5301 Final Project

US 2020 vs 2024 Election Results Analysis

Austin Steel
Shalini Dronadula
Karunakar Kalvala

Executive Summary

	Diff D	Diff D %	Diff R	Diff R %	Diff Other	Diff Other %	Diff Total
Mean	-129,746	-2.16%	56,394	2.27%	-2,145	-0.11%	-75,498
95% Min	-213,548	-2.54%	26,907	1.93%	-8,784	-0.26%	-151,996
95% Max	-45,944	-1.78%	85,882	2.60%	4,492	0.04%	1,001
P-Value	0.0031	0.0001	0.0003	0.0001	0.5191	0.1542	0.053
Null Hyp	Reject	Reject	Reject	Reject	Fail To Reject	Fail To Reject	Reject, note P
Alt Hyp	Evidence For	Evidence For	Evidence For	Evidence For			Evidence For

For our project, we decided to look at the results from the 2020 and 2024 US presidential election. Our dataset contains the vote count and percent for each state for each party (Republican, Democrat, and Other) party in both respective elections.

Both elections had very unique circumstances, such as Covid in 2020 and the late Biden-Kamala switch in the Democrat candidate in 2024, which made them unlike any elections prior. At the same time, both elections had many similarities, such as the Republican candidate and popular issues among voters like immigration, abortion, and the economy.

Goal: We thought it would be of interest to see if there were statistically significant differences in the vote totals, and vote percents for each party in the two elections. To do this, we performed several paired t-test, comparing the mean votes for each party in both elections.

Data Preprocessing

- Data was collected from NBC.News
- We have 15 columns and 52 rows (each state, DC and Total)
- No missing values
- Cleaning the table up
- Added columns to calculate the difference for each political party in each election. (Ex: Difference in D Votes = 2024 D votes – 2020 D votes)
- All values are numerical entries except the State column.
- We then started by looking basic data analysis of our dataset.

Initial Data Exploration Discovery

The MEANS Procedure

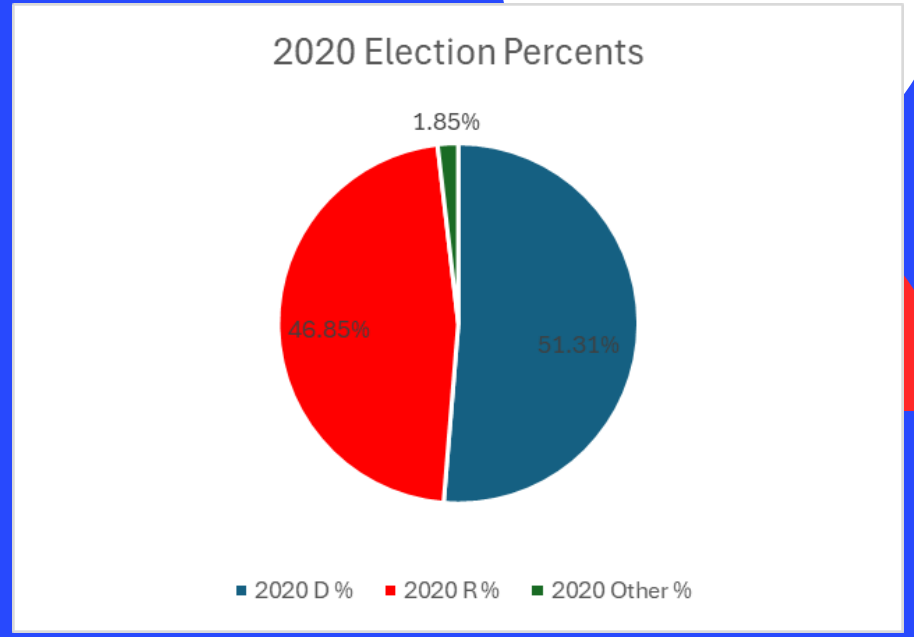
Variable	Label	Mean	Std Dev	Minimum	Maximum	Median
2020 D Votes	2020 D Votes	1593794.14	1916171.84	73491.00	11110639.00	856034.00
2020 D %	2020 D %	0.4863000	0.1204811	0.2655000	0.9215000	0.4936000
2020 R Votes	2020 R Votes	1455372.06	1413252.57	18586.00	6006518.00	1020280.00
2020 R %	2020 R %	0.4920353	0.1198747	0.0540000	0.6994000	0.4924000
2020 Other Votes	2020 Other Votes	57297.16	59618.78	7475.00	384223.00	36685.00
2020 Other %	2020 Other %	0.0216588	0.0075623	0.0092000	0.0439000	0.0200000
2020 Total Votes	2020 Total Votes	3106463.35	3306713.21	276765.00	17501380.00	2148062.00
2024 D Votes	2024 D Votes	1464047.82	1654612.71	69527.00	9253875.00	772487.00
2024 D %	2024 D %	0.4647047	0.1163001	0.2584186	0.9027707	0.4750049
2024 R Votes	2024 R Votes	1511766.65	1493414.65	21076.00	6393036.00	1036213.00
2024 R %	2024 R %	0.5147088	0.1186751	0.0646763	0.7159800	0.5086213
2024 Other Votes	2024 Other Votes	55151.22	73271.75	6888.00	503686.00	32950.00
2024 Other %	2024 Other %	0.0205865	0.0082181	0.0050684	0.0406149	0.0190003
2024 Total Votes	2024 Total Votes	3030965.69	3139476.97	269048.00	15818118.00	2070101.00
Change_D_Votes		-129746.31	297958.52	-1856764.00	74385.00	-53086.00
Change_R_Votes		56394.59	104842.91	-53728.00	502689.00	20284.00
Change_Other_Votes		-2145.94	23600.71	-79883.00	119463.00	-2413.00
Change_Total_Votes		-75497.67	271991.03	-1683262.00	251177.00	-18487.00
Change_D_Percent		-0.0215953	0.0134863	-0.0556468	0.0015807	-0.0211313
Change_R_Percent		0.0226736	0.0118741	0.0049457	0.0615957	0.0183685
Change_Other_Percent		-0.0010723	0.0052932	-0.0142636	0.0105514	-0.0013532

