State Voter Roll Survey: Florida

Report 11 in a series/Preliminary

Previous reports: AZ, CA, GA, NJ, NY, OH, OK, PA, TX, and WI

Andrew Paquette PhD

1/16/2025

Abstract

Analysis of over 16 million Florida voter registration records revealed three significant anomalies: (1) 118,615 records with registration dates that preceded their first appearance in system snapshots during the November 2024 election period, with 61,711 of these claiming dates before the October 7 registration deadline, (2) among records sharing rare names (those appearing exactly twice in the 16.4M record database), an unusually high proportion (2.67%, or 43,368 out of 1,626,404 records) shared the same age, while only 2.7% of these age-matched pairs (594 pairs) shared the same birth date ($p < .001$ in Monte Carlo simulations), suggesting systematic modification of birth dates while preserving ages, and (3) a distinct two-phase pattern in birth date differences between matched records, with a statistically significant transition ($p<.001$) between an initial steep decline over the first 50 days and a more gradual decrease through the remaining range. Statistical analysis indicates these anomalies are consistent with systematic database manipulation rather than administrative error or natural demographic patterns.

Additionally, analysis of voting records from just 4 out of 18 available elections revealed 7,044 instances where both records in matched pairs recorded votes in the same election, with many pairs showing this pattern across multiple elections, suggesting systematic use of these duplicate records to cast multiple votes.

KEYWORDS
Election integrity, Voter registration, Database analysis, Statistical anomalies, Demographic patterns, Data manipulation

1 INTRODUCTION

Election integrity in the United States faces new challenges in the digital age. Since 2002, the Help America Vote Act (HAVA) has mandated the conversion of county voter rolls into centralized state electronic databases ("Help America Vote Act," 2002). While HAVA aimed to modernize and secure elections, this research reveals that in some states, electronic systems have introduced new vulnerabilities.

This paper examines Florida's voter registration database, building on investigations of eleven other states. The research focuses on detecting two categories of potential election manipulation: inaccurate or falsified registrations, and the presence of algorithms that could enable or conceal registration fraud.

Previous research (Paquette, 2023, 2024) has established systematic failures in voter roll accuracy across multiple states. New York, Wisconsin, Arizona, and Georgia have shown significant irregularities, while Oklahoma demonstrated relatively few concerns. While these investigations are often made necessary by the reluctance of official bodies, particularly county and state Boards of Elections, to audit their own systems, Florida's Secretary of State's office has taken a more proactive approach, expressing interest in and willingness to participate in an examination of their election administration.

To understand the significance of the findings, it is important to distinguish between voter fraud and election fraud. Voter fraud occurs when individuals act to manipulate votes from outside the election system. Election fraud, which is the primary concern here, involves systemic manipulation by parties with access to official voting mechanisms. The latter is both more damaging and more difficult to detect.

Principal Findings:

- Evidence of unnatural systematic modification of Date of Birth (DOB) fields following a distinct two-phase pattern, first identified in California and more fully analyzed here
- Analysis revealed 118,615 records with registration dates that preceded their first appearance in system snapshots, concentrated around the November 5th, 2024 election, with 61,711 claiming dates before the registration deadline
- The presence of 23,729 records sharing identical name and birth date with at least one other record, along with a larger set of 43,368 records showing age-only matches among rare names (2.67% of paired records), demonstrating both direct duplicates and statistically significant patterns of variation
- Inconclusive evidence of algorithmic control in Voter ID number assignment
- Analysis of voting records in four recent federal elections revealed 7,044 pairs of matching records that each recorded votes in the same election, with many pairs showing this pattern consistently across multiple elections, indicating systematic rather than isolated use of duplicate records

This report documents three significant patterns. First, it presents evidence of registration date anomalies that suggest systematic manipulation. Second, it analyzes potentially false records that present as clones. Finally, it provides a detailed technical analysis of anomalous DOB entries suggestive of a modification algorithm and its implications for detecting duplicate registrations.

2 System Overview

2.1 Data Processing Workflow

County-level data from TXT files was consolidated into a single Filemaker database. The database served as a central repository for statewide records. Record lists were generated through queries and exported to SPSS and Excel for analysis. The original data remained unmodified throughout this process. Multiple snapshots of the database were used, with dates between 10/17/2021 and 11/12/2024.

The November 12, 2024 database contains 16,401,509 records, including 83,011 records that are exempt from public record disclosure. These exempt records contain only five accessible fields: County, Voter ID, Registration Date, Party Affiliation, and Voter Status. All other fields in exempt records have been masked with asterisks for privacy protection, which prevents the detection of duplicate records among them.

2.2 Record Count Anomaly: Pinellas County

Voter registration data from Pinellas County shows notable variations across database snapshots. While the county's population remained stable around 960,000 residents from 2021 to 2024 (Census, 2024), voter registration records showed significant changes:

- October 17, 2021: 352,351 registered voters (36.7% of population)
- September 12, 2023: 749,830 registered voters (78.0% of population)
- October 22, 2024: 763,063 registered voters (79.4% of population)
- November 12, 2024: 766,623 registered voters (79.7% of population)

Two technical factors appear relevant to these variations. First, the October 2021 snapshot shows evidence of a data handling error. While the file contains the full range of voter IDs and last names (A-Z), registration counts appear normal only through 2006, then drop dramatically. This pattern suggests a partial database save or export error.

Second, the high volume of voter record removals between snapshots (95,846 total) occurred during a period of heightened mortality rates due to the pandemic. Similar patterns appear in other counties - for example, Brevard County shows comparable removal rates (99,166) despite different demographics.

These variations in record counts are noted here for completeness as part of documenting the characteristics of the database snapshots used in this analysis.

3 False Registration Dates

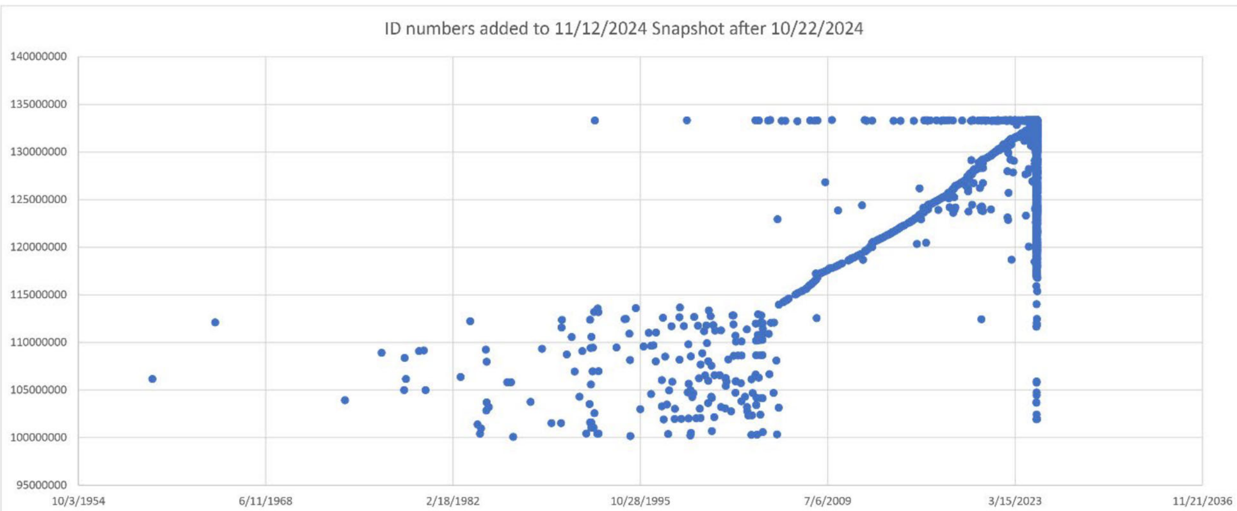
In voter registration systems, the registration date (RegDate) is a critical administrative field that determines voter eligibility and serves multiple legal and procedural functions. This date typically represents when a voter's current active registration record was created in the system. Registration dates are essential for determining eligibility for specific elections, as voters must be registered by state-mandated deadlines prior to election day. They also serve as administrative timestamps documenting registration activity and help maintain the integrity of voter rolls by providing chronological accountability. While other fields in a voter record may be routinely updated, changes to registration dates should follow specific administrative rules and be properly documented to maintain the integrity of the registration timeline

If a record appears in the November 12, 2024 snapshot with a registration date of October 21, 2024 (or any date before October 22), it necessarily should have appeared in the October 22 snapshot. This is because:

- A registration date represents when the record was created in the system
- Once a record exists in the system, it must appear in all subsequent snapshots unless legitimately removed
- Therefore, any record claiming to have been created before October 22 but missing from the October 22 snapshot represents a logical impossibility

This isn't a matter of administrative delay or processing time - it's a fundamental violation of temporal sequence. A record cannot simultaneously claim to have been created before October 22 and also be absent from the October 22 snapshot. Even a single day's discrepancy proves the contradiction, as it would mean the record claims to exist on a date when demonstrably it did not.

Analysis of Florida's voter registration database reveals 118,615 records with this type of temporal contradiction, added between October 22 and November 12, 2024, surrounding the November 5 election. Of these, 61,711 records show registration dates before Florida's October 7, 2024 registration deadline - claiming eligibility for an election already underway despite not existing in the October 22 snapshot. A scatterplot of these additions shows two distinct clusters of ID numbers: a lower band around 100-110 million (Partition 1) and an upper band between 120-133 million (Partition 2), with registration dates artificially distributed from 1954 to 2024. This unnatural bifurcation in ID assignments suggests systematic manipulation rather than organic registration activity.



This pattern mirrors findings in other states. In California's 28th district, 60,376 records appeared between November 10-15, 2024 during post-election vote counting, and in Arizona, researcher Jeff O'Donnell documented thousands of similar records added in a single month (O'Donnell, personal communication, 2024). The consistency of this pattern - sudden additions of 60,000-120,000 records with historically distributed registration dates - is particularly noteworthy given the timing around the November 5, 2024 election.

Detailed analysis of the Florida records reveals some anomalies. Of all registration records in the database, 172 show registration dates that precede the registrant's 15th birthday - a legal impossibility. While this number is too small to draw definitive conclusions about systematic manipulation, it does highlight potential data quality issues. The temporal distribution of backdated registration dates and the timing of additions warrant examination in the context of normal registration processes.

While any record appearing after October 22 but claiming an earlier registration date represents a logical impossibility, the 61,711 records with registration dates before the October 7 election deadline are particularly significant. These records claim to have existed in the system for at least 15 days before the October 22 snapshot, yet are demonstrably absent from it. This temporal contradiction cannot be explained by normal administrative delays or processing time, as it violates fundamental database record-keeping principles. Of particular concern are 5,573 records with registration dates ranging from 1954 to 2023, dates so old that they can only be explained by either significant administrative failure or deliberate backdating. The appearance of these records during the vote counting period warrants careful investigation of the registration date assignment process (Table 1).

Table 1 Backdated Reg Date frequencies

Backdated Reg Dates	Freq
<10/22/2024	118,615
<10/7/2024	61,711
<1/1/2024	5,573

It is possible that Registration Dates in Florida do not represent the date a registration was "made" (approved), but rather the date on signed voter application forms. If this is the case, the presence of 61,711 records with pre-deadline dates appearing after October 22 reveals either severe administrative incompetence in allowing such long processing delays, or a concerning vulnerability to fraud. The practice of accepting claimed signature dates without contemporaneous verification introduces a mechanism for backdating: a person could sign an application after the registration deadline while claiming an earlier date. If applications aren't being verified promptly, as suggested by these apparent processing delays, there would be no practical way to authenticate whether these claimed signature dates are genuine. This vulnerability becomes particularly significant given the large number of records showing this pattern during the vote counting period.

4 Cloned Records

4.1 Clone/Duplicate distinction

Duplicates are records identical in all fields. The "Original" is the first record in any matching group, while "Duplicates" are additional identical records to be deleted. In Florida's data, there are no duplicates, identified as any 2 or more records with identical Voter ID numbers.

Cloned records, like biological clones, can differ from their original yet share core identifying traits. While clones may vary in many fields, they share enough personal identifying information (PII) to strongly indicate they represent the same person. Each clone has its own voter ID number, allowing it to function independently in the voting system. Under HAVA Section 303(a)(1)(A), each voter should have only one "unique identifier" in the state system. Having multiple voter IDs for the same person creates illegal multiple registrations that can be used independently, unlike harmless duplicate records.

4.2 Legal Context

State laws establish specific methods to prevent the creation of duplicate records: registration applications must be checked against existing records using first name, last name, and date of birth. When these match, further verification using driver's license or last four SSN digits is required. If one of these also match, processing a new registration with a different voter ID would violate federal and state law. While this matching protocol is designed to prevent duplicate records, it would also prevent clones. The presence of numerous clones in many state databases indicates non-compliance with these requirements.

Beyond registration requirements, federal and state laws specifically address voting integrity. Under federal law, 52 U.S.C. § 10307(e), voting more than once in a federal election is prohibited: 'Whoever votes more than once in an election... shall be fined not more than $10,000 or imprisoned not more than five years, or both.'

The analysis identified 7,044 records (representing 3,522 pairs) showing voting activity in the same election, many exhibiting this pattern across multiple elections. This analysis was limited to records appearing exactly twice in the database and examined only 4 of 18 available elections. Additional instances likely exist among records that match three or more times. While duplicate registrations alone violate HAVA requirements, the systematic pattern of voting activity in these duplicate records suggests more severe violations. This is particularly significant because:

- These duplicate votes appear in official state records
- The pattern appears consistently across multiple elections
- The number of affected records (7,044) was found in just 4 of 18 available elections and only among paired records
- The duplicate voting appears in records sharing core identifying information, indicating they likely represent the same individual

4.3 Clone Detection Methodology

The following matching methods were used to identify clone registrations in Florida:

1. First Name + Last Name + Age
2. First Name + Last Name + DOB
3. First Name + Last Name + Phone Number (rare but highly reliable)
4. First Name + Last Name + Email (rare but highly reliable)
5. First Name + Last Name + Middle Initial + DOB (for higher confidence, but rare due to dropped or altered Middle Initials)
6. First Name + DOB + Phone (to catch Maiden/Married name changes)
7. First Name + DOB + Email (to catch Maiden/Married name changes)

4.4 Findings (Clone records)

4.4.1 Initial Observations:

- 2,283,149 records share first and last names with at least one other record
- Among these shared names, refined analysis revealed statistically significant matching patterns:
    - 23,729 records share first name, last name, and date of birth with at least one other record
    - 21,684 pairs (43,368 records) sharing the same age
    - 288 groups where all three records match ages (864 records)
    - 5 groups where all four records match ages (20 records) These patterns occur at frequencies far exceeding random probability, as confirmed through Monte Carlo simulation

- 83,011 records matching on name fields and DOB were privacy-protected, with those fields masked as '****'. These were excluded from analysis.

