# Election Audit Precinct Selection 2022

# Three Anomalies Proving It Was Not Random

**Prepared by:**
*Tim Ragsdale*
*M.S. Computer Systems and Software Engineering (Air Force Institute of Technology)*
*B.S. Computer Engineering (University of Arizona)*
*30+ Years Modeling Simulation & Analysis Experience*

**15 Mar 2024**

# TABLE OF CONTENTS

## *1.0 INTRODUCTION*

The *AGREED UPON PROCEDURES REPORT Evaluation of the Accuracy of Voting Machine Tabulators used for the 2022 General Election Held on November 8, 2022*[1] (**AUP22**) posted to the Secretary of State's website defines a specific process for selecting precincts for audit. This selection process for precinct audits is the sole focus of this paper. Furthermore, the scope of this paper is limited to evaluating the *randomness* of the results.

If the selection process is required to be random and it can be demonstrated that the results of the 2022 audit selection process was not random, then corrective actions by the Secretary of State of New Mexico need to be taken.

## *1.1 THE AUDIT SELECTION PROCESS*

The **AUP22** document spells out the administrators' process for conducting precinct selection. This process was precisely followed to recreate the needed data required to model the process.

1.1.1. "Obtained a listing of congressional districts by county precincts and downloaded the Statewide Voter Turnout and the Media file from the Secretary of State's website."

1.1.2. "We created a spreadsheet from the above data to cumulate the number of ballots cast to determine the population for sampling, beginning alphabetically with the first Bernalillo County precinct and ending with the last Valencia County precinct."

> Their spreadsheet resulted in records for 714,797 indexed voters. When the process was reproduced, the total was 714,754. This is likely due to voting results being updated since the selection was done and is not an area of concern.

1.1.3. "The required sample of precincts was chosen randomly using a 'dice roll' method whereby seven ten-sided dice were rolled, and the **resulting number normalized so that a particular voter (unnamed) could be**

[1] *AGREED UPON PROCEDURES REPORT, accessed 17 Feb 2024:* https://www.sos.nm.gov/voting-and-elections/voter-information-portal-nmvote-org/election-audits-2/

**identified within all the voting precincts in the state.**[2] "[*emphasis added*]

1.1.3.1. The normalization process here converts a rolled value [0...9,999,999] to select a particular voter index (position in the list) from the spreadsheet described in 1.1.2.

Though the dice rolls were publicly performed and the normalized values are provided in the report, *the process of selecting that voter's precinct is not provided* and therefore cannot be evaluated for accuracy.


This process of selecting a precinct number from a dice roll resides in a tool that has not been made available for inspection; therefore, the process is not transparent and observable as required.

1.1.3.2. The election results are used to determine the number of precincts to be selected for audit. For the 2022 selection, the following number of audits were required.

| 5 | 109 | 10 | 13 |
|---|---|---|---|
| Office Selected for Audit | | | |
| US Rep Dist 1 | US Rep Dist 2 | Gov. & Lt. Gov | SC Justice Position 1 |

[2] *Ibid*

1.1.3.3. The process is to roll the dice to select precincts until all of the requirements are met. Any rolls that result in a precinct that has already been selected or does not meet the requirements for any remaining audit (for example, a district 1 precinct while looking for district 2 precincts) are ignored.

### *1.2 Modeling the Process*

This process is simple to model. Java language scripts were used to read the precinct data into a list of precinct objects. The precinct objects are marked with the beginning/ending voter index for that precinct and the Java language's random number generator was iteratively used to select between 1 and 714,754 (see 1.1.2) until all of the audit requirements are met.

The pseudo-random number generator is more random than the use of dice because dice rolls are limited to one of 10 million discrete values, while the random number generator potential values are double precision floating point numbers that can be partitioned precisely across the population weighted precincts.

*A set of runs*: Begin a series of random voter draws, match to their precinct, check each audit requirement to see if it can fulfill it (incomplete and the correct district), and continue until all audit requirements are complete.

For comparison, three sets of 100 thousand sets of runs (equivalent of approximately 133 million dice rolls) were completed.

- Run Sets A and B weighted the precincts by cast votes (AUP22)
- Run Set C weighted the precincts by registered voters for comparison (using the same random seed as B)

***Only some of the statistics that follow rely on the modeling runs. Anywhere modeling is used it will be clearly marked "Modeling:"***

### *1.3 Recognizing the Problem*

Appendix 3 of the AUP22 (A-3) provides the values rolled, the normalized values, the precinct selected, and marks the precincts for audit that are a match. It is worth noting the first few rolls result in multiple audits for the selected precinct[3].