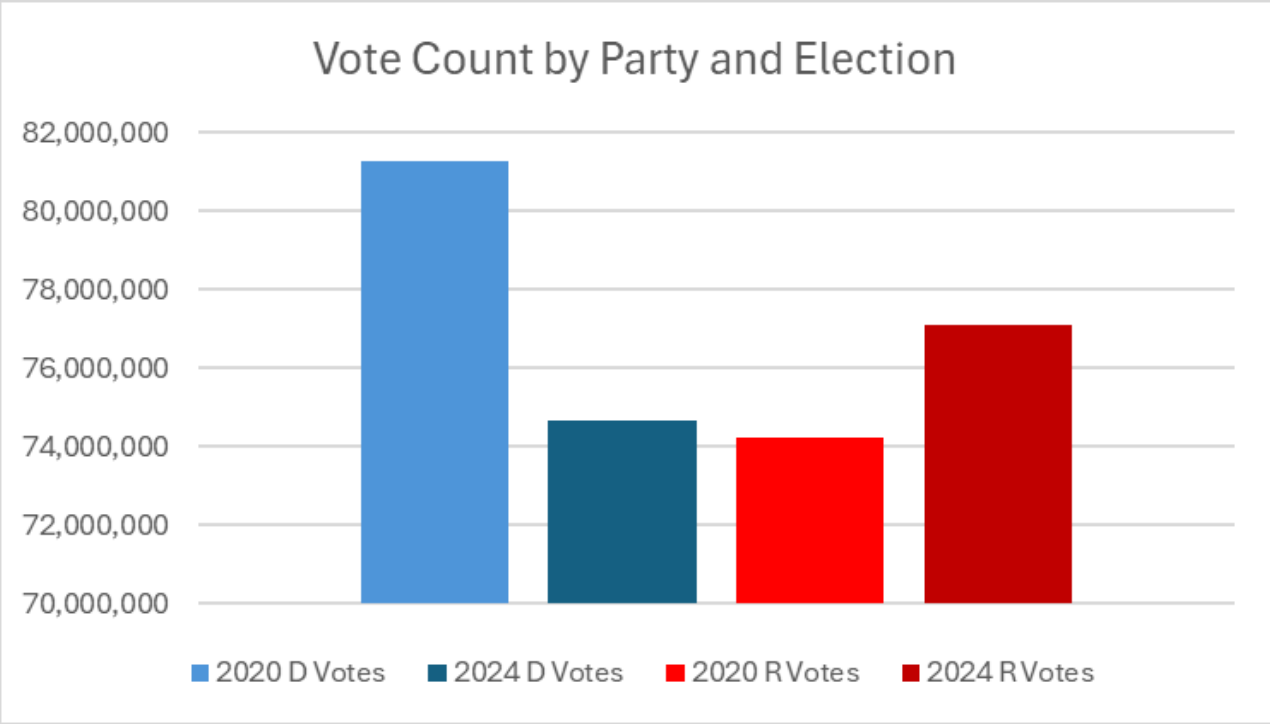
Variable	Label	Mean
Change_D_Votes		-129746.31
Change_R_Votes		56394.59
Change_Other_Votes		-2145.94
Change_Total_Votes		-75497.67
Change_D_Percent		-0.0215953
Change_R_Percent		0.0226736
Change_Other_Percent		-0.0010723