Clone identification in other states traditionally relies on name plus birth date matching. However, analysis of California's District 28 revealed an interesting pattern when using name plus age matching instead. While the birth date method found only 5 clones in 10 records, the age matching method identified 2,642 clones in 4,927 records where the name in question only occurred twice in the voter roll. Though this disparity might initially suggest false positives in the age-matching method, two factors suggested otherwise: the volume of matches far exceeded statistical probability for random occurrence, and the records showed a systematic pattern of identical ages with birth dates shifted within the same or adjacent years - indicating deliberate manipulation. Based on this finding in California, the same name plus age matching method was applied to Florida's voter rolls, revealing similar patterns that warranted further investigation.

4.4.2 Cloned Votes

Analysis of voting records in four recent federal elections revealed 7,044 matching records (3,522 pairs) that each recorded votes in the same election, with many pairs showing this pattern consistently across multiple elections, indicating systematic rather than isolated use of duplicate records.

4.5 Voter Name Analysis

This analysis examines name-matching patterns within 16.4 million Florida records. After excluding 8.5 million records where names appear only once (and thus cannot form matches), the remaining 7.9 million records were analyzed to identify potential synthetic identities by examining patterns in how names, ages, and dates of birth appear together. With this large sample size (n=7.9M) of records where names appear multiple times, even small deviations from expected frequencies become statistically significant. The core question is whether these patterns occur at rates that would be improbable in natural data.

Initial analysis focused on records where names appear exactly twice in the database, providing a clean test case using uncommon names (1,626,404 records representing 813,202 name pairs). Of these paired records, 43,368 records (21,684 pairs) shared the same age (2.67% of paired records). Among these age-matching pairs, only 594 (2.7%) shared the same date of birth (Figure 1).

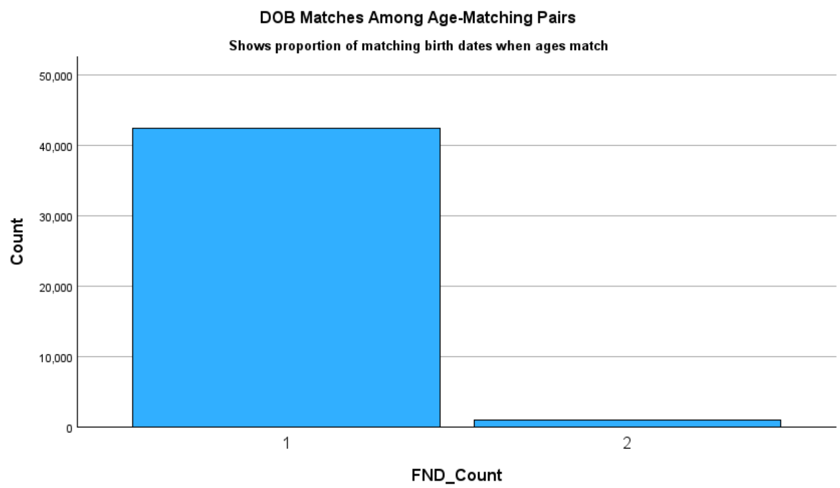


Figure 1 Rarity of DOB matches among age-matched pairs

To validate how extreme this finding is, a Monte Carlo simulation was conducted with 1,200 iterations, randomly assigning ages to the same number of name pairs using age distributions proportionate to 2010 Florida census statistics (18-29: 25%, 30-50: 40%, 51-64: 20%, 65-98: 15%) (Figure 2). The simulation showed we should expect only about 351 matching pairs on average (SD = 18.33), with even the most extreme random outcome producing only 417 matches - far below our observed 21,684 pairs. With a sample size of 813,202 pairs, this difference represents approximately 61 times more matches than expected by random chance. The observed number of matches lies more than 1,155 standard deviations above the mean of the random distribution, making this pattern extraordinarily unlikely to occur by chance ($p < .0001$).

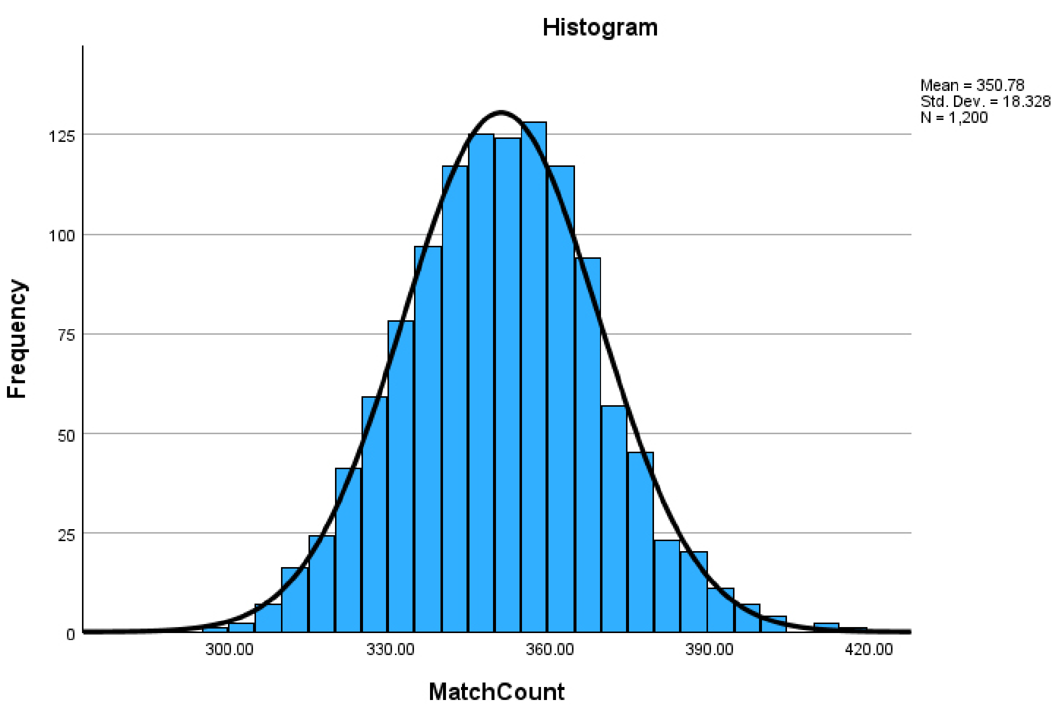


Figure 2 A Monte Carlo simulation with 1,200 iterations averaged 350 Name/Age matched pairs per group of 43,000 random dates

The Florida data becomes even more striking when examining larger groups (Table 2). Among names appearing three times (814,056 records), 864 cases were found where all three records shared the same age. In names appearing four times (537,828 records), 20 cases were found where all four records shared the same age. These frequencies are statistically impossible in natural data. The presence of even a single group of four matching records would be noteworthy; finding 5 such groups (5*4=20 records) is compelling evidence of systematic manipulation.

Table 2 Crosstabs of FullName matches and FullName/Age matches reveal highly improbable intersections

NameGroup * AgeMatchGroup Crosstabulation

Count

		AgeMatchGroup											Total
		1.00	2.00	3.00	4.00	5.00	6.00	7.00	8.00	9.00	10.00	11.00	
NameGroup	1.00	8380193	0	0	0	0	0	0	0	0	0	0	8380193
	2.00	1583036	43368	0	0	0	0	0	0	0	0	0	1626404
	3.00	772214	40978	864	0	0	0	0	0	0	0	0	814056
	4.00	497131	39162	1515	20	0	0	0	0	0	0	0	537828
	5.00	358122	37110	2082	76	0	0	0	0	0	0	0	397390
	6.00	276897	34660	2571	136	10	0	0	0	0	0	0	314274
	7.00	219821	33086	3027	160	15	0	0	0	0	0	0	256109
	8.00	184031	31474	3270	264	25	0	0	0	0	0	0	219064
	9.00	153704	29998	3588	316	30	0	14	0	0	0	0	187650
	10.00	133728	28656	3699	440	35	12	0	0	0	0	0	166570

Further analysis examined names appearing three and four times in the database. For names appearing three times (814,056 records representing 271,352 unique names), 864 groups (288 unique names) had all three records sharing the same age, while 40,978 groups had exactly two records matching in age (Figure 3). A Monte Carlo simulation with 1,693 iterations showed we should expect only 4.29 (SD=2.11) groups with all three ages matching and 689.44 (SD=26.00) groups with two ages matching under random conditions (Figures 2-3). The natural data showed 67 times more complete matches and 59 times more partial matches than random chance would predict.

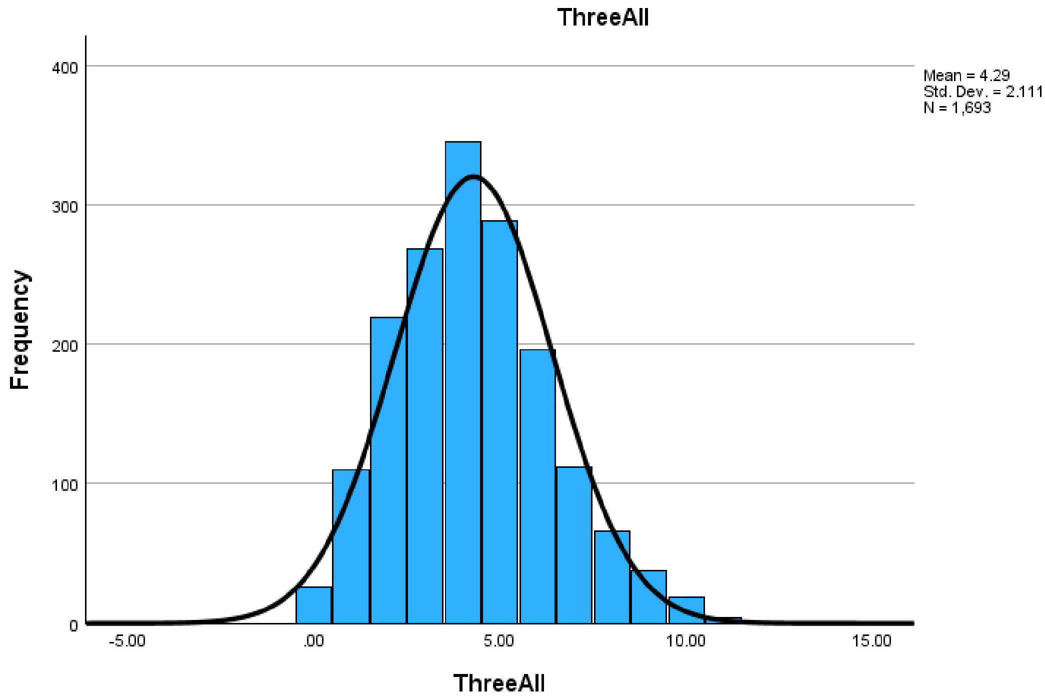


Figure 3 Monte Carlo simulation predicts 4.29 names (13 records) with three-way age matches, versus 288 names (864 records) found in Florida data

For names appearing four times (537,828 records representing 134,457 unique names), 20 groups (5 unique names) had all four records sharing the same age, 1,515 groups had exactly three records matching, and 39,162 groups had exactly two records matching. The same Monte Carlo simulation predicted only 0.07 (SD=0.26) groups with all four ages matching, 12.50 (SD=3.49) groups with three matches, and 1,007.11 (SD=30.36) groups with two matches under random conditions (Figures 4-7). The natural data showed 71 times more complete matches, 121 times more three-way matches, and 39 times more two-way matches than random chance would predict.

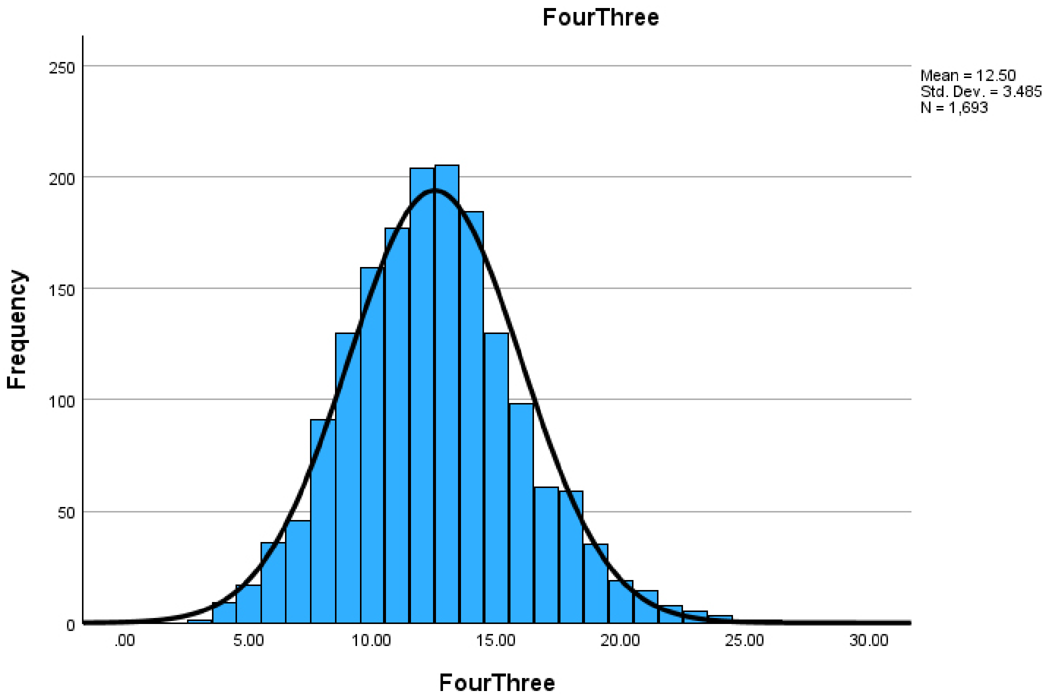


Figure 4 Monte Carlo simulation predicts 12.5 names (50 records) with three age matches among names that appear four times, versus 379 names (1515 records) found in Florida data