[3] *Recommended modification*: Do not fulfill audit requirements with the same sets of rolls. This disproportionately targets the precincts rolled first for multiple audits.

| Roll # | Dice 1 | Dice 2 | Dice 3 | Dice 4 | Dice 5 | Dice 6 | Dice 7 | Aggregation | Roll Number Normalized | Row Matched | County | Precinct # | US Rep District | US Rep Dist 1 | US Rep Dist 2 | Gov. & Lt. Gov | SC Justice Position 1 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 0 | 2 | 8 | 2 | 5 | 1002825 | 71,682 | 234 | Bernalillo | PCT 307 | DISTRICT 1 | X | | X | X |
| 2 | 2 | 4 | 6 | 6 | 9 | 3 | 7 | 2466937 | 176,336 | 581 | Bernalillo | PCT 380 | DISTRICT 1 | X | | X | X |
| 3 | 8 | 3 | 5 | 2 | 5 | 4 | 9 | 8352549 | 597,038 | 1,838 | Santa Fe | PRECINCT 026 | DISTRICT 3 | | | X | X |
| 4 | 4 | 3 | 1 | 0 | 3 | 4 | 2 | 4310342 | 308,102 | 913 | Dona Ana | PRECINCT 024 | DISTRICT 2 | | X | X | X |

With 2163 precincts to choose from and only 486 rolls (selections), the repeated (41) selections of Bernalillo 307 (BERN307) and others were indicators that further analysis was needed.

### *1.4 Anomaly One--Three in a Row*

BERN307 was selected three times in a row. It is important to note the weighting factor of the individual precinct and calculate its probability of being selected on an individual roll:

| In-a-row | registered | cast |
|---|---|---|
| 1 | 1:1995 | 1:1430 |
| 2 | 1:3.98 M | 1:2.0 M |
| 3 | 1:7.94 B | 1:2.9 B |

1:3.5 M "5-sigma"

***BERN307 cast votes: 500 out of 714,754. 500/714,754 = 0.00069954 or a 1 in 1430 chance.*** To be selected twice in a row, this value is squared and to be selected three times in a row, it is cubed. The chart above gives the one-in-X statistics for BERN307 if the precincts are weighted by registered voters or by cast votes. This anomaly alone proves[4] the result of the selection process was not random.

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 229 | 9 | 1 | 4 | 3 | 0 | 7 | 5 | 9143075 | 653,544 | 1,978 | Santa Fe | PRECINCT 166 | DISTRICT 3 |
| 230 | 1 | 8 | 0 | 4 | 9 | 9 | 9 | 1804999 | 129,021 | 234 | Bernalillo | PCT 307 | DISTRICT 1 |
| 231 | 1 | 0 | 0 | 5 | 9 | 2 | 8 | 1005928 | 71,904 | 234 | Bernalillo | PCT 307 | DISTRICT 1 |
| 232 | 0 | 8 | 9 | 3 | 1 | 4 | 5 | 0893145 | 63,842 | 234 | Bernalillo | PCT 307 | DISTRICT 1 |
| 233 | 5 | 3 | 5 | 9 | 1 | 8 | 6 | 5359186 | 383,073 | 1,185 | Lea | PRECINCT 029 | DISTRICT 3 |

1 in 2.9 Billion

---

[4] The line between "impossible" and "astronomically improbable" is a matter of mathematical and philosophical debate. The "5 sigma rule" is sometimes used as a threshold for "impossible." That corresponds to $3 \times 10^{-7}$ or about 1 in 3.5 million.

Ref: *https://blogs.scientificamerican.com/observations/five-sigmawhats-that/*

How Improbable is Anomaly One?

If you were locked in a room until you could flip a coin to the same face three times in a row, it would take a while but you'd get out. For four in a row you may give up because it would be twice as hard as three. And five would seem hopeless as it is twice as hard as four…

You'd have a better chance of getting 31 in a row than rolling BERN307 three times in a row in this process.