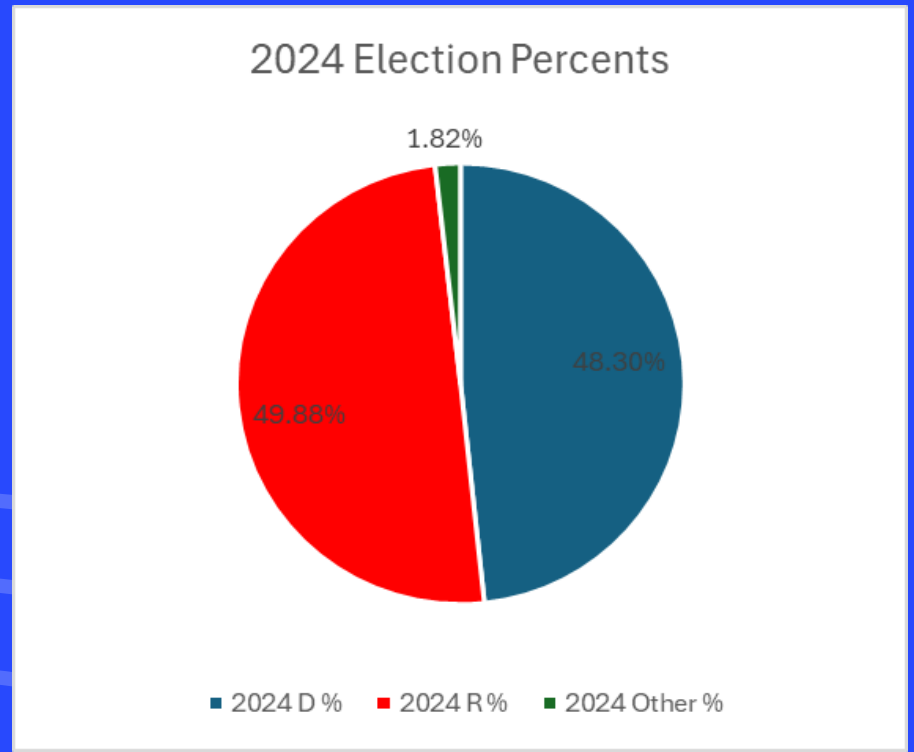
We will be testing the mean votes for each party and mean percent for each party for the 2020 versus 2024 elections.

Initial Data Exploration Discovery

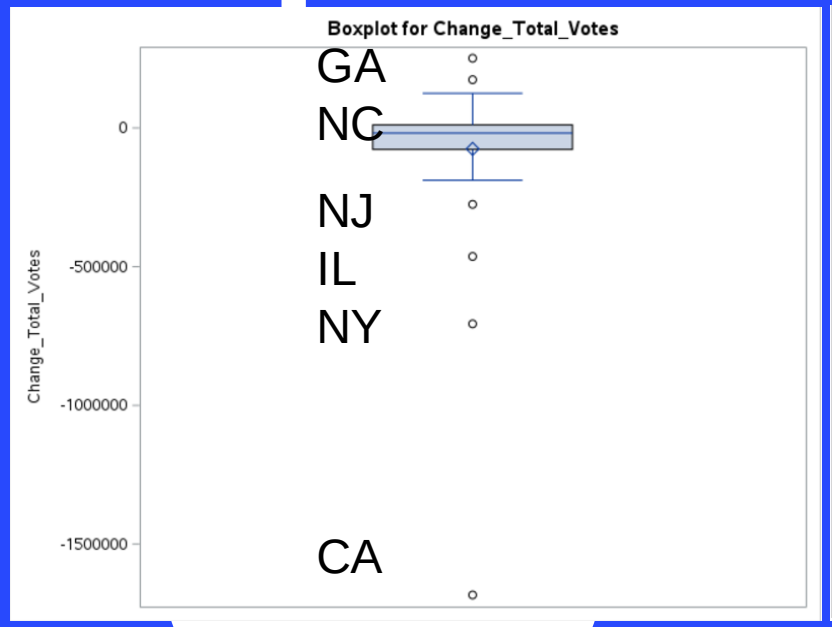
	D Votes	R Votes	Other Votes	Total
2020	81,283,501	74,223,975	2,922,155	158,429,631
2024	74,666,439	77,100,099	2,812,712	154,579,250
Change	(6,617,062)	2,876,124	(109,443)	(3,850,381)