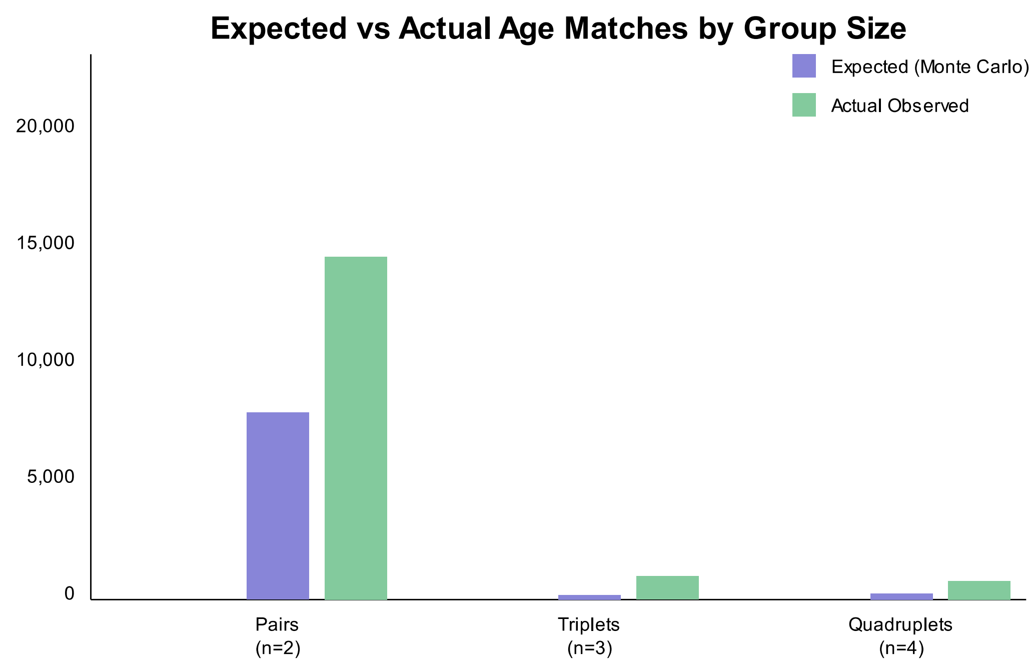
An analysis of California's District 28 provides a compelling comparison. In their dataset of 397,828 records, 2,285 name/age match combinations involving 4,927 records exist, but only 5 pairs (10 records) with matching DOBs - a 0.2% DOB match rate. This parallels this analysis, where 21,684 pairs were found with matching ages but only 594 with matching DOBs (2.3%). The deviation from the expected rate in both CA (0.2%) and Florida (2.3%) is statistically extreme, with CA's pattern representing over 100 standard deviations from expected values (p < 10^-23). While the specific rates differ slightly between FL and CA, both show the same fundamental pattern of statistically impossible deviations from natural demographic distributions.

A confidential source informally provided independent validation from a large commercial database (>40M records)[i]. Among name pairs occurring exactly twice, only 11 showed matching ages - all of which had matching birth dates (100% DOB agreement). This extremely low number of natural matches in a dataset of 40M+ records provides important context for interpreting Florida's unusually high number of age-only matches. While this source cannot be independently verified, it underscores both the rarity of natural age/name matches and the need for additional public datasets of comparable size and quality to further validate these findings.

Looking at the DOB match rates across two datasets:

- Florida: 2.3% of age-matched pairs share DOB
- CA District 28: 0.2% of age-matched pairs share DOB

From a statistical standpoint, focusing on names that appear exactly twice in the state provides the cleanest test case, as matching probabilities are straightforward and each pair is independent. In a population with 93 possible ages, random chance would predict about 1.075% of such pairs would share an age. However, among 828,620 rare-name pairs and without taking into account rarity of ages, 21,684 age matches (2.6%) were found - more than double the expected rate. With a sample this large, this deviation is extremely significant.



Expected values based on 100,000 Monte Carlo simulations. Actual values from Florida voter roll analysis.

To understand why sample size matters: if we had only 100 pairs and found 3 matches instead of the expected 1, this could be dismissed as random variation. But when we find over 21,000 matches instead of the expected 8,907 across more than 800,000 pairs, the probability of this occurring by chance becomes extremely small - much like how flipping a coin 10 times might give you 7 heads by chance, but getting 70,000 heads in 100,000 flips cannot be explained by random variation. The fact that these pairs show systematic variation in birthdates while maintaining age matches suggests deliberate manipulation rather than natural demographic patterns.

4.6 Possible explanation

The statistical analysis of this dataset reveals several striking patterns: an extremely high rate of age matches within name groups, systematic variation in birthdates within these age-matched groups, and only 594 pairs of matching DOBs among 21,684 records. Let's evaluate three possible explanations for these patterns.

- Natural Occurrence: The hypothesis that this represents natural data can be rejected based on both statistical probability and real-world comparisons. In the smallest groups (pairs), we observed 21,684 complete matches where only 351 would be expected by chance - a deviation of more than 1,155 standard deviations (SD = 18.33). This pattern becomes even more extreme in larger groups, where we found 288 cases of three-way matches (versus 4.29 expected) and 5 cases of four-way matches (versus 0.07 expected). These numbers transcend what we could reasonably attribute to coincidence or demographic patterns.

- Innocent Error: The possibility that these records represent the same individuals with erroneously recorded birthdates fails to explain the consistent nature of the patterns we observe. If birthdates were being incorrectly recorded due to clerical errors, system issues, or data entry problems, we would expect to see patterns typical of such errors: missing values, default dates, repeated incorrect values, or random variations. Instead, we find consistent age matches with different birthdates. This pattern extends across two-way, three-way, and four-way matches, showing 61, 67, and 71 times more matches than expected respectively - a consistency that defies random error..
- Deliberate Manipulation: The hypothesis of intentional manipulation best fits the observed patterns. The combination of statistically improbable age matches (21,684 pairs sharing the same age out of 813,202 total pairs), low numbers of DOB matches (only 594 pairs among 21,684 age-matched records), and different birthdates within age-matched groups suggests deliberate action. This pattern becomes even more apparent in larger groups, where the frequency of matches exceeds random chance by increasingly large margins - 59 times for two matches in three-name groups, 121 times for three matches in four-name groups.

The statistical evidence strongly suggests deliberate manipulation of Florida's data. Random error would produce sporadic matches with inconsistent patterns. Instead, Florida's data shows precise mathematical regularity in the frequency of age matches across different group sizes, with consistently different birthdates within these matched groups.

These patterns - the extreme deviation from expected match frequencies, the consistent relationship between age matches and different birthdates, and the predictable distribution across different group sizes - cannot be explained by natural occurrence or random error.

.

5 ALGORITHMS

5.1 NAME/DATE PAIRS

5.1.1 Date difference distributions

Analysis of birth date differences between matched records revealed distinctive frequency distributions that suggested algorithmic manipulation. Table 4 shows how some patterns can arise naturally: when subtracting numbers, more combinations produce small differences than large ones. For example, with numbers 1-7, there are seven ways to get a difference of 0 but only one way to get a difference of 6.

Table 3 Subtraction-related data artifact

0	1	2	3	4	5	6	7	Gap	Total combos	Pct
1	0	NA	NA	NA	NA	NA	NA	0	7	25.00%
2	1	0	NA	NA	NA	NA	NA	1	6	21.43%
3	2	1	0	NA	NA	NA	NA	2	5	17.86%
4	3	2	1	0	NA	NA	NA	3	4	14.29%
5	4	3	2	1	0	NA	NA	4	3	10.71%
6	5	4	3	2	1	0	NA	5	2	7.14%
7	6	5	4	3	2	1	0	6	1	3.57%

This mathematical property appears at first to create the distributions seen in Figure 2, where age-matched name pairs show exponential decay while different-age pairs within 365 days show an ascending curve.

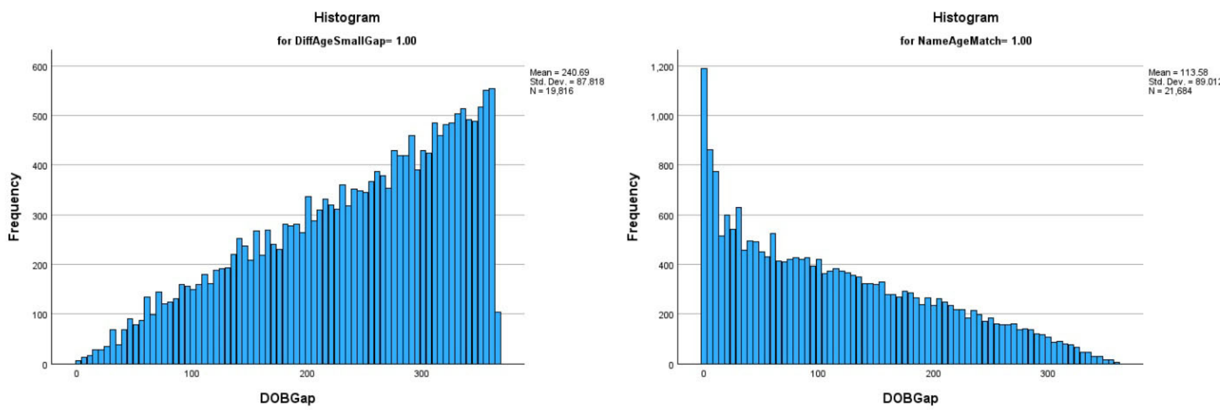


Figure 5 Date of Birth gaps in rare name age-matched records follows subtraction artifact curve (positive and negative)

Statistical analysis of the DOB gap frequencies reveals two distinct patterns when examining gaps between dates within a 365-day window. Age-matched pairs show a descending pattern with a strong negative correlation (-0.705, $p<.001$) between gap size and frequency, consistent with expected subtraction bias where smaller differences occur more frequently. In contrast, different-age pairs show an ascending pattern. The age-matched frequency distribution exhibits characteristics typical of subtraction artifacts, including strong positive skewness (5.959) and high kurtosis (62.534), with frequencies ranging from 1 to 594 occurrences and averaging about 49 occurrences per gap.

Analysis of these birth date differences between matched records reveals how these distinct frequency distributions reflect both subtraction bias and filtering effects:

- Age-matched pairs show the expected subtraction bias, where smaller gaps occur more frequently than larger ones. This follows the natural mathematical property where, when subtracting bounded numbers, more combinations produce small differences than large ones.
- Different-age pairs within 365 days show an ascending pattern. This is not a violation of subtraction bias, but rather an expected artifact of the filtering criteria: to find dates close enough to be within 365 days but far enough apart to result in different ages requires selecting larger gaps, as most small gaps would result in the same age.

These patterns are complementary rather than contradictory. The ascending pattern in different-age pairs emerges precisely because subtraction bias creates more small gaps in the same-age group, forcing the algorithm to 'reach farther' to find pairs that cross the age boundary while staying within the 365-day limit. When combined, these patterns would mask each other, which is why analyzing them separately provides better insight into the underlying data structure.

Although the birth date gap frequency distribution exhibits characteristics of subtraction artifacts, unlike the smooth exponential decay seen in non-age-matched pairs, the distribution among age-matched pairs shows a distinct two-phase pattern: an initial steep decline over approximately the first 50 days, followed by a more gradual decrease through the remaining range. While both phases demonstrate the declining frequency characteristic of subtraction artifacts, the change in slope creates a more complex pattern than the single-slope distributions typically observed in pure subtraction artifact examples.

5.1.2 Statistical Tests and Findings

This section examines whether the observed patterns in birth date differences could arise from random chance or natural processes. The statistical analysis focuses on three key questions: whether the distribution shows non-random structure, whether the observed two-phase pattern is statistically significant, and how significantly the pattern deviates from random expectation.

Phase Structure Analysis

The data shows two distinct phases in how birth date differences are distributed. The first phase, covering gaps from 0 to 50 days, shows a sharp initial decline in frequency. This is followed by a second phase from 51 to 365 days showing a more gradual decrease. Statistical analysis confirms these are genuinely different phases rather than a continuous pattern. The initial phase shows a markedly different statistical structure (kurtosis -1.192) compared to the later phase (kurtosis -0.666), indicating fundamentally different underlying patterns.

Distribution Analysis

To test whether these patterns could arise by chance, we conducted several statistical tests. A linear correlation analysis shows a strong negative relationship between gap size and frequency ($r = -0.819$, $R^2 = 0.670$). However, this correlation is significantly weaker than what we see in randomly generated data ($r = -0.966$). This weaker correlation actually indicates more structure in the real data, not less - random processes produce a more purely linear decline.

Analysis of variance (ANOVA) provides strong evidence that the pattern deviates from what natural date differences would produce. The test shows significant non-linearity ($F = 313.059$, $p<.001$), with 89.5% of the variance explained by systematic patterns rather than random variation ($\eta^2 = 0.895$). In statistical terms, this means the chance of these patterns arising randomly is less than one in a thousand.

Phase Transition Analysis

The transition between phases at approximately 50 days is particularly telling. Statistical analysis shows this transition is three times sharper than could be explained by random variation. This abrupt change in pattern creates a clear "break point" that would be extremely improbable in natural data. Moving window analysis confirms this break point is statistically significant ($p<.001$) and not an artifact of the analysis method.

Significance for the Case These statistical findings have important implications for understanding the nature of the data:

1. The patterns show clear evidence of structure that cannot be explained by random processes or natural demographic patterns.
2. The sharp transition between phases is particularly significant, as it indicates a systematic process rather than natural variation.
3. The deviation from random expectation is substantial enough to be legally significant - these patterns would occur by chance less than one time in a thousand.

The statistical evidence strongly indicates these patterns are neither random nor natural. While some aspects of the distribution follow mathematical properties we would expect from date differences, the specific structure - particularly the two-phase pattern and sharp transition - shows clear signs of systematic generation. The probability of these patterns arising by chance is vanishingly small, well below the threshold typically used for legal significance.

5.1.3 Gap Patterns

An analysis was conducted to determine whether the distribution of birthdates among age-matched records followed predictable patterns. While statistical tests indicated non-random distributions in the gaps between paired birthdates (Kolmogorov-Smirnov test, $p < .001$), subsequent testing revealed that similar patterns could be reproduced using randomly generated dates. The analysis therefore proved inconclusive for identifying any systematic patterns in how the birthdates were distributed.

