  |  |
| --- | --- | --- |
| Field | Type | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_Meaning\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |
| county\_name | Text String | Name of County |
| contest\_name | Text String | Name of contest |
| random\_sequence\_index | Integer | Index in the random sequence (starting with 1) |
| imprinted\_id | Text String | combination of scanner, batch and record ids that uniquely identifies the ballot card and may be imprinted on the card |
| ballot\_type | Text String | BallotType from Dominion CVR export file, a code for the set of contests that should be present on the physical ballot card |
| choice\_per\_voting\_computer | List of Text Strings | List of voter choices in the given contest on the given ballot card, as interpreted by the vote-tabulation computer system (note: overvotes recorded as blank votes) |
| choice\_per\_audit\_board | List of Text Strings | List of voter choices in the given contest on the given ballot card, as interpreted by the Audit Board (note: overvotes recorded as a too-long list of choices) |
| did\_audit\_board\_agree | YES/NO | "YES" if the Audit Board came to consensus on the interpretation of the given ballot card; "NO" if not; blank if the card has not been reviewed by the Audit Board. |
| audit\_board\_comment | Text String | Text of comment entered by Audit Board about the given contest on the given ballot card, or indication that the ballot was not found. |
| timestamp | Timestamp | Date and time of Audit Board's submission of their interpretation to the RLA Tool |

#### m\_ballot\_list\_for\_review.sql

List of ballot cards assigned to Audit Board for review. (This list could be created from the random sequence by removing duplicates and ordering by tabulator, batch and position within the batch.) Within each county, the list is ordered by rounds and, within each round, by tabulator, batch and position within the batch.

|  |  |  |
| --- | --- | --- |
| Field | Type | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_Meaning\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |
| county\_name | Text String | Name of County |
| round | Integer | The audit round number in which the ballot card is assigned to the given County's Audit Board for review. |
| scanner\_id | Integer | TabulatorNum from Dominion CVR export file, identifying the tabulator used to read the physical ballot card |
| batch\_id | Integer | BatchId from Dominion CVR export file, identifying the batch of physical ballot cards in which the card was scanned |
| record\_id | Integer | RecordId from Dominion CVR export file, indicating the position of the card in its batch of physical ballot cards |
| imprinted\_id | Text String | combination of scanner, batch and record ids that uniquely identifies the ballot card and may be imprinted on the card |
| ballot\_type | Text String | BallotType from Dominion CVR export file, a code for the set of contests that should be present on the physical ballot card |

#### m\_selected\_contest\_static.sql

List of contests selected by the Secretary of State for audit, with information about the contest that doesn't change during the audit, namely the reason for the audit, the number of winners allowed in the contest, the tabulated winners of the contest, the numbers of ballots cards recorded as cast in the county (total number as well as the number containing the given contest) and the value of the error inflation factor (gamma).

|  |  |  |
| --- | --- | --- |
| Field | Type | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_Meaning\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |
| county\_name | Text String | Name of County |
| contest\_name | Text String | Name of contest |
| audit\_reason | Text String | Reason for audit ("STATE\_WIDE\_CONTEST" and "COUNTY\_WIDE\_CONTEST" refer to the types of contests that must be chosen to drive the audit, per Rule 25.2.2(i)) |
| votes\_allowed | Integer | Maximum number of choices that can be recorded on one ballot in the given contest |
| winners\_allowed | Integer | Number of winners allowed for the contest (required to calculate diluted margin) |
| winners | List of Text Strings | List of all winners of the given contest in the given County. (Note that for multi-county contests this list includes the highest vote-getters within the County, even if these were not the winners across all Counties.) |
| min\_margin | Integer | The smallest margin between any winner and any loser |
| county\_ballot\_card\_count | Integer | The number of ballot cards recorded in the given County in the election (including cards that do not contain the contest in question) |
| contest\_ballot\_card\_count | Integer | The number of ballot cards recorded in the given County that contain the contest in question |
| gamma | Number | Error inflation factor defined in Stark's paper, Super-simple simultaneous single-ballot risk-limiting audits, which is cited in Lindeman and Stark's paper, A Gentle Introduction to Risk Limiting Audits, which is cited in Rule 25.2.2(j)) |

#### m\_selected\_contest\_dynamic.sql

List of contests selected by Secretary of State for audit, with current status. Which contests has the Secretary selected for audit? Which contests (if any) has the Secretary selected for hand count? How many discrepancies of each type?

|  |  |  |
| --- | --- | --- |
| Field | Type | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_Meaning\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |
| county\_name | Text String | Name of County |
| contest\_name | Text String | Name of contest |
| current\_audit\_type | Text String | "COMPARISON", "HAND\_COUNT", "NOT\_AUDITABLE" or "NONE |
| computerized\_audit\_status | Text String | "NOT\_STARTED", "NOT\_AUDITABLE", "IN\_PROGRESS" or "ENDED". |