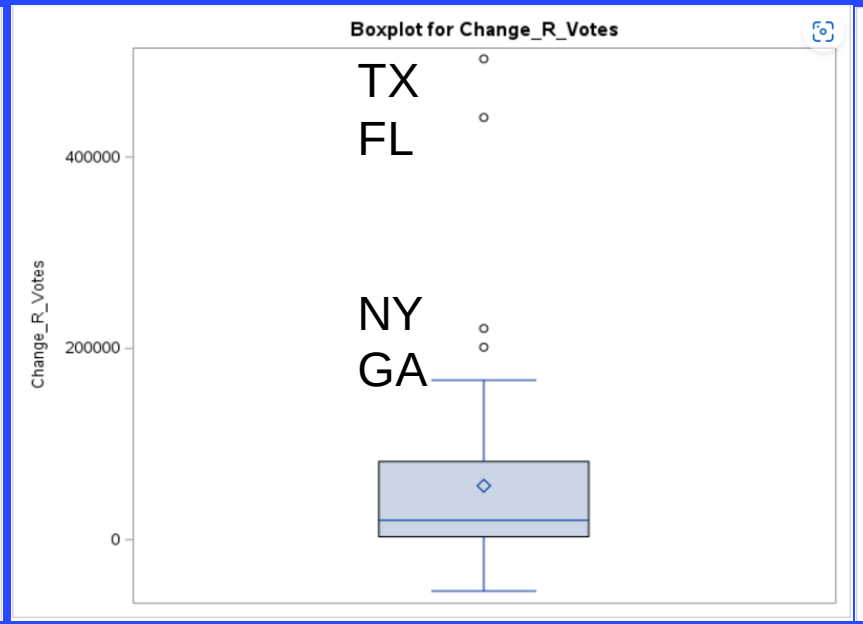


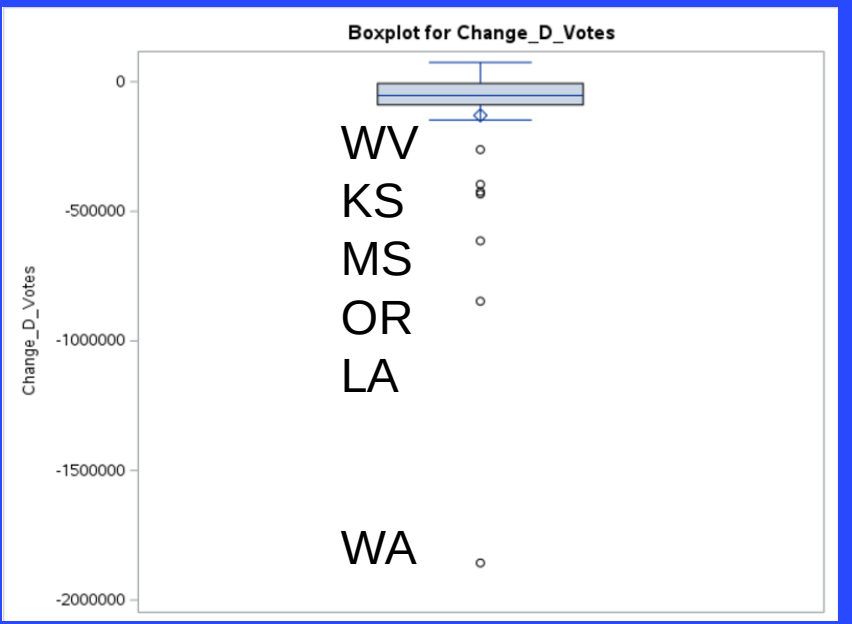


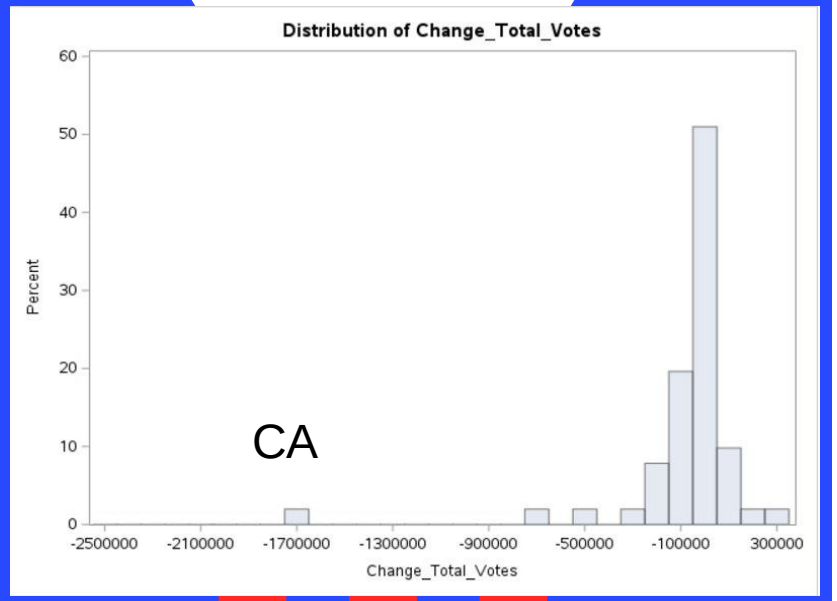


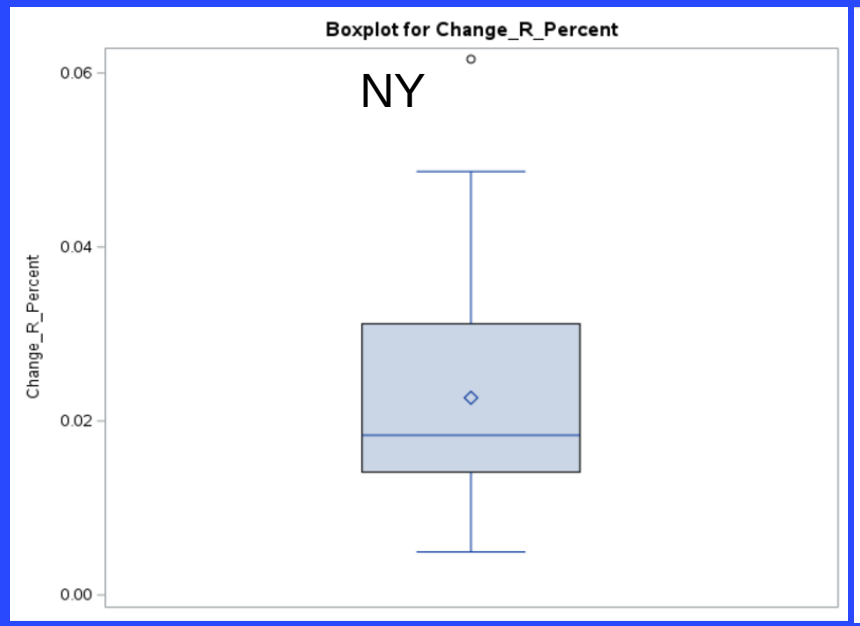
Boxplots and Distributions

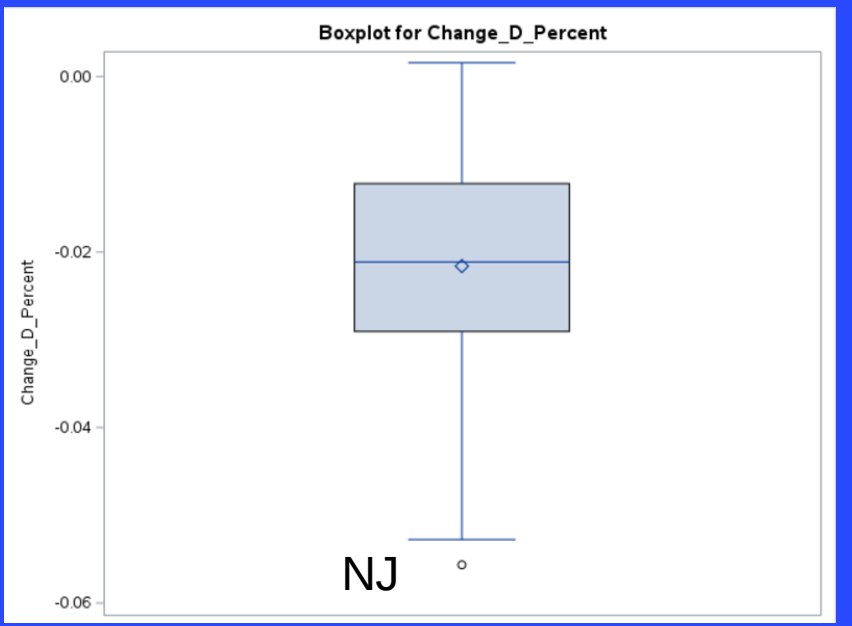




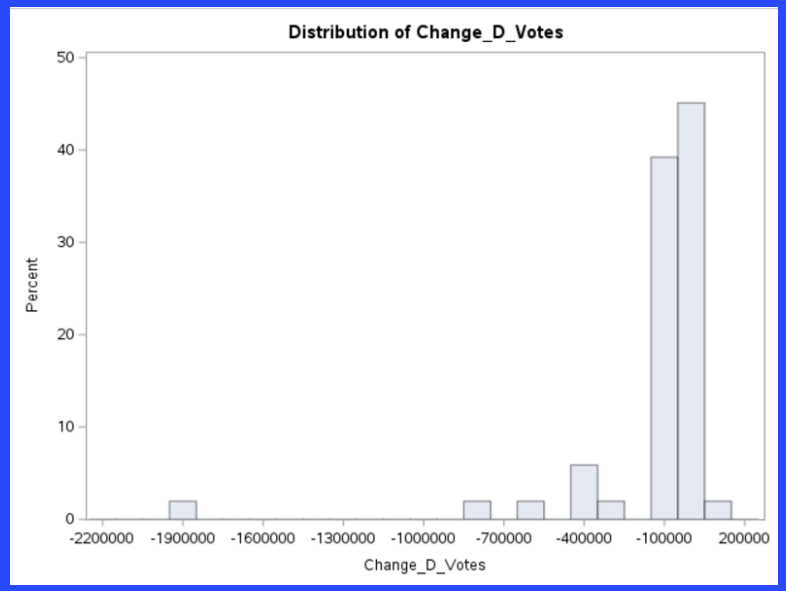


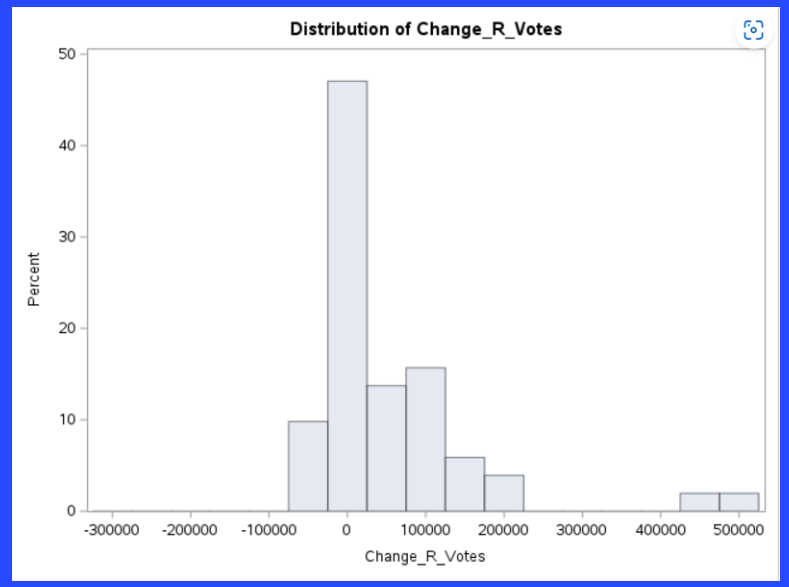


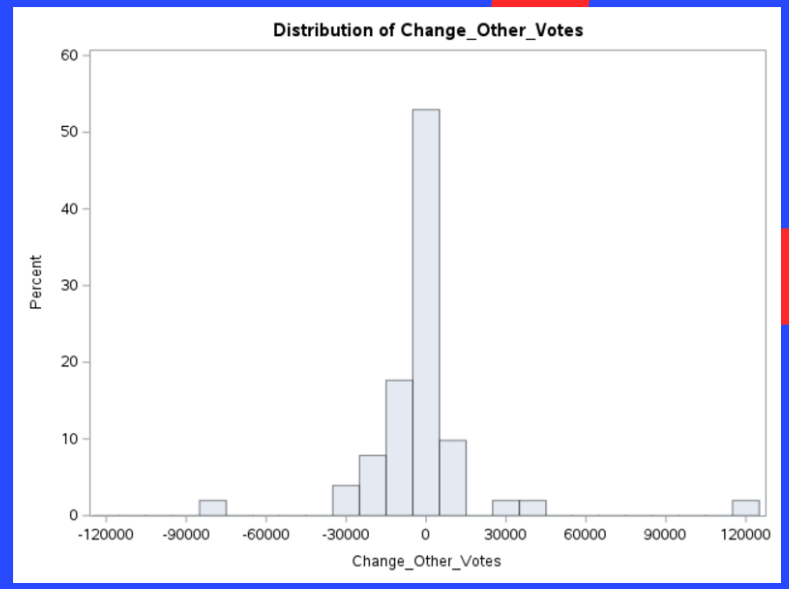


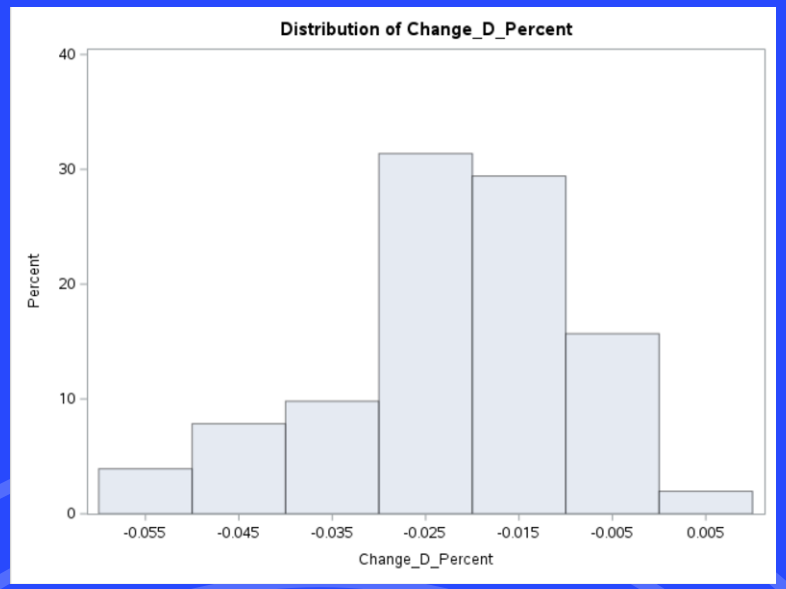


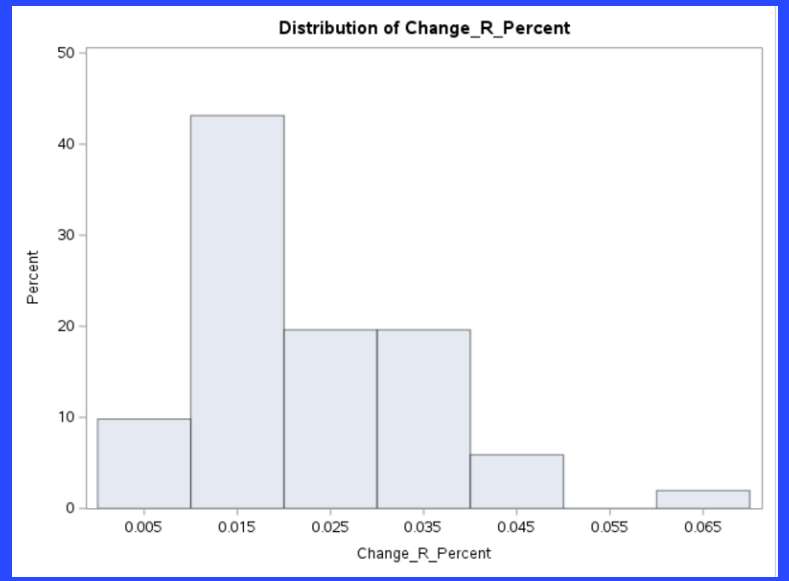
Distributions of the Change in Votes, and Percent Change by Party

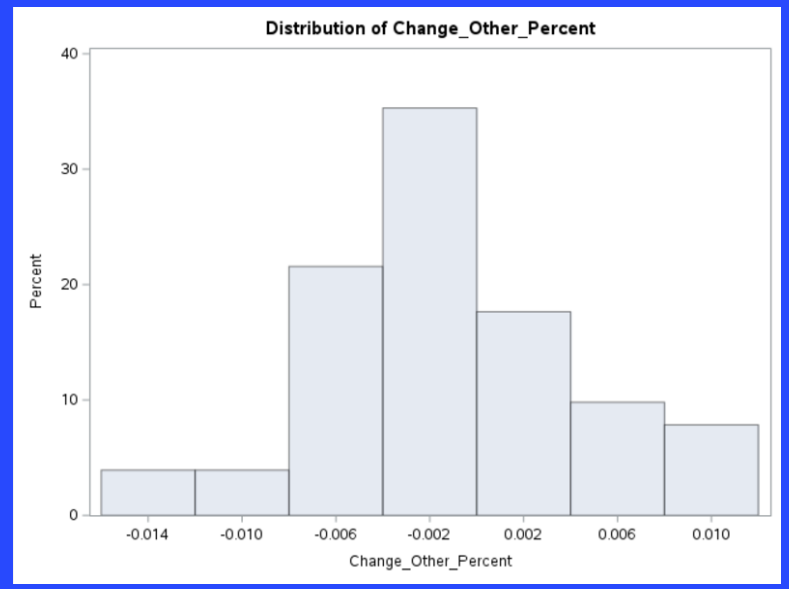












Paired t-test 7x!!!