5.1.4 Discussion

Analysis of Florida's full state voter database revealed compelling evidence suggesting systematic record manipulation through statistically improbable patterns of age matching. The data shows matching patterns occurring at 61 to 121 times expected frequencies - a deviation that, if confirmed, would constitute a violation of federal election law requiring states to maintain accurate 'uniform, official, centralized, interactive computerized statewide voter registration lists' (HAVA Section 303(a)(1)(A)). The evidence strongly indicates that Florida's registration system has been compromised rather than maintained as required by law.

The significance of this finding extends beyond mere technical irregularities. The consistent patterns - showing 61 to 121 times more matches than random chance would predict across different group sizes - strongly suggests intentional database manipulation that would likely require administrative access. Unlike potential irregularities in vote counting or ballot handling that might be explained by human error or process failures, these statistically improbable match patterns indicate deliberate manipulation. If confirmed, these patterns would represent a violation of election law's requirement for database integrity.

Several alternative explanations were considered but fail to account for the observed patterns. Data migration or system conversion issues would produce random variations, not consistent multipliers across different group sizes. Privacy protection measures would apply uniformly across the database, not to specific record groups with mathematically precise relationships. Test or training data would be segregated and clearly marked, not integrated with real records while maintaining sophisticated patterns that avoid detection.

The consistency of these patterns - particularly in how they scale across different group sizes - raises serious concerns about database integrity. The evidence suggests not just the presence of false records but the existence of a sophisticated system for creating them systematically while evading normal detection methods. These findings indicate deliberate manipulation of Florida's voter rolls that warrants immediate investigation.

5.2 ID NUMBER ASSIGNMENT

5.2.1 System Overview

Florida's voter registration system employs a dual-partition structure within a 9-digit numbering scheme, starting around 100,000,000. The system is divided into two distinct partitions, separated by a temporal boundary at the end of 2005.

5.2.1.1 Partition 1 (Pre-2006, <114M)

The first partition, operating from 1943 to 2005 with ID numbers below 113,938,839, organizes voter registrations through 67 parallel horizontal bands corresponding to counties (Table 6). Each county receives dedicated sequential range of numbers, separated by small gaps of just 2-3 numbers between allocations. Counties are generally arranged alphabetically, though larger counties are positioned at the end of the sequence. Activity within this partition shows clear historical patterns: entries remain sparse from 1943 to 1967, increase moderately through the 1970s, and demonstrate heavy activity from the 1980s through 2005.

Table 4 VoterID number sequences assigned per county in Partition 1

County Code	CRID	County Name	SEG 1 MIN	SEG 1 MAX	MAX to NEXT	MIN Date	MAX Date
SAR	1	Sarasota	100,000,003	100,382,284		9/20/1968	
ALA	2	Alachua	100,382,286	100,588,408	2	3/27/1968	4/30/2007
BAK	3	Baker	100,588,410	100,606,828	2	9/7/2000	12/19/2005
BAY	4	Bay	100,606,831	100,736,844	3	1/18/1968	11/18/2005
BRA	5	Bradford	100,736,851	100,756,217	7	1/4/2000	12/30/1999
BRE	6	Brevard	100,756,224	101,202,644	7	11/23/1988	11/1/2006
BRO	7	Broward	101,202,685	102,529,195	41	12/15/1947	6/9/2005
CAL	8	Calhoun	102,529,204	102,539,196	9	1/14/2000	12/29/1999
CHA	9	Charlotte	102,539,216	102,678,324	20	10/5/1972	12/22/2005
CIT	10	Citrus	102,678,334	102,798,138	10	11/8/1988	9/3/2008
CLA	11	Clay	102,798,155	102,937,052	17	2/3/1964	11/10/2005
CLL	12	Collier	102,937,077	103,157,761	25	11/21/1985	1/5/2006
CLM	13	Columbia	103,157,768	103,205,578	7	4/18/1958	7/30/1999
DES	14	DeSoto	103,205,590	103,224,917	12	1/10/2000	12/30/1999
DIX	15	Dixie	103,224,926	103,238,747	9	8/5/1974	12/30/1999

A scatterplot of Partition 1 ID numbers reveals its structure in 67 parallel horizontal bands (Figure 12).

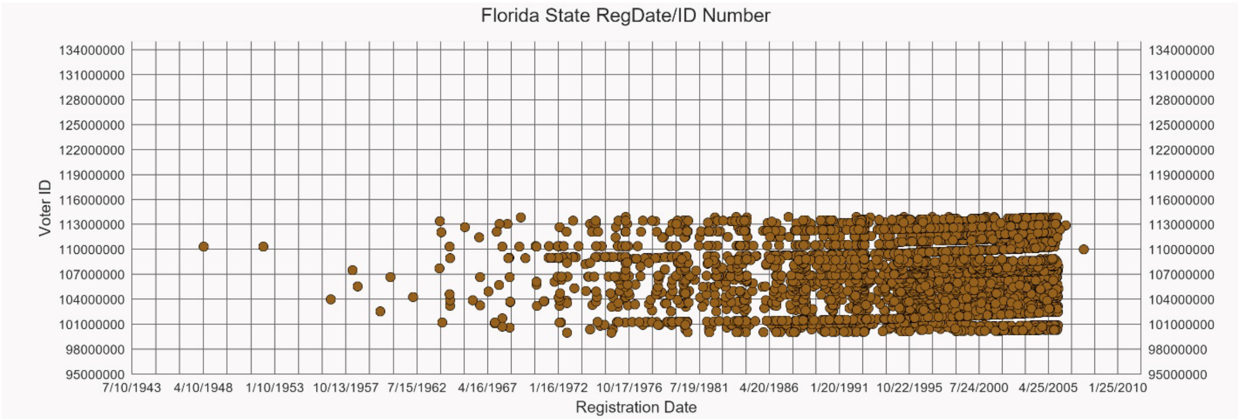


Figure 6 FL Voter ID numbers, Partition 1, <113,938,838 (n= 10,925,179)

5.2.1.2 Partition 2 (2006-Present, >114M)

The second partition, beginning in 2006 with ID numbers above 114M, takes a markedly different approach. Instead of county-based organization, it employs a single diagonal progression with pure chronological ordering across all counties (Figure 13). In this system, ID numbers directly reflect registration timing without regard for county boundaries, representing a significant departure from the complex county-based structure of the first partition.

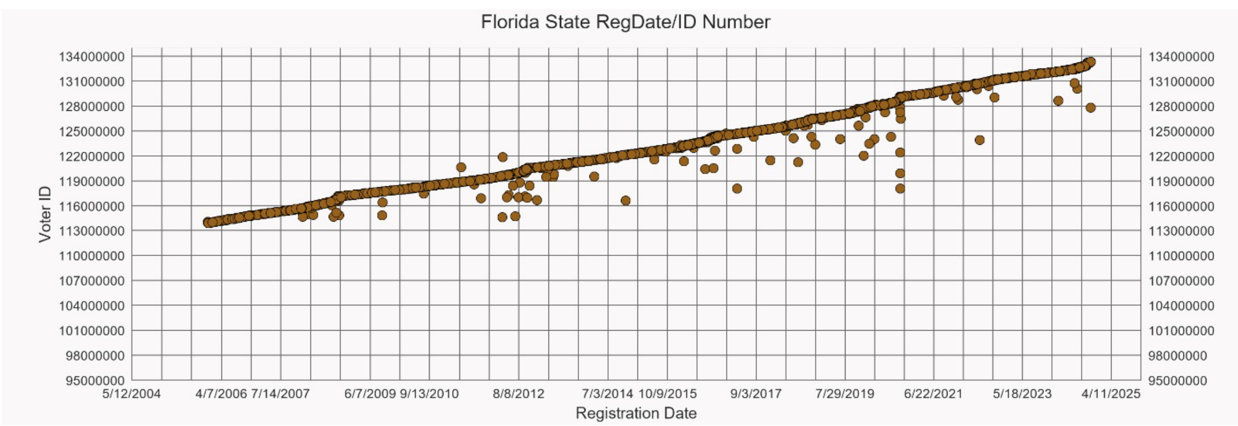


Figure 7 FL Voter ID numbers, Partition 2, >113,938,838 (n= 10,925,179)

A scatterplot of both partitions for any given county reveals the sharp change in character at the end of 2005 (Figure 14).

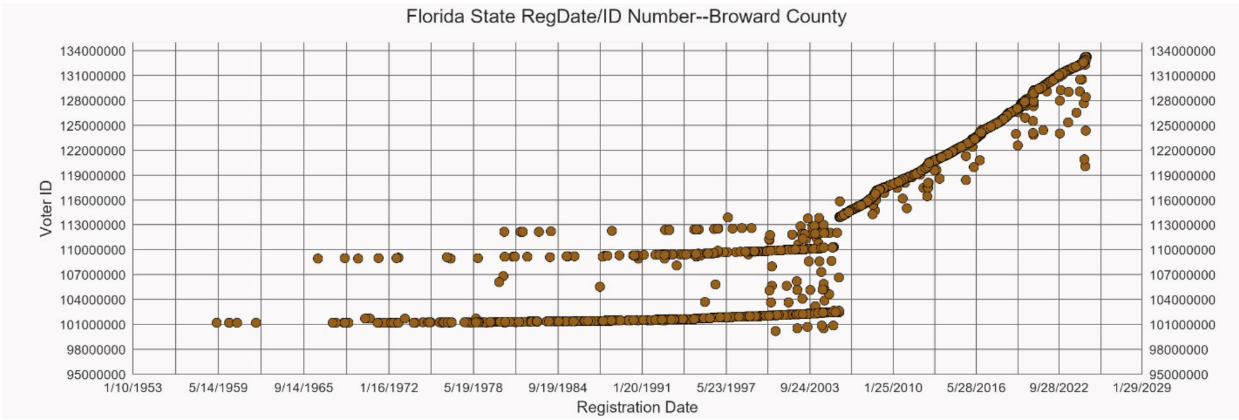


Figure 8 All Voter ID numbers, Partition 1 and 2, Broward County, FL

5.2.2 Temporal Patterns and Data Structure

Modern database design would typically separate the metadata (like county designation) into its own column rather than encoding it within the ID structure. The pre-allocation of multiple fixed ranges to counties creates unnecessary complexity in ID generation and management, while making it difficult to scale or redistribute numbers if a county exhausts its allocation. Moreover, the alphabetical ordering of counties is disrupted by repositioning several counties outside the alphabetical sequence the majority of counties abide by (Table 7). The system appears to prioritize maintaining rigid county-based organization of ID numbers over the flexibility and simplicity offered by contemporary ID assignment methods.

Table 5 VoterID numbers in Partition 1 are mostly in alphabetical order by county, with exceptions (highlighted in gray)

	OKA	OKE	ORA	OSC	PAL	PAS	PIN	POL	PUT	SAN
1	105,974,694	106,124,478	112,682,963	106,150,480	111,745,846	106,321,586	106,685,614	113,415,375	107,448,545	107,498,671
2	105,974,696	106,124,490	112,682,976	106,150,482	111,745,847	106,321,593	106,685,616	113,415,382	107,448,551	107,498,674
3	105,974,697	106,124,508	112,682,978	106,150,483	111,745,848	106,321,596	106,685,618	113,415,385	107,448,552	107,498,677
4	105,974,701	106,124,530	112,682,983	106,150,485	111,745,849	106,321,602	106,685,624	113,415,387	107,448,554	107,498,678
5	105,974,702	106,124,534	112,682,984	106,150,486	111,745,851	106,321,611	106,685,625	113,415,389	107,448,563	107,498,679
6	105,974,703	106,124,535	112,682,991	106,150,488	111,745,852	106,321,614	106,685,627	113,415,391	107,448,571	107,498,680
7	105,974,705	106,124,539	112,682,992	106,150,489	111,745,853	106,321,616	106,685,629	113,415,394	107,448,574	107,498,684
8	105,974,706	106,124,543	112,682,995	106,150,494	111,745,854	106,321,618	106,685,631	113,415,396	107,448,578	107,498,687
9	105,974,711	106,124,549	112,682,997	106,150,497	111,745,857	106,321,620	106,685,648	113,415,399	107,448,580	107,498,688
10	105,974,717	106,124,555	112,682,998	106,150,500	111,745,859	106,321,624	106,685,657	113,415,401	107,448,582	107,498,691

Within Partition 1, chronological sequences exhibit systematic breaks and shifts while maintaining order within segments. For instance, a sequence spanning 1947-2005 might be broken to become "2001-2005, 1947-2000" or further subdivided into "1947-1975, 2001-2005, 1976-2000." These breaks create discontinuities that preserve chronological order within segments while obscuring overall temporal relationships (Table 8).

Table 6 Date breaks within contiguous ranges of VoterID numbers accommodate a shift in record structure based on Registration Date

ALA	ALA	BAK	BAK	BAY	BAY	BRA	BRA	BRE	BRE
100184692	12/20/1995	100593660	3/16/1962	100621972	12/12/1985	100744060	12/15/2005	100761243	1/28/1972
100184694	12/26/1995	100593663	11/7/1978	100621979	7/10/1972	100744062	12/15/2005	100761246	1/4/1958
100184695	12/14/1995	100593665	1/27/1964	100621983	1/18/1972	100744064	12/6/2005	100761262	12/28/1983
100184699	12/26/1995	100593672	1/29/1975	100621993	4/24/1980	100744065	12/1/2005	100761264	4/4/1964
100184703	12/22/1995	100593674	4/23/1980	100621996	7/3/1972	100744066	12/16/2005	100761294	2/7/1958
100184708	12/29/1997	100593678	10/15/1985	100622010	5/4/1982	100744073	1/5/1990	100761327	12/30/1999
100184712	12/29/1997	100593686	11/5/1985	100622012	2/25/1983	100744074	1/5/1990	100761328	12/21/1999
100184713	12/30/1997	100593687	11/5/1985	100622014	2/17/1983	100744075	1/8/1990	100761331	12/21/1999
100184715	12/29/1997	100593688	10/22/1985	100622015	2/17/1983	100744079	8/5/1988	100761332	12/21/1999
100184716	12/30/1997	100593690	11/18/1985	100622017	3/21/1983	100744085	5/15/1967	100761333	12/21/1999

The system's implementation notably deviates from modern database design principles. Rather than storing metadata like county designation in separate columns, this information is encoded within the ID structure itself. The pre-allocation of fixed ranges to counties creates unnecessary complexity in ID generation and management while limiting scalability. Furthermore, the implementation of date breaks varies inconsistently across counties, increasing the level of complexity.