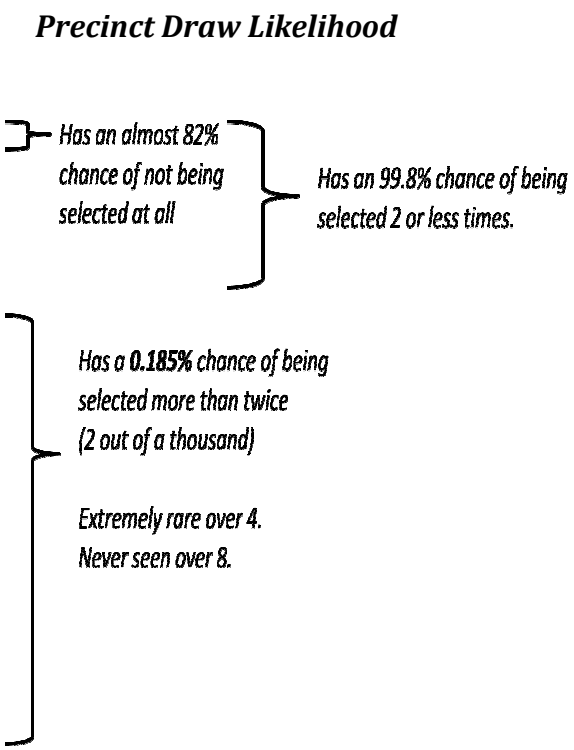
### *1.5 Anomaly Two--BERN307 Drawn 41 Times in 486 Rolls*

Modeling the process hundreds of thousands of times with multiple starting random number seeds and using two different ways of weighting the precincts (by cast votes and by registered voters) gives consistent results within expected statistical bounds.

#### *1.5.1 Modeling Anomaly Two*

Here are the results of a single set of 100K sets of runs showing the frequency of any precinct getting selected 0 times to 8 or more times.

| **total precinct selected** | **frequency** | *100K runs precinct selection rate* |
|---|---|---|
| 0 | 176,914,864 | *81.79143042%* |
| 1 | 34,841,315 | *16.10786639%* |
| 2 | 4,142,891 | *1.91534489%* |
| 3 | 371,884 | *0.17192973%* |
| 4 | 27,191 | *0.01257097%* |
| 5 | 1,737 | *0.00080305%* |
| 6 | 115 | *0.00005317%* |
| 7 | 2 | *0.00000092%* |
| 8 | 1 | *0.00000046%* |
| **>8** | **0** | *0* |



Summary of modeling results--precinct averages:

- Almost 82% chance of not being selected.
- 99.8% chance of being selected two times or less.
- Approximately 2 in a thousand are selected more than twice.
- Extremely rare over 4 and never more than 8

### 1.5.2 Modeling: Three Sets of 100 Thousand Runs vs AUP22

The following illustration compares three sets of 100K sets of runs. The first two sets weight the precinct sizes for normalization according to cast vote count (as in AUP22) with different starting random seeds, and the third uses weights the precincts by registered voter counts and the same random number seed as one of the first two (relevant comparisons where needed). The results converge so well that when charted they essentially overlay each other.