Null Hypothesis: H0: There is no significant change in the mean number of Democratic votes between 2020 and 2024. ($\mu_1 = \mu_2$)

Alternative Hypothesis: Ha: There is a significant change in the mean number of Democratic votes between 2020 and 2024. ($\mu_1 \neq \mu_2$)

Null Hypothesis: H0: There is no significant change in the mean number of Republican votes between 2020 and 2024. ($\mu_1 = \mu_2$)

Alternative Hypothesis: Ha: There is a significant change in the mean number of Republican votes between 2020 and 2024. ($\mu_1 \neq \mu_2$)

Null Hypothesis: H0: There is no significant change in the mean number of Other votes between 2020 and 2024. ($\mu_1 = \mu_2$)

Alternative Hypothesis: Ha: There is a significant change in the mean number of Other votes between 2020 and 2024. ($\mu_1 \neq \mu_2$)

Null Hypothesis: H0: There is no significant change in the mean number of Total votes between 2020 and 2024. ($\mu_1 = \mu_2$)

Alternative Hypothesis: Ha: There is a significant change in the mean number of Total votes between 2020 and 2024. ($\mu_1 \neq \mu_2$)

Paired t-test 7x!!!

Null Hypothesis: H0: There is no significant change in the mean number of Democratic Percent Votes between 2020 and 2024. ($\mu_1 = \mu_2$)

Alternative Hypothesis: Ha: There is a significant change in the mean number of Democratic Percent Votes between 2020 and 2024. ($\mu_1 \neq \mu_2$)

Null Hypothesis: H0: There is no significant change in the mean number of Republican Percent Votes between 2020 and 2024. ($\mu_1 = \mu_2$)

Alternative Hypothesis: Ha: There is a significant change in the mean number of Republican Votes between 2020 and 2024. ($\mu_1 \neq \mu_2$)

Null Hypothesis: H0: There is no significant change in the mean number of Other Percent Votes between 2020 and 2024. ($\mu_1 = \mu_2$)

Alternative Hypothesis: Ha: There is a significant change in the mean number of Other Percent Votes between 2020 and 2024. ($\mu_1 \neq \mu_2$)