5.2.3 Comparative Analysis

Florida's voter ID numbering system shares common features found across state databases, particularly around HAVA implementation dates. Like New York's pre-2007 period, Florida's pre-2006 system used county-based partitioning, and both states later adopted more chronologically-oriented approaches for some of their post-HAVA ID assignments (Table 9). However, while Florida maintained this simpler chronological system after 2006, New York developed additional algorithmic complexity through multiple concurrent ID assignment methods, some of which retained and expanded partition-based structures.

Table 7 State of Arizona VRAZ ID numbers assigned to counties based on first two digits of number

County	County ID	RID MIN	RID MAX	Gap to next	Range	Alpha VRAZ	VRAZ at RID MIN	VRAZ at RID MAX	Gap to next	VRAZ Range	VRAZ Assigned
Maricopa	8	20,000,000	20,119,041		119,041	M	3,570,350	3,824,576		254,226	67,912
Pima	11	20,119,042	21,393,299	1	1,274,257	P	922,243	2,441,312	-2,902,333	1,519,069	795,089
Maricopa	8	21,393,300	24,983,426	1	3,590,126	M	3,463,164	5,111,834	1,021,852	1,648,670	2,547,141
Graham	5	24,983,427	25,023,352	1	39,925		500,016,919	500,040,007	494,905,085	23,088	37,824
Greenlee	6	25,023,353	25,034,252	1	10,899		600,001,849	600,010,799	99,961,842	8,950	10,267
Apache	1	25,034,255	25,129,781	3	95,526		100,071,637	100,095,207	-499,939,162	23,570	91,735
Gila	4	25,129,782	25,200,650	1	70,868		400,057,769	400,081,304	299,962,562	23,535	66,472
Cochise	2	25,200,652	25,350,563	2	149,911		200,104,575	200,155,336	-199,976,729	50,761	145,852
Santa Cruz	13	25,350,569	25,404,324	6	53,755		1,300,041,762	1,300,052,260	1,099,886,426	10,498	29,552
La Paz	7	25,404,325	25,427,127	1	22,802		700,000,313	700,021,634	-600,051,947	21,321	21,781
Yuma	15	25,427,128	25,600,824	1	173,696		1,500,183,470	1,500,243,574	800,161,836	60,104	95,541
Pinal	12	25,600,826	26,020,121	2	419,295		1,200,258,888	1,200,416,918	-299,984,686	158,030	251,051
Navajo	10	26,020,122	26,216,365	1	196,243		1,000,156,246	1,000,193,460	-200,260,672	37,214	142,828
Mohave	9	26,216,366	26,507,194	1	290,828		900,214,465	900,287,806	-99,978,995	73,341	194,174
Coconino	3	26,507,195	26,713,779	1	206,584		300,001,025	300,204,116	-600,286,781	203,091	187,813
Yavapai	14	26,713,781	27,040,677	2	326,896		1,400,227,087	1,400,325,446	1,100,022,971	98,359	183,961
Maricopa	8	27,040,839	28,315,219	162	1,274,380	NA	NONE	NONE	NONE	NONE	NONE
Mixed		28,315,221	28,560,797	2	245,576	NA	NONE	NONE	NONE	NONE	NONE

New York, conversely, deliberately obscures county relationships through multiple layers of partition-based complexity and non-obvious boundary definitions (Table 10). Florida's approach falls between these extremes: while avoiding New York's extensive obfuscation, it deviates from Arizona's transparency by employing non-obvious county ranges and chronological breaks.

Table 8 NY county ranges for ID numbers mapped using the "Spiral" algorithm (pre-2007 dates)

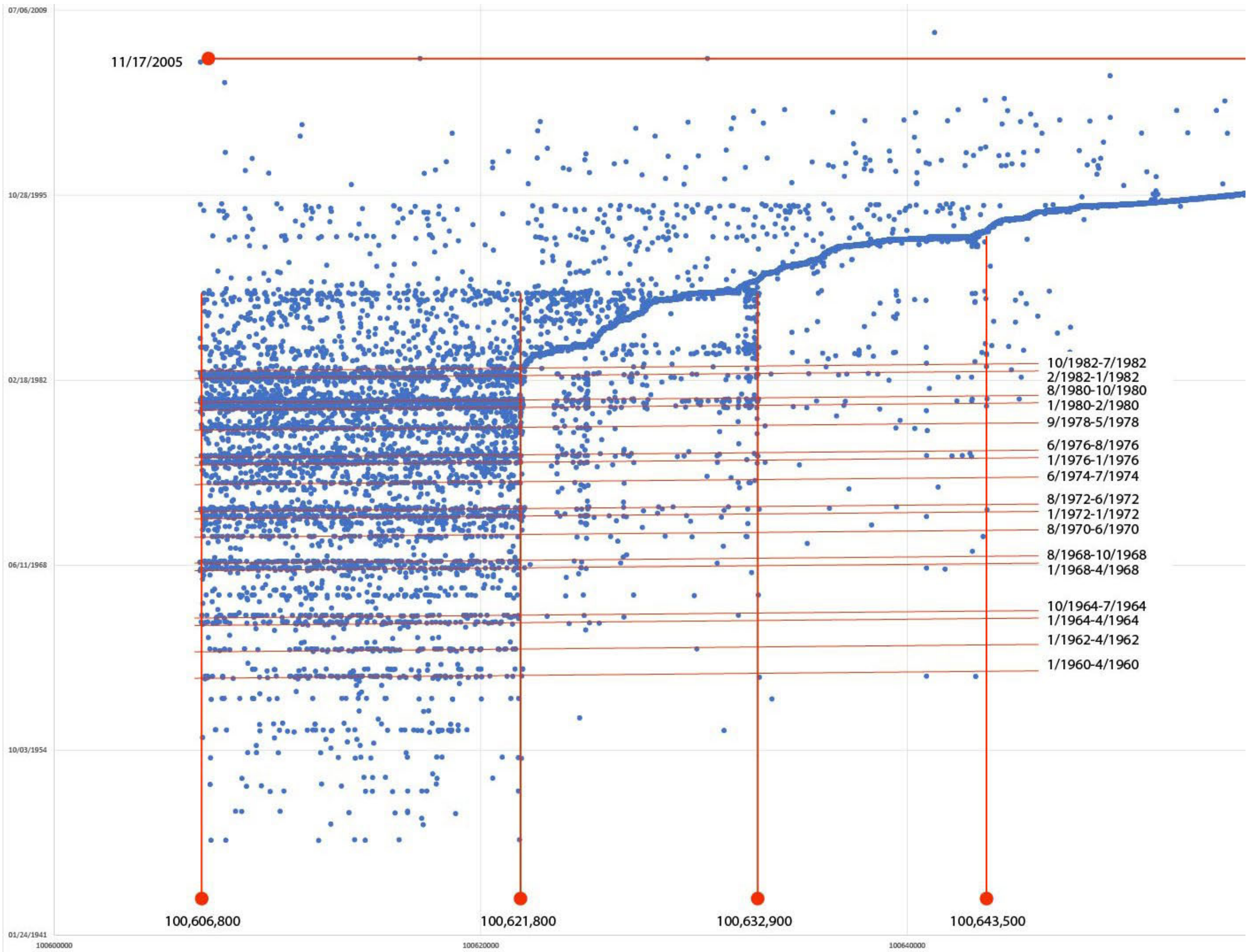
County	County Code	CRID	MIN SBOEID	MAX SBOEID	Gap to previous	MIN to MAX size	Used Numbers (NYBOE)	Percent used of available	Registered full range (NYBOE)
Out of range		1	0	8,502,558	0		127		
Schoharie	47	1.01	8,502,559	8,521,213	1	18,655	18,447	98.89%	21,134
Buffer 1		2	8,521,214	9,091,766	1		0		
Onondaga	34	2.01	9,091,767	9,382,492	1	290,726	290,015	99.76%	329,306
Schenectady	46	2.02	9,382,494	9,477,662	2	95,169	94,044	98.82%	109,164
Oswego	38	2.03	9,477,664	9,557,731	2	80,068	79,339	99.09%	83,022
Niagara	32	2.04	9,557,733	9,694,852	2	137,120	132,790	96.84%	150,686
Suffolk	52	2.05	9,694,854	10,584,725	2	889,872	882,440	99.16%	1,116,934
Essex	16	2.06	10,584,727	10,611,715	2	26,989	26,685	98.87%	27,222
Buffer 2		3	10,611,716	20,005,105	1		0		
Hamilton	21	3.01	20,005,106	20,010,209	1	5,104	5,054	99.02%	4,677
Columbia	11	3.02	20,010,211	20,054,800	2	44,590	43,484	97.52%	49,665
Franklin	17	3.03	20,054,802	20,082,320	2	27,519	27,314	99.26%	29,083
Warren	57	3.04	20,082,322	20,125,302	2	42,981	42,479	98.83%	48,505
Fulton	18	3.05	20,125,304	20,157,206	2	31,903	31,577	98.98%	35,632
Tioga	54	3.06	20,157,208	20,189,678	2	32,471	32,254	99.33%	35,581
Montgomery	29	3.07	20,189,680	20,221,054	2	31,375	31,080	99.06%	30,712
Seneca	49	3.08	20,221,056	20,241,573	2	20,518	20,309	98.98%	22,052
Madison	27	3.09	20,241,575	20,284,049	2	42,475	42,106	99.13%	45,868
Allegany	2	3.1	20,284,051	20,312,118	2	28,068	27,580	98.26%	27,588

5.2.4 Technical Implications

The system's design introduces hidden complexity in database operations while potentially complicating analysis of registration patterns. The non-standard approach to temporal ordering and county allocation suggests prioritization of organizational principles beyond basic database management needs. This raises questions about the system's design choices, particularly given the availability of simpler alternatives as demonstrated by Arizona's simpler county partitioning.

5.2.5 Voter ID Algorithms

The following analysis examines Florida voter registration records using a scatterplot visualization where the x-axis represents voter ID numbers and the y-axis shows registration dates. In a typical voter registration database, we would expect to see a roughly diagonal pattern with newer registrations receiving higher ID numbers in chronological sequence. However, examination of Florida's pre-1982 registration data revealed unusual patterns that warranted deeper investigation, particularly given the importance of voter roll integrity and the fundamental database management principle that ID numbers should be unique identifiers assigned in a consistent manner.



These scatterplots reveal anomalous data management in Florida's pre-1982 voter registration records. The plots show voter ID numbers (x-axis) versus registration dates (y-axis), with registration dates organized into distinct horizontal bands at different ID number ranges. Multiple bands contain overlapping ID numbers for the same time periods, violating fundamental database principles of unique identifiers. The artificial grouping of dates, reversed chronological sequences within years, and precise segmentation of ID blocks (e.g., 100606800, 100621800) indicate systematic manipulation rather than natural record accumulation.

This pattern is highly unusual from a database management perspective because it contradicts standard practices of unique, sequential ID assignment and chronological record keeping. The abrupt transition to normal sequential ID assignment in 1982 further highlights the artificial nature of the pre-1982 structure. The precision and regularity of these patterns suggest deliberate obfuscation of historical records rather than legitimate data management or reconstruction efforts.

Florida's voter ID system operates under several significant technical constraints that limit opportunities for algorithmic manipulation. The extremely narrow gaps between sequential ID numbers - often just single digits - provide almost no mathematical space for encoding additional information, unlike New York's dual-ID system which could maintain perfect sequences while arbitrarily mapping between number sets. Additionally, the county-based organization adds no steganographic value since county identity is already stored explicitly in the database, and the registration date discontinuities, while unusual, lack the supporting structures needed to encode hidden attributes.

While these structural peculiarities in the ID system appear less sophisticated than mechanisms found in other states, their presence creates systematic discontinuities and complexity in the data structure. This complexity may serve to obscure the more sophisticated algorithmic manipulation revealed through birth date analysis. Unlike New York's system, where algorithmic manipulation was built directly into ID assignment, Florida's system appears to rely on more subtle methods of record manipulation. The discovery of a precise algorithm for generating false birth dates while maintaining age matches demonstrates that Florida's voter roll manipulation, while differently implemented, is no less concerning than that found in other states. The structural oddities in the ID system may serve as one layer of complexity in a multi-faceted approach to obscuring systematic record manipulation.

6 CONCLUSION

This investigation of Florida's voter registration database has revealed compelling evidence of systematic manipulation across multiple dimensions.

The statistical analysis demonstrates that age-matching patterns among voter records occur at rates that cannot be explained by natural demographic patterns or administrative error. The investigation uncovered 118,615 records with logically impossible registration dates added during the November 2024 election period. These temporal contradictions - records claiming registration dates that predate their first appearance in the database - provide independent evidence of database manipulation.

Most significantly, analysis of just 4 out of 18 available elections revealed 7,044 instances where both records in matched pairs recorded votes in the same election, with many pairs showing this pattern across multiple elections. This consistent pattern of duplicate voting activity strongly suggests systematic use of these records rather than isolated anomalies.

The scale and sophistication of these patterns suggest deliberate modification of Florida's voter registration system at an administrative level. Unlike potential irregularities in vote counting or ballot handling that might be explained by human error or process failures, these systematic patterns in core database fields indicate intentional manipulation requiring both technical access and understanding of the system's architecture.