Because declaring a hand count ends the computerized portion of the audit, a contest that is being hand-counted will have the value "ENDED" in this field. one\_vote\_over\_count | Integer | The number of ballot cards in the random sequence so far (with duplicates) on which there is a one-vote overstatement (per Lindeman & Stark's A Gentle Introduction to Risk Limiting Audits). one\_vote\_under\_count | Integer | The number of ballot cards in the random sequence so far (with duplicates) on which there is a one-vote understatement (per Lindeman & Stark's A Gentle Introduction to Risk Limiting Audits). two\_vote\_over\_count | Integer | The number of ballot cards in the random sequence so far (with duplicates) on which there is a two-vote overstatement (per Lindeman & Stark's A Gentle Introduction to Risk Limiting Audits). two\_vote\_under\_count | Integer | The number of ballot cards in the random sequence so far (with duplicates) on which there is a two-vote understatement (per Lindeman & Stark's A Gentle Introduction to Risk Limiting Audits).

#### m\_cvr\_hash.sql

Hashes of CVR files

|  |  |  |
| --- | --- | --- |
| Field | Type | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_Meaning\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |
| county\_name | Text String | Name of County |
| hash | Text String | Hash value entered by the given county after uploading the cast vote record file to be used in the audit |

#### m\_manifest\_hash.sql

Hashes of ballot manifest files

|  |  |  |
| --- | --- | --- |
| Field | Type | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_Meaning\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |
| county\_name | Text String | Name of County |
| hash | Text String | Hash value entered by the given county after uploading the ballot manifest file |

to be used in the audit

#### m\_random\_sequence.sql

Random sequence of ballot cards used for the audit. (This random sequence is generated with replacement and thus may include duplicates.)

|  |  |  |
| --- | --- | --- |
| Field | Type | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_Meaning\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |
| county\_name | Text String | Name of County |
| review\_index | Integer | The position in the random sequence for the given County |
| scanner\_id | Integer | TabulatorNum from Dominion CVR export file, identifying the tabulator used to read the physical ballot card with the given review-index |
| batch\_id | Integer | BatchId from Dominion CVR export file, identifying the batch of physical ballot cards in which the card with the given review-index was scanned |
| record\_id | Integer | RecordId from Dominion CVR export file, indicating the position of the card with the given review-index in its batch of physical ballot cards |
| imprinted\_id | Text String | Combination of scanner, batch and record ids that uniquely identifies the ballot card with the given review-index and may be imprinted on the card |
| ballot\_type | Text String | BallotType from Dominion CVR export file, a code for the set of contests that should be present on the physical ballot card with the given review-index |

#### m\_tabulate.sql

|  |  |  |
| --- | --- | --- |
| Field | Type | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_Meaning\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |
| county\_name | Text String | Name of County |
| contest\_name | Text String | Name of contest |
| choice | Text String | Name of candidate or for a ballot question "Yes" or "No" |
| votes | Integer | Number of votes recorded for the given choice in the given contest in the given County |

### Other Data Exports

The exports in this section, while not strictly necessary for independent verification of the audit calculations, will be of interest and value to the public.

#### all\_contest\_static.sql

List of all contests, with information about the contest that doesn't change during the audit, namely the reason for the audit, the number of winners allowed in the contest, the tabulated winners of the contest, the numbers of ballots cards recorded as cast in the county (total number as well as the number containing the given contest) and the value of the error inflation factor (gamma).

|  |  |  |
| --- | --- | --- |
| Field | Type | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_Meaning\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |
| county\_name | Text String | Name of County |
| contest\_name | Text String | Name of contest |
| audit\_reason | Text String | Reason for audit ("STATE\_WIDE\_CONTEST" and "COUNTY\_WIDE\_CONTEST" refer to the types of contests that must be chosen to drive the audit, per Rule 25.2.2(i); other reasons from the Rule include CLOSE\_CONTEST, TIED\_CONTEST, GEOGRAPHICAL\_SCOPE, CONCERN\_REGARDING\_ACCURACY, and COUNTY\_CLERK\_ABILITY; the audits of other contests on the ballot are OPPORTUNISTIC\_BENEFITS) |

votes\_allowed | Integer | Maximum number of choices that can be recorded on one ballot in the given contest winners\_allowed | Integer | Number of winners allowed for the contest (required to calculate diluted margin) winners | List of Text Strings | List of all winners of the given contest in the given County. (Note that for multi-county contests this list includes the highest vote-getters within the County, even if these were not the winners across all Counties.) min\_margin | Integer | The smallest margin between any winner and any loser county\_ballot\_card\_count | Integer | The number of ballot cards recorded in the given County in the election (including cards that do not contain the contest in question) contest\_ballot\_card\_count | Integer | The number of ballot cards recorded in the given County that contain the contest in question gamma | Number | Error inflation factor defined in Stark's paper, Super-simple simultaneous single-ballot risk-limiting audits, which is cited in Lindeman and Stark's paper, A Gentle Introduction to Risk Limiting Audits, which is cited in Rule 25.2.2(j))