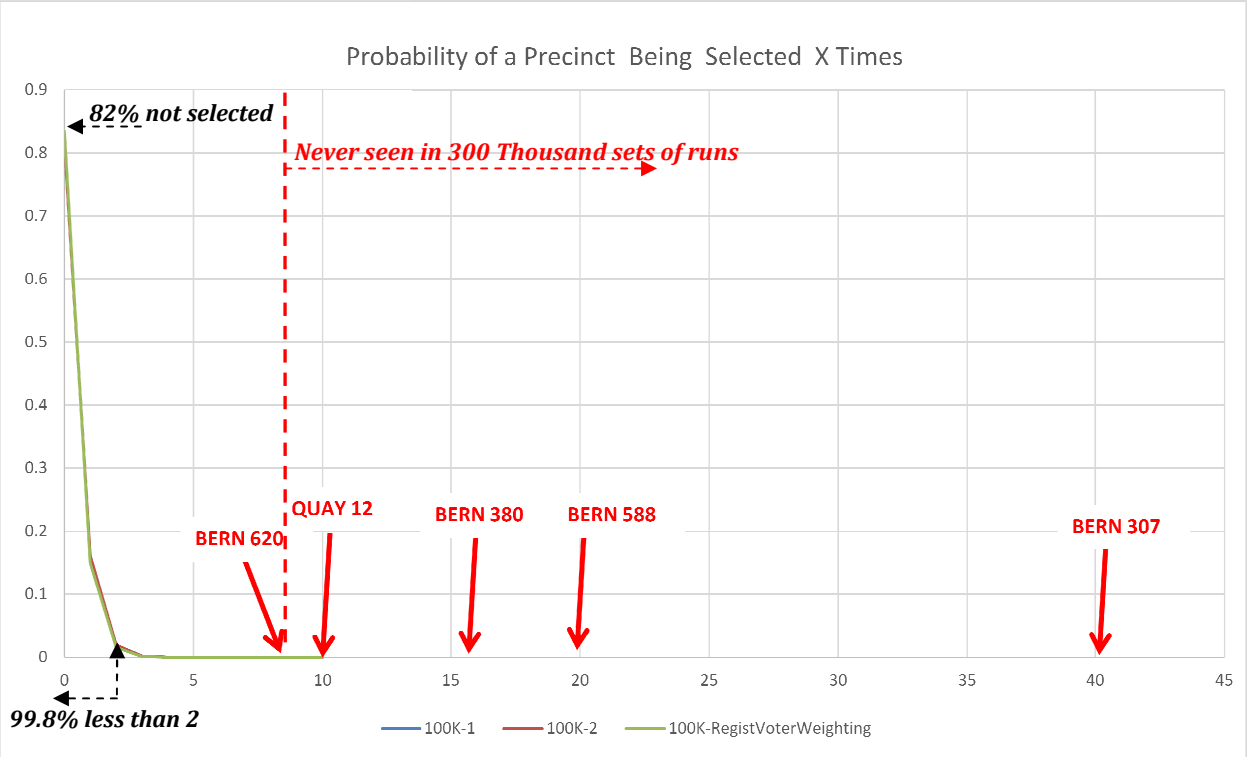
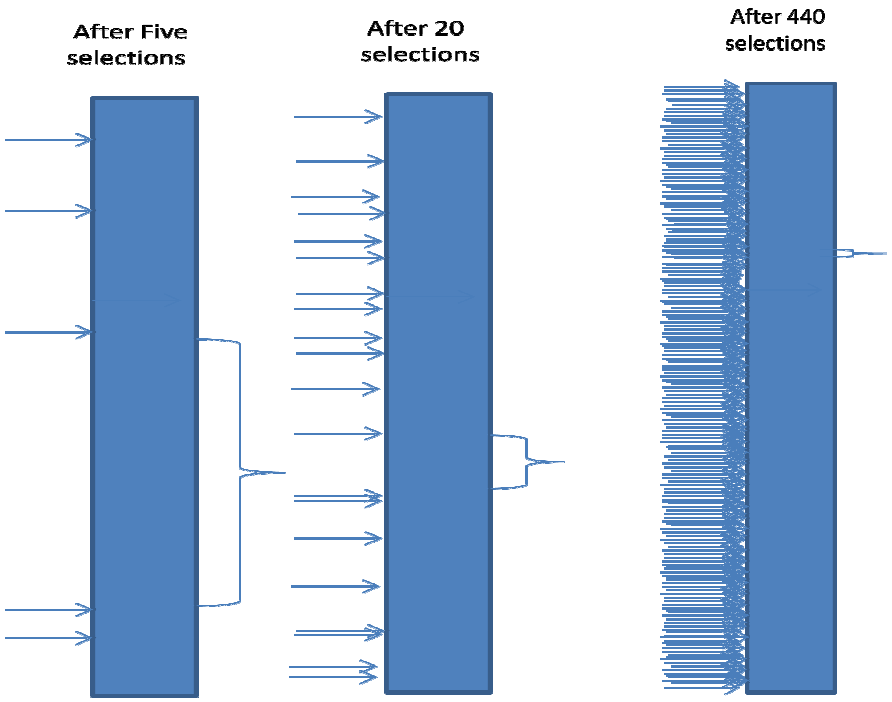



Figure 1. AUP22 Extreme Outliers

The selection count is along the bottom axis and the ratio of selection as height. Zero selections has a value of almost 0.82 (82%). Once the count reaches two, the ratio is almost zero. BERN307 isn't the only county selected an astronomically improbable number of times. Anomaly two alone is sufficient to demonstrate the result of the selection process was not random.

### *1.6 Anomaly Three--Consecutive Unselected Rows*

Before the first roll all 2163 precincts are unselected so there are 2163 consecutive unselected rows. As random selections are made, the groups of consecutive unselected get smaller. The following illustrates how it shrinks.