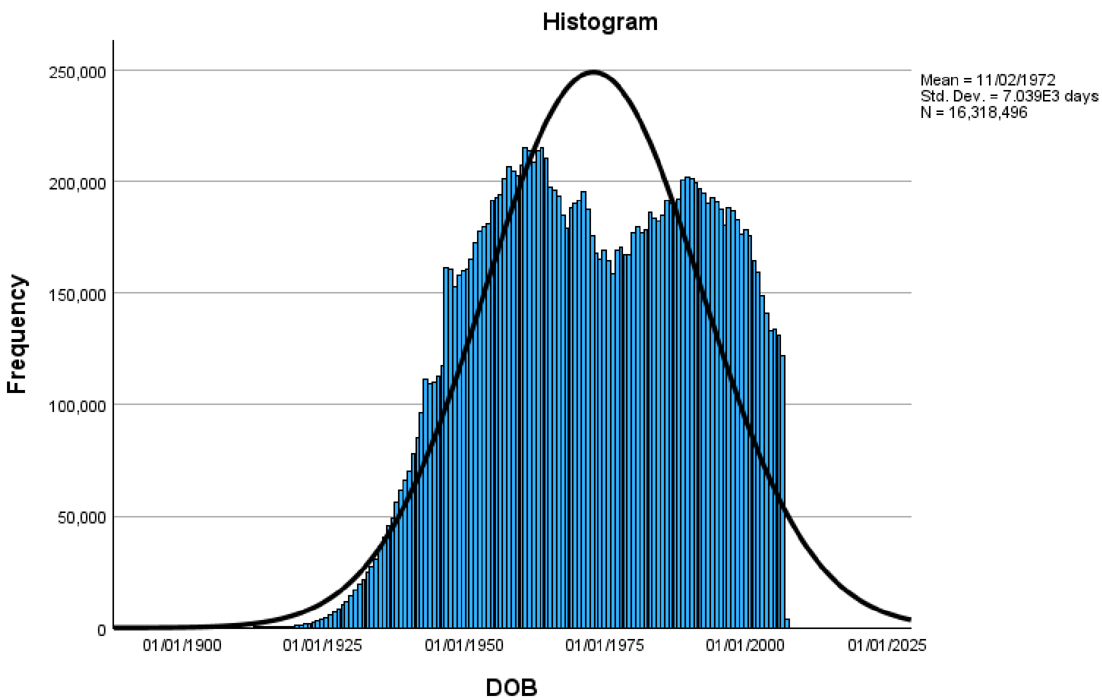
These findings warrant immediate investigation by appropriate authorities. The patterns documented here represent potential violations of the Help America Vote Act's requirement for maintaining accurate voter registration databases. More broadly, they raise serious concerns about the integrity of Florida's voter registration system and its vulnerability to systematic manipulation.

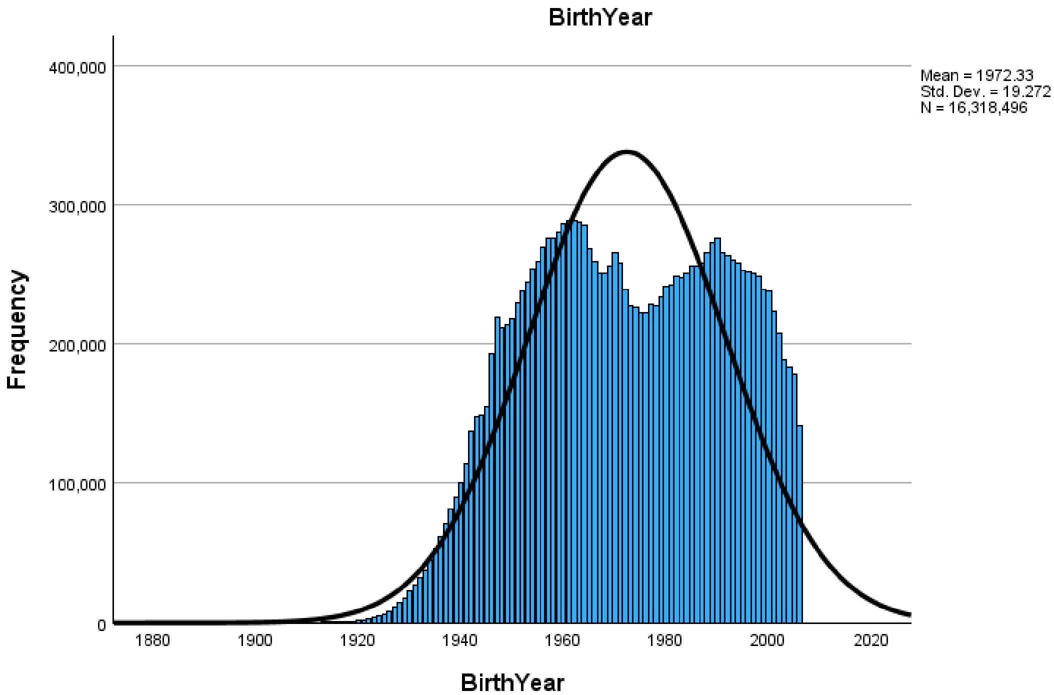
7 APPENDIX

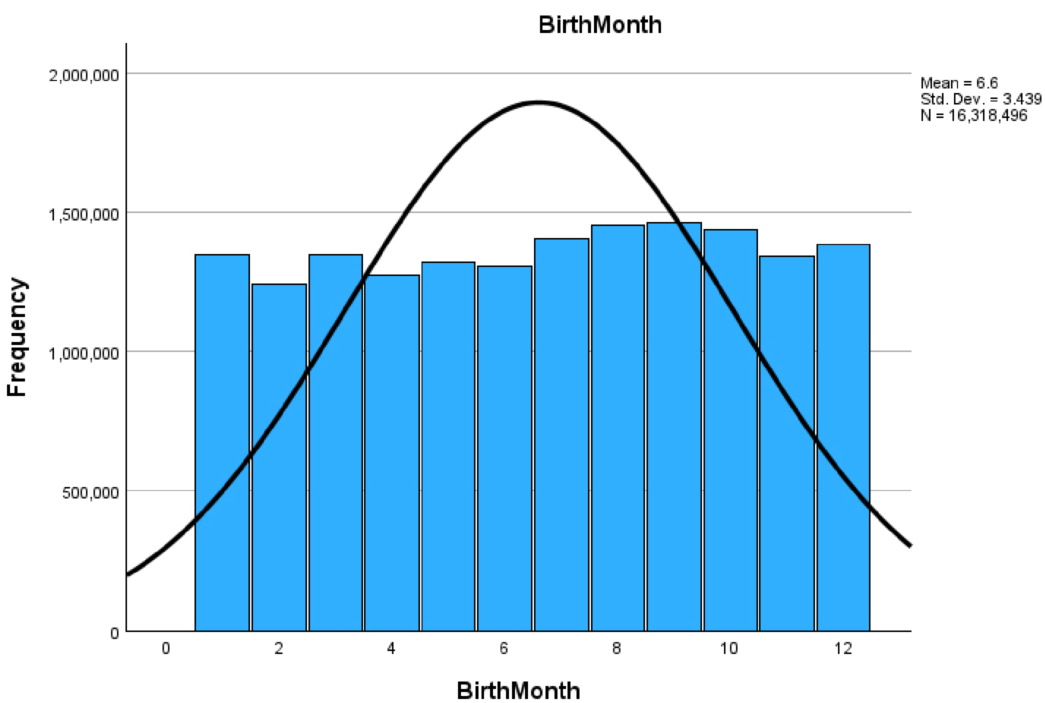
7.1 DOB AND DOB DECOMPOSED FOR FULL DATASET (16M RECORDS)

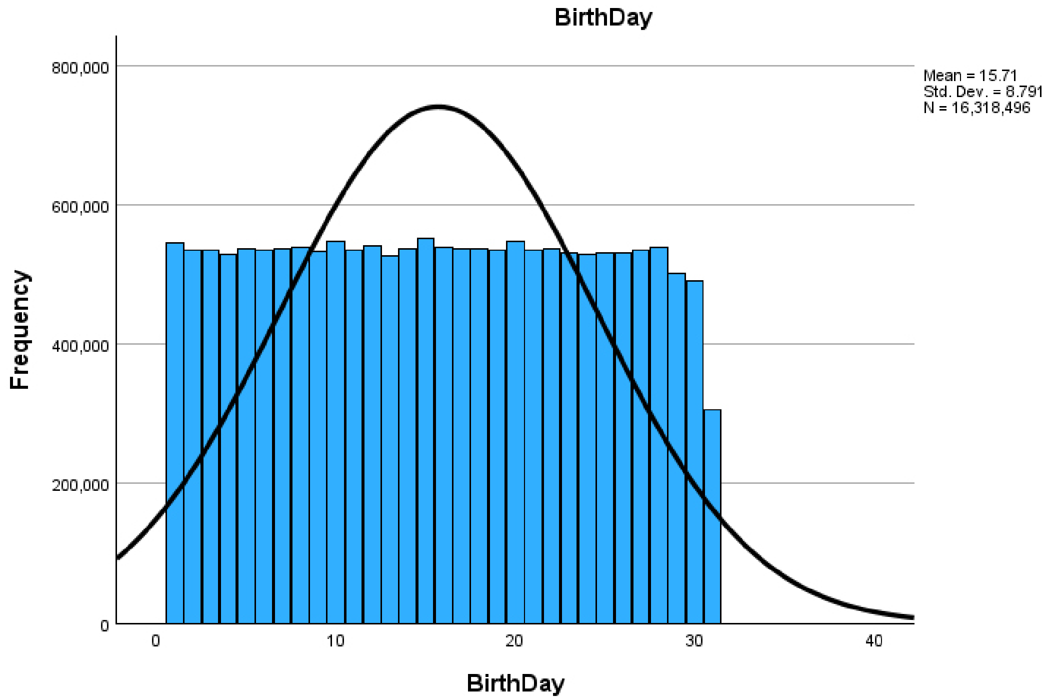
The full Florida voter registration database (N=16,318,496) shows expected demographic distributions across all date-of-birth components:

- Birth Years follow a natural population curve centered around 1972 (mean=1972.33, SD=19.272), with a gradual rise from 1940, peak in the 1960-1980 period, and decline toward 2000, reflecting the expected age distribution of registered voters.
- Birth Months show an approximately uniform distribution across all twelve months (mean=6.6, SD=3.439), with minor natural seasonal variations consistent with known birth patterns.
- Birth Days demonstrate a remarkably even distribution from days 1-30 (mean=15.71, SD=8.791), with the expected dropoff at day 31 due to varying month lengths. The uniformity across days 1-30 is consistent with natural birth patterns.

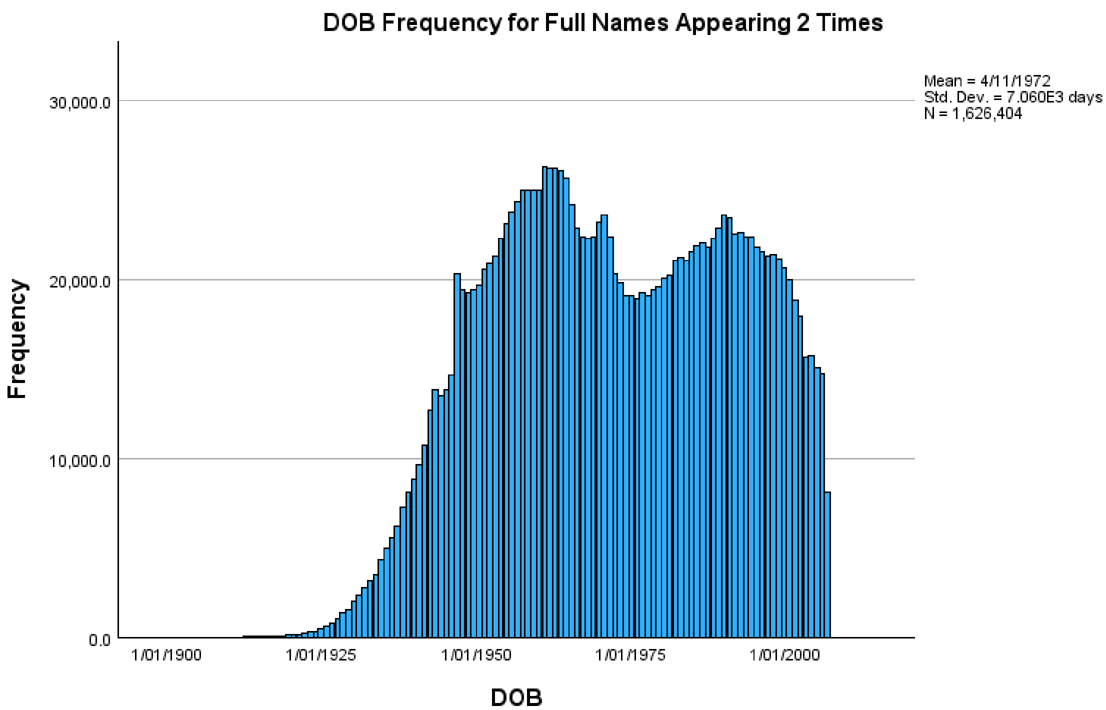


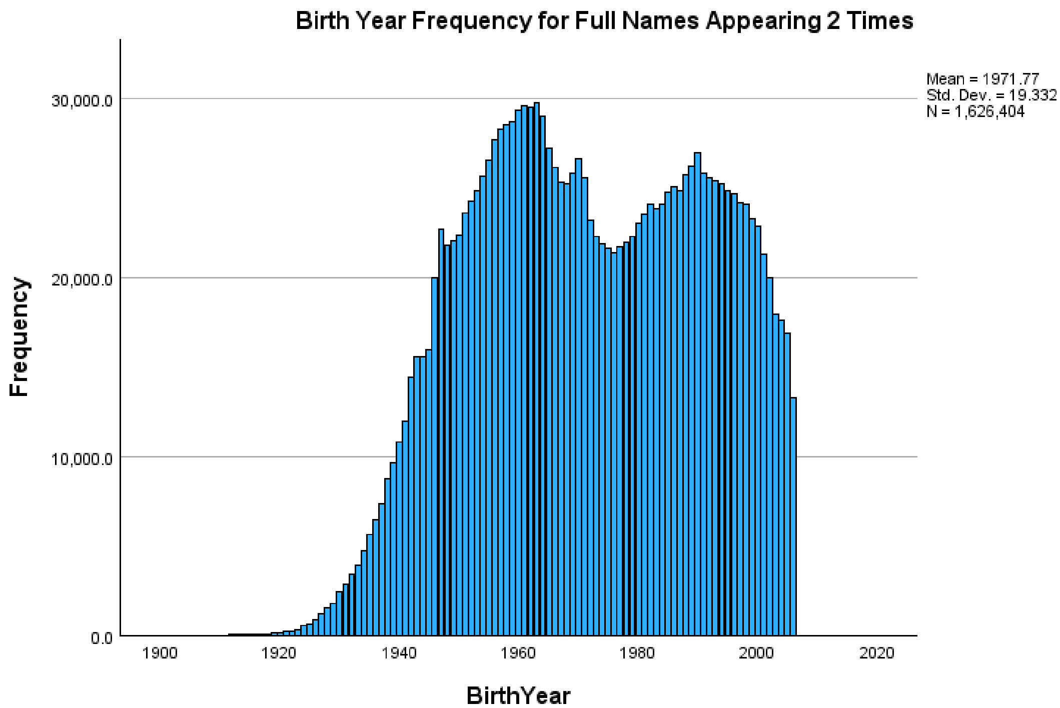

BirthYear
Mean = 1972.33
Std. Dev. = 19.272
N = 16,318,496
Frequency
400,000
300,000
200,000
100,000
0
1880
1900
1920
1940
1960
1980
2000
2020
BirthYear