#### all\_contest\_dynamic.sql

List of contests with current status. Which contests has the Secretary selected for audit? Which contests (if any) has the Secretary selected for hand count? How many discrepancies of each type have been found so far?

|  |  |  |
| --- | --- | --- |
| Field | Type | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_Meaning\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |
| county\_name | Text String | Name of County |
| contest\_name | Text String | Name of contest |
| current\_audit\_type | Text String | Comparison audit, ballot polling audit or hand count |
| computerized\_audit\_status | Text String | Not started, in progress, risk limit achieved, or ended. Because declaring a hand count ends the computerized portion of the audit, a contest that is being hand-counted will have the value "ended" in this field. |
| one\_vote\_over\_count | Integer | The number of ballot cards in the random sequence so far (with duplicates) on which there is a one-vote overstatement (per Lindeman & Stark's A Gentle Introduction to Risk Limiting Audits). |
| one\_vote\_under\_count | Integer | The number of ballot cards in the random sequence so far (with duplicates) on which there is a one-vote understatement (per Lindeman & Stark's A Gentle Introduction to Risk Limiting Audits). |
| two\_vote\_over\_count | Integer | The number of ballot cards in the random sequence so far (with duplicates) on which there is a two-vote overstatement (per Lindeman & Stark's A Gentle Introduction to Risk Limiting Audits). |
| two\_vote\_under\_count | Integer | The number of ballot cards in the random sequence so far (with duplicates) on which there is a two-vote understatement (per Lindeman & Stark's A Gentle Introduction to Risk Limiting Audits). |

#### all\_contest\_audit\_details\_by\_cvr.sql

For each contest, and for each cast vote record presented to the Audit Board for verification that contains the given contest, the RLA system's record of the Audit Board's review of the physical ballot for that contest

|  |  |  |
| --- | --- | --- |
| Field | Type | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_Meaning\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |
| county\_name | Text String | Name of County |
| contest\_name | Text String | Name of contest |
| random\_sequence\_index | Integer | Index in the random sequence (starting with 1) |
| imprinted\_id | Text String | combination of scanner, batch and record ids that uniquely identifies the ballot card and may be imprinted on the card |
| ballot\_type | Text String | BallotType from Dominion CVR export file, a code for the set of contests that should be present on the physical ballot card |
| choice\_per\_voting\_computer | List of Text Strings | List of voter choices in the given contest on the given ballot card, as interpreted by the vote-tabulation computer system (note: overvotes recorded as blank votes) |
| choice\_per\_audit\_board | List of Text Strings | List of voter choices in the given contest on the given ballot card, as interpreted by the Audit Board (note: overvotes recorded as a too-long list of choices) |
| did\_audit\_board\_agree | YES/NO | "YES" if the Audit Board came to consensus on the interpretation of the given ballot card; "NO" if not; blank if the card has not been reviewed by the Audit Board. |
| audit\_board\_comment | Text String | Text of comment entered by Audit Board about the given contest on the given ballot card, or indication that the ballot was not found. |
| timestamp | Timestamp | Date and time of Audit Board's submission of their interpretation to the RLA Tool |

#### auditboards.sql

|  |  |  |
| --- | --- | --- |
| Field | Type | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_Meaning\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |
| county\_name | Text String | Name of County |
| member | Text String | Name of audit board member |
| sign\_in\_time | Timestamp | Beginning of an audit board member's RLA Tool session |
| sign\_out\_time | Timestamp | End of the given session for the given audit board member |

#### prefix\_length.sql

|  |  |  |
| --- | --- | --- |
| Field | Type | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_Meaning\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |
| county\_name | Text String | Name of County |
| audited\_prefix\_length | Integer | Length of the longest prefix of the random sequence of cvr selections containing only cvrs that have been audited |

#### seed.sql

|  |  |  |
| --- | --- | --- |
| Field | Type | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_Meaning\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |
| seed | 20-Digit String | the random seed for the pseudo-random number generator |

#### upload\_status.sql

|  |  |  |
| --- | --- | --- |
| Field | Type | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_Meaning\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |
| county\_name | Text String | Name of County |
| filename | Text String | Name of file |
| hash\_status | Text String | "VERIFIED", "MISMATCH", or "NOT\_CHECKED" |
| approx\_count | Integer | Approximate number of lines in the file |
| size | Integer | Size of file in bytes |
| status | Text String | "IMPORTED\_AS\_BALLOT\_MANIFEST", "IMPORTED\_AS\_CVR\_EXPORT" or "NOT\_IMPORTED" |
| timestamp | Timestamp | Date and time of the most recent update to the upload status of the given file |

## State and County Audit Reports

Other export files are the same as the files available via the GUI interface, for example state\_report.xlsx.