#### *1.6.1 Modeling Anomaly Three--Consecutive Unselected Rows*

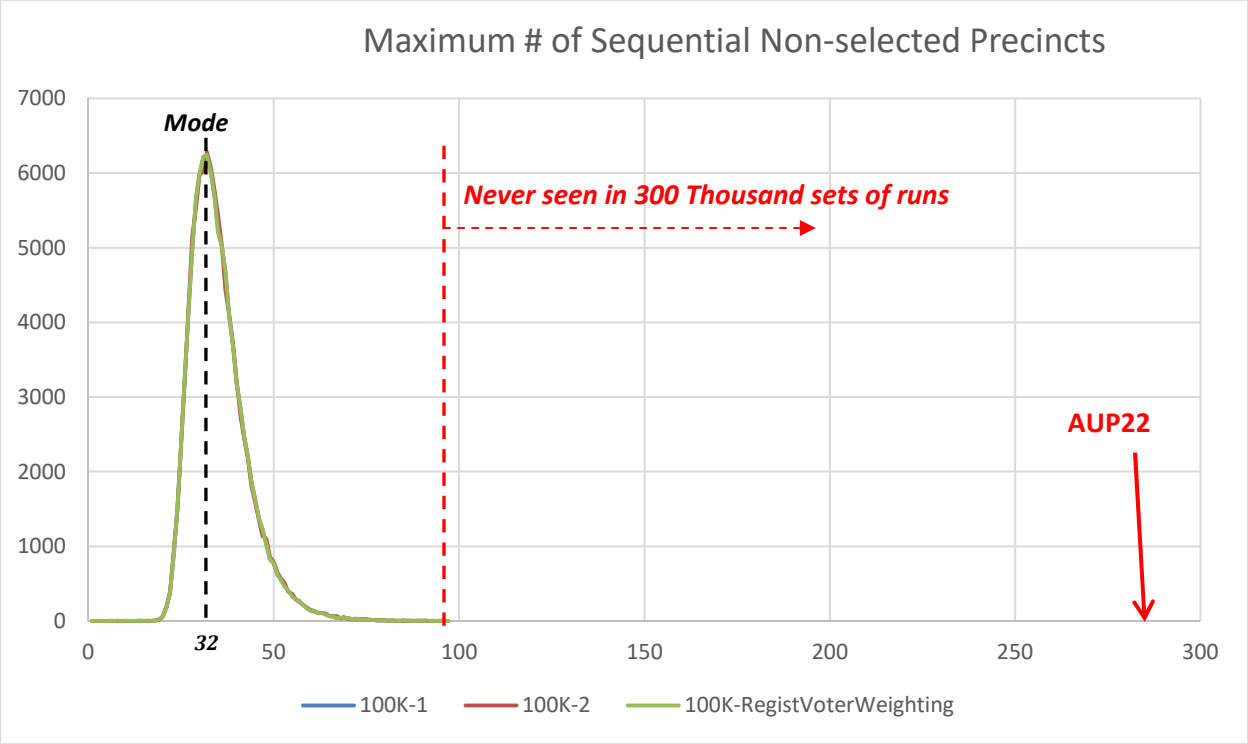
In a set of runs the random draws continue until the audit requirements are complete. This number varies depending on the "luck" of the draw. The AUP22 scenario it is driven by the District 2 requirement of 109. For the 100K sets of runs, the number of draws required ranged from 619 to 298. This low value had a lot of "lucky" district 2 rolls and finished early. The average number of rolls required was 444.

| Rolls to fill audit criteria | |
|---|---|
| Highest: | 619 |
| Average: | 443.6 |
| Lowest: | 298 |

| Consecutive Unselected Rows | |
|---|---|
| Lowest: | 18 |
| Most Common: | 32 |
| Highest: | 97 |

At the end of each set of runs the largest number (block) of unselected precincts was collected for comparison with AUP22 results. The normalized values from AUP22, once sorted, has the two values farthest apart being precinct index 234 and 516. A block of 282 consecutive unselected rows.

| RollNumber | Roll | prec_row | county | pct | distance |
|---|---|---|---|---|---|
| 230 | 1804999 | 234 | Bernalillo | 307 | 0 |
| 174 | 1818800 | **234** | Bernalillo | 307 | 0 |
| 87 | 1865747 | **516** | Bernalillo | 229 | **282** |
| 188 | 1896459 | 522 | Bernalillo | 231 | 6 |





This illustrates what a 13% gap looks like after 484 selections.

A 13% gap after 484 selections in a random selection is astronomically improbable.

Anomaly three alone is sufficient to demonstrate the result of the selection process was not random.

## *1.7 Conclusion*

Three different anomalies have been presented to demonstrate the 2022 precinct selection results were not random and therefore invalid. The first two prove some counties were selected too many times, and the last reveals how entire blocks of counties were missed for selection.

No assumptions or assertions have been presented about how or why these anomalies occurred. An examination of the tool used for converting rolls to precincts in AUP22 is required for the problem to be fully understood.