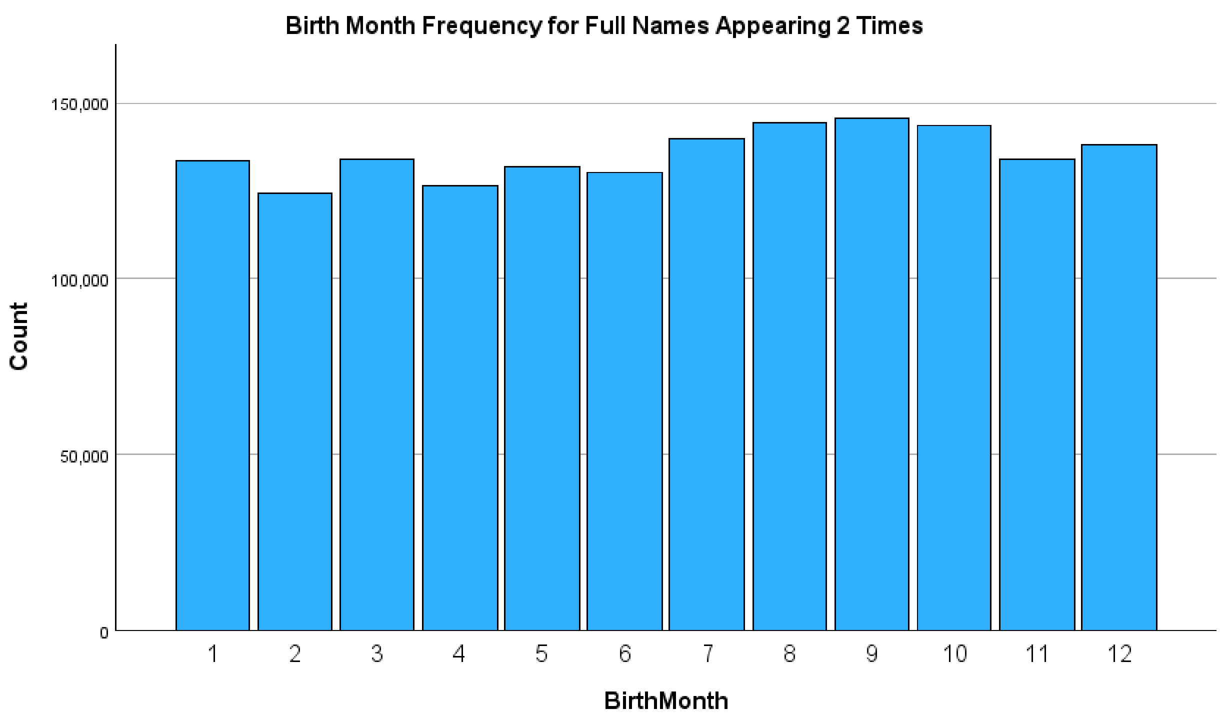

BirthMonth
Mean = 6.6
Std. Dev. = 3.439
N = 16,318,496
Frequency
2,000,000
1,500,000
1,000,000
500,000
0
0
2
4
6
8
10
12
BirthMonth

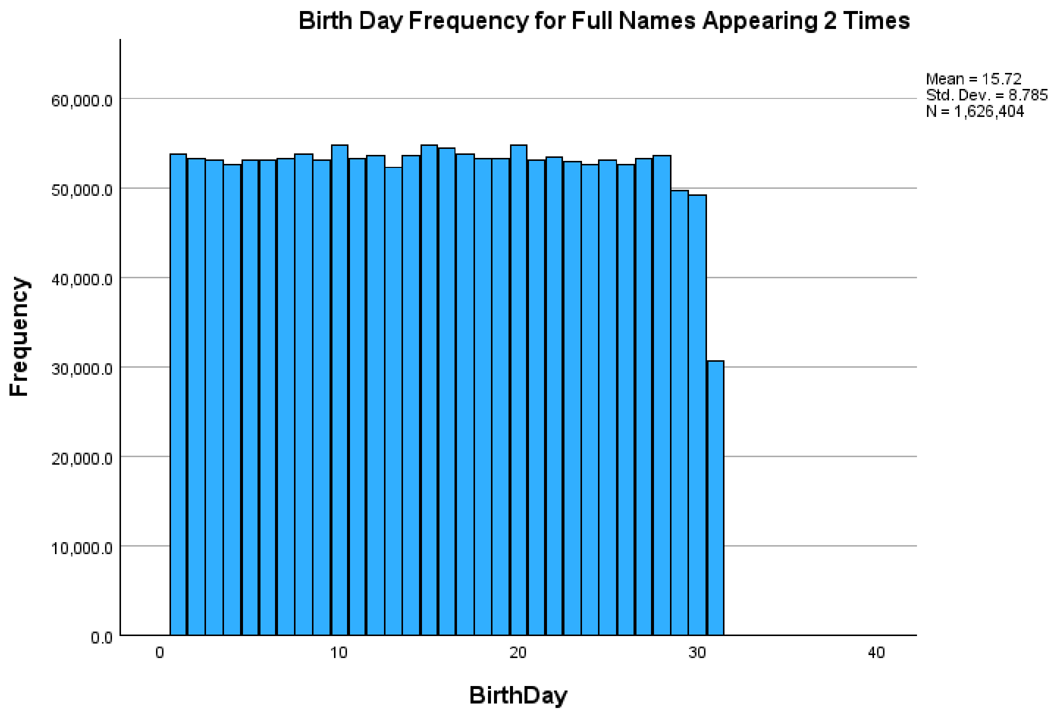


7.2 Frequencies for all full names that appear twice in the database



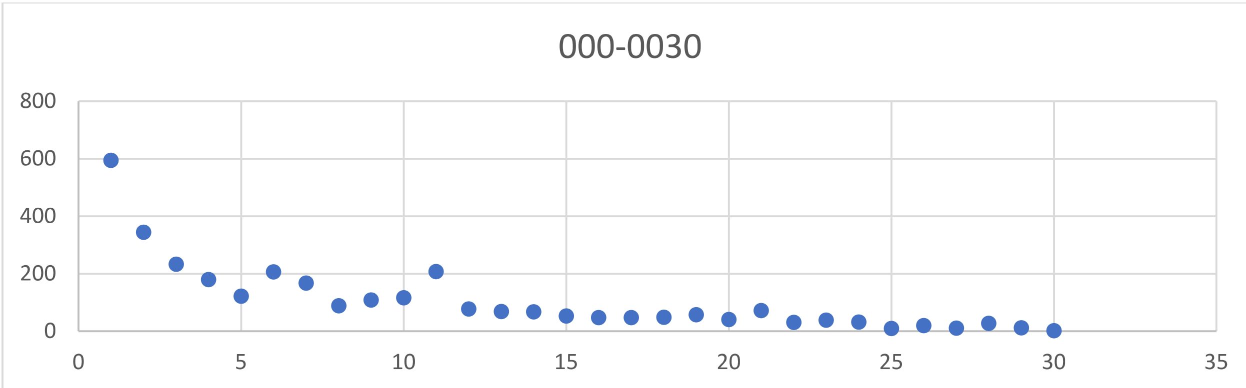

Birth Year Frequency for Full Names Appearing 2 Times
Mean = 1971.77
Std. Dev. = 19.332
N = 1,626,404
Frequency
30,000.0
20,000.0
10,000.0
0.0
1900
1920
1940
1960
1980
2000
2020
BirthYear

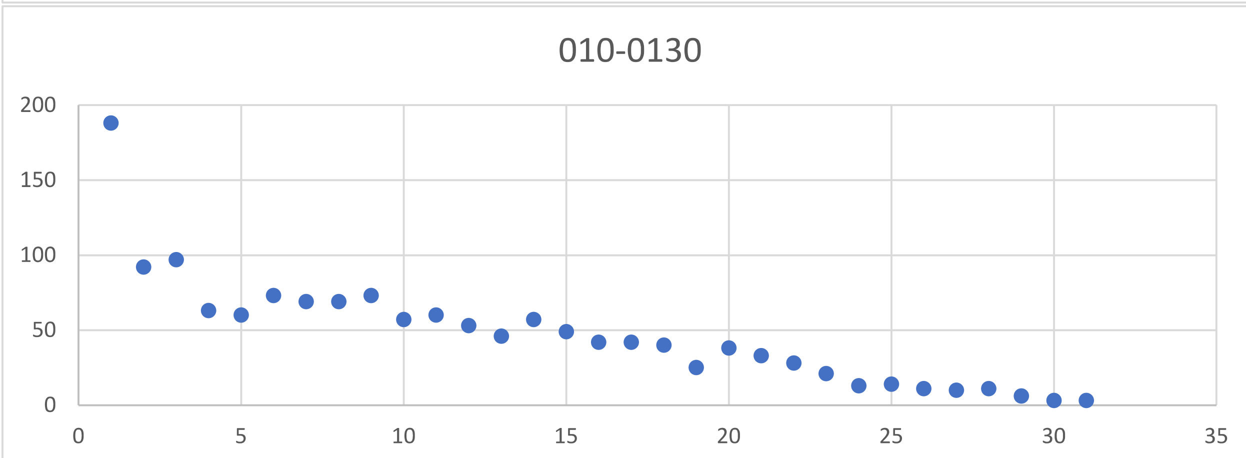

Birth Month Frequency for Full Names Appearing 2 Times
Count
150,000
100,000
50,000
0
1
2
3
4
5
6
7
8
9
10
11
12
BirthMonth

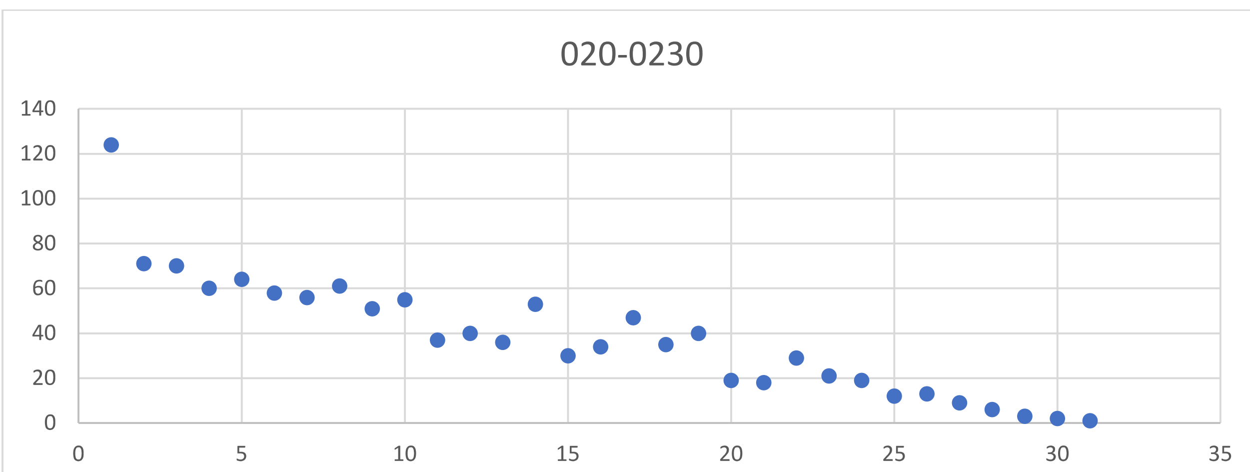


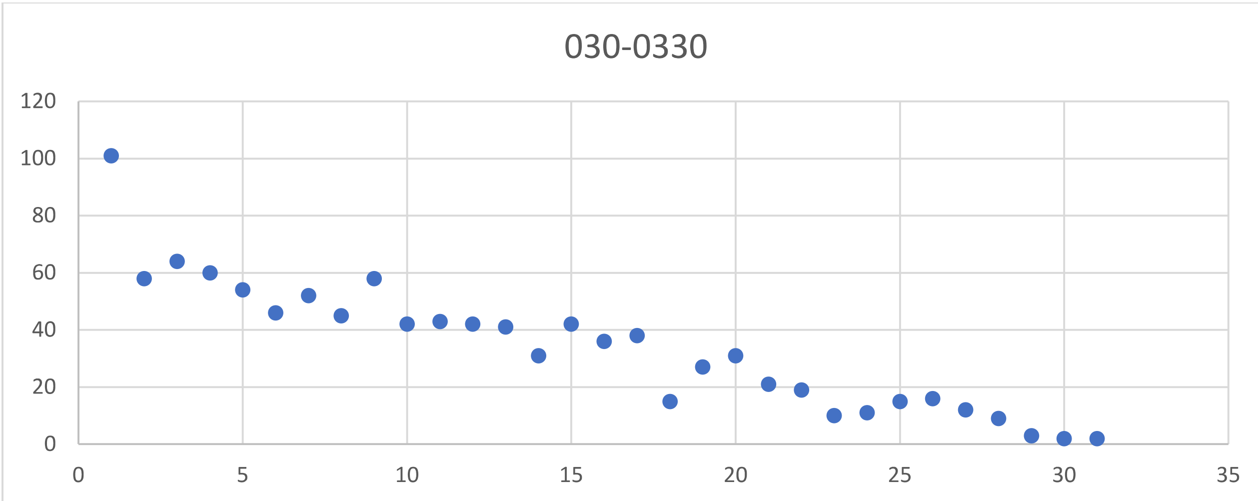
7.3 GAP PATTERN FREQUENCY SCATTERPLOTS

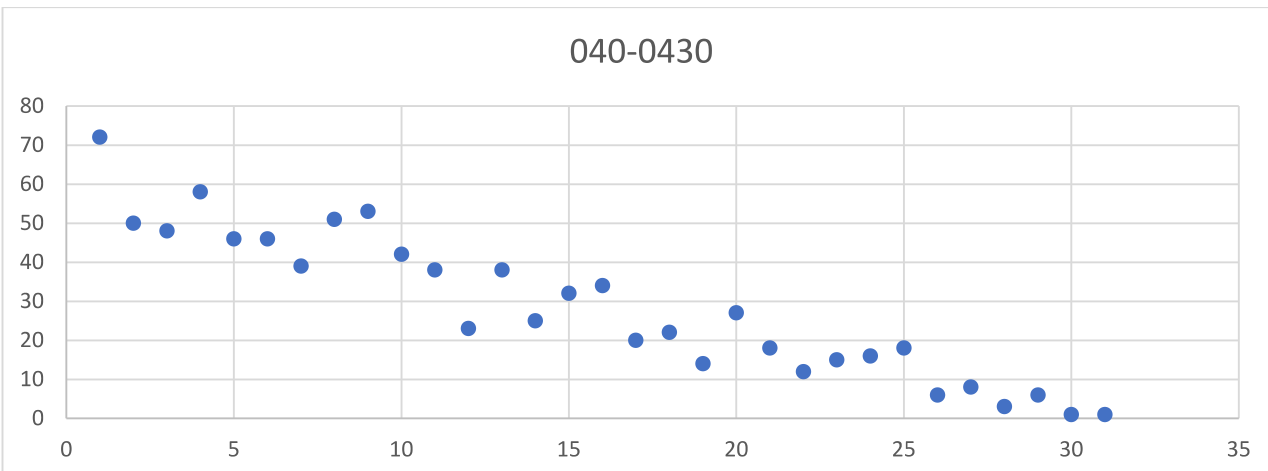
The following figures show the frequency distribution of date-of-birth gaps between record pairs, grouped by month pattern (000-030, 010-0130, etc.). Each chart demonstrates the same exponential decay pattern at different scales, with frequency decreasing as the time gap increases. The consistent pattern across different month groupings suggests systematic record generation rather than natural occurrence.

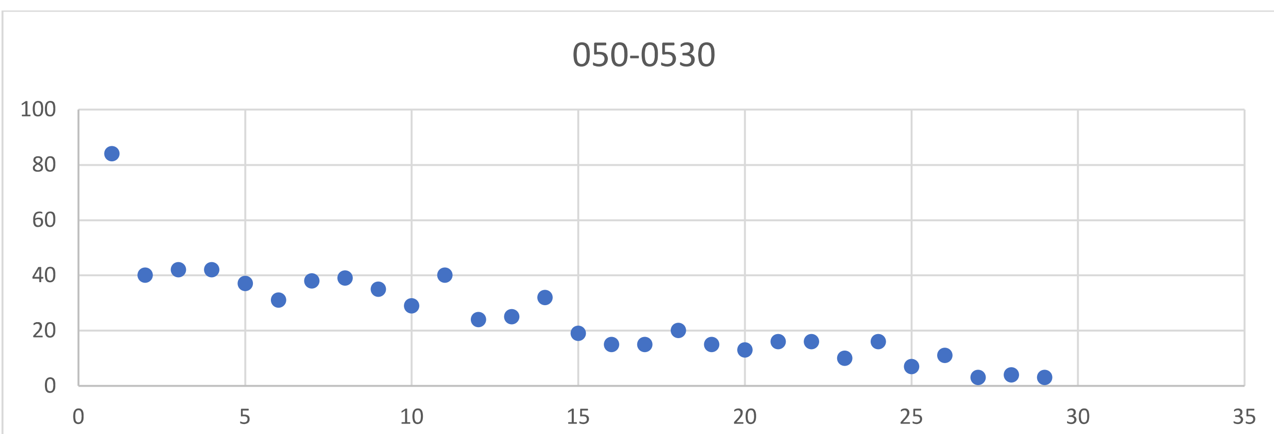

000-0030
800
600
400
200
0
0
5
10
15
20
25
30
35

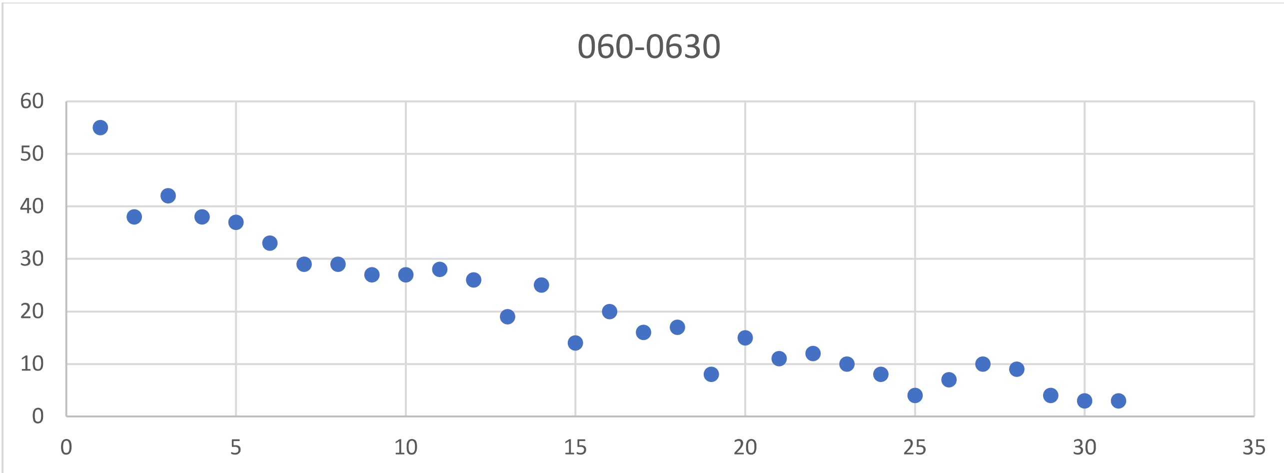

010-0130
200
150
100
50
0
0
5
10
15
20
25
30
35

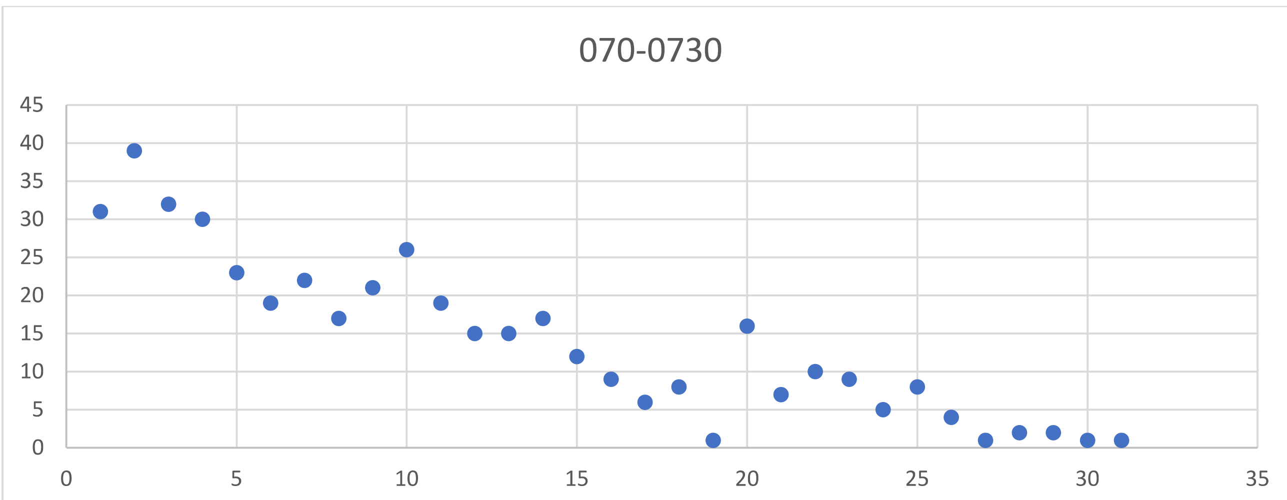

020-0230
140
120
100
80
60
40
20
0
0
5
10
15
20
25
30
35

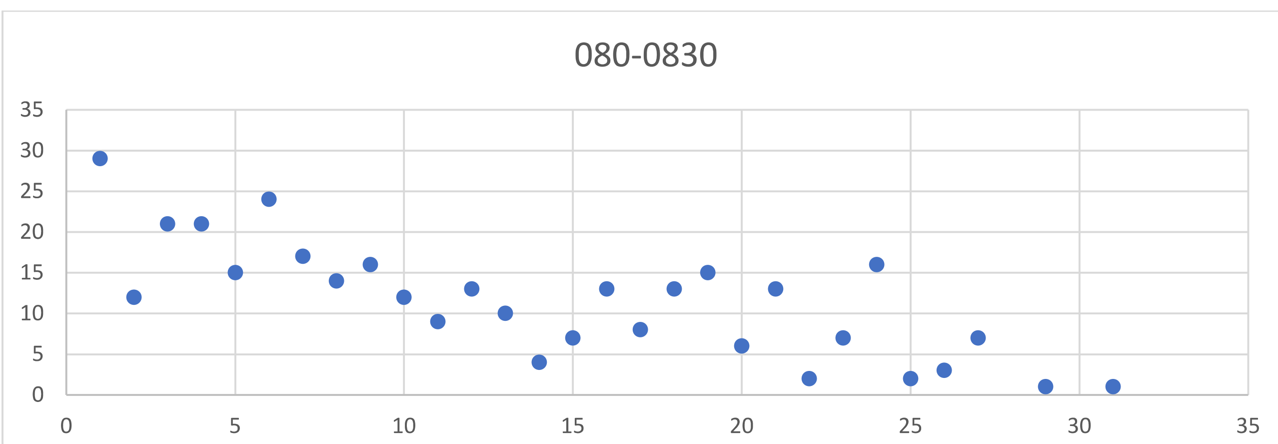

030-0330
120
100
80
60
40
20
0
0
5
10
15
20
25
30
35

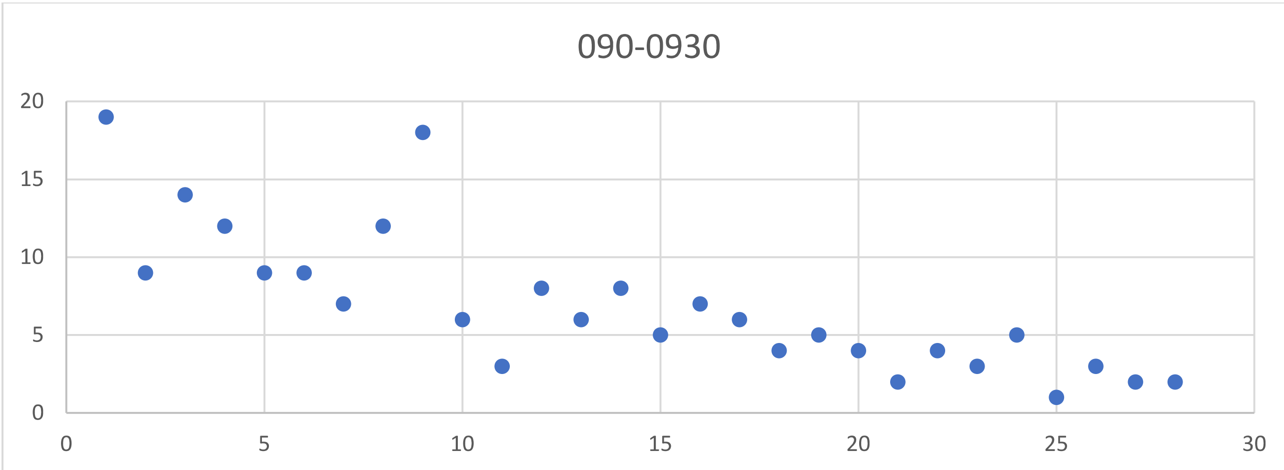

040-0430
80
70
60
50
40
30
20
10
0
0
5
10
15
20
25
30
35

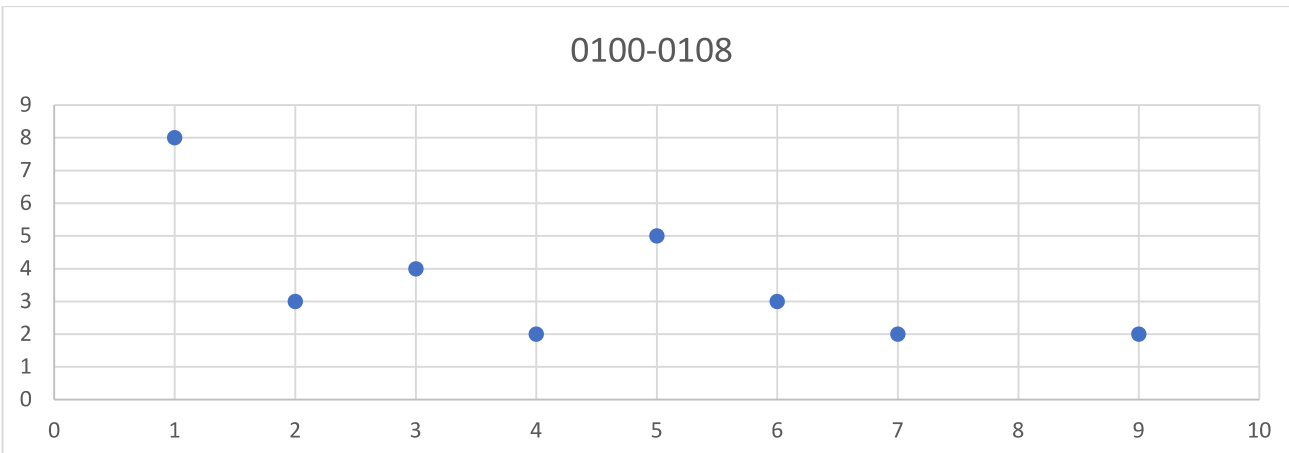

050-0530
100
80
60
40
20
0
0
5
10
15
20
25
30
35


060-0630
60
50
40
30
20
10
0
0
5
10
15
20
25
30
35


070-0730
45
40
35
30
25
20
15
10
5
0
0
5
10
15
20
25
30
35


080-0830
35
30
25
20
15
10
5
0
0
5
10
15
20
25
30
35





8 REFERENCES

Help America Vote Act, 107-252, United States, 107 Stat. 66 (2002 2002).

Paquette, A. (2023). The Caesar cipher and stacking the deck in New York State voter rolls *Journal of Information Warfare, 22*(2), 86-105.

Paquette, A. (2024). New Jersey voter ID numbers reconfigured with shift cipher. *(In-Press)*.

---

[i] This item is not for public dissemination. It is presented here solely as a datapoint that indicates the value of obtaining formal access to a well-kept non-governmental database for comparison.