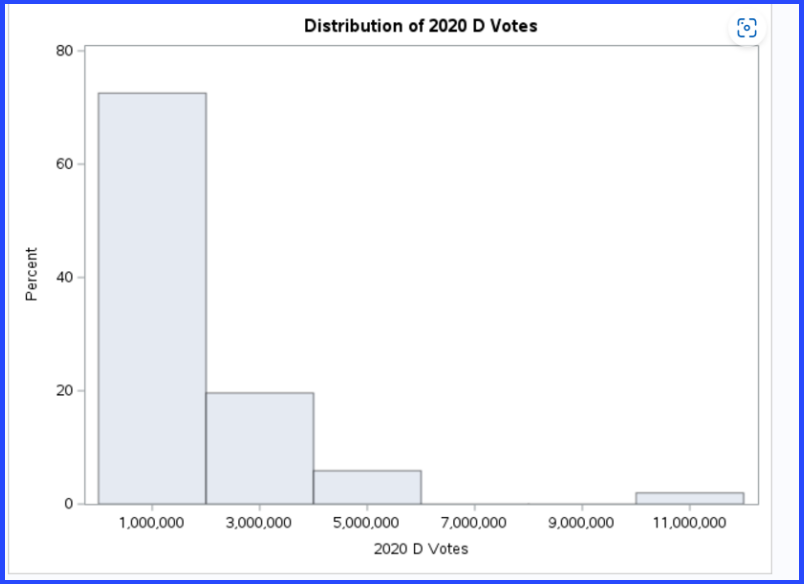
Check Assumptions

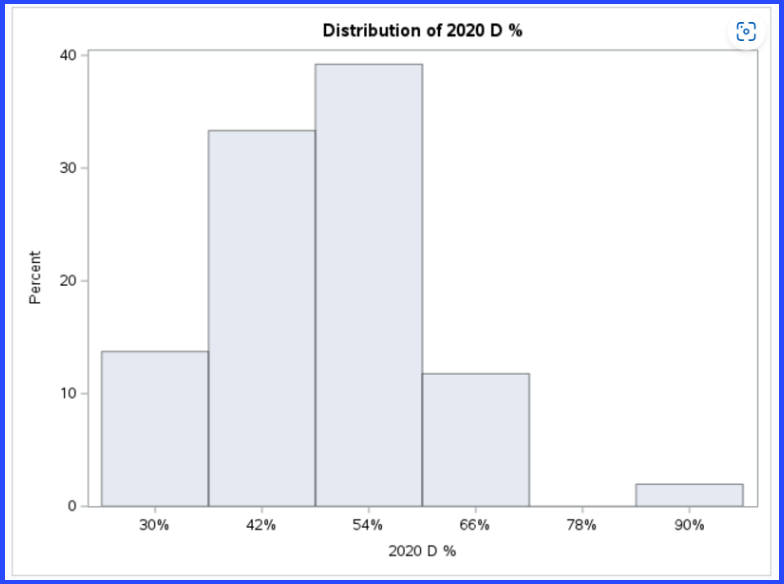
1. ***Randomness***: The sample can be classified as random.
2. ***Independence***: Each entry (row) is a state, and therefore is independent of all the other entries.
3. ***Normality***: We need to check the normality for each variable we will be testing. View histograms and use *proc univariate* looking for P-value < $\alpha = 0.05$.

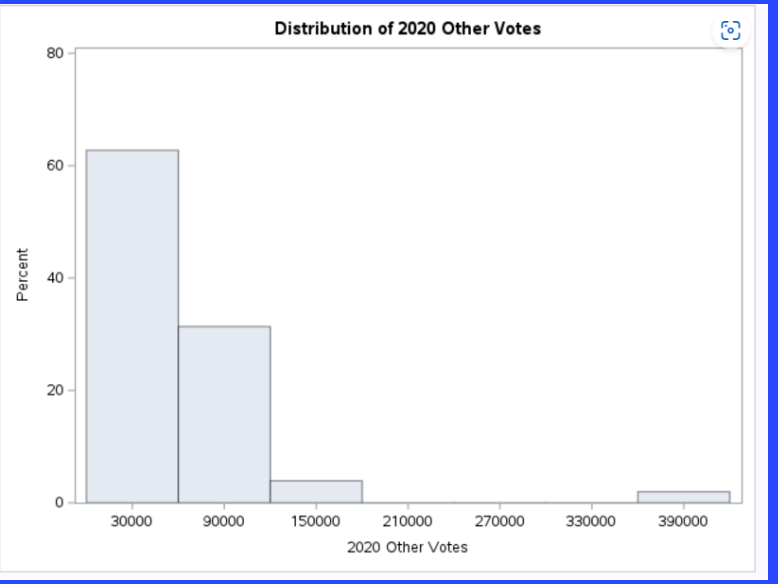
Normality Test	
Variable	*P-Test*
2020 D Votes	<0.0001
2020 D %	0.0457
2020 R Votes	<0.0001
2020 R %	0.0288
2020 Other Votes	<0.0001
2020 Other %	0.0049
2020 Total Votes	<0.0001
2024 D Votes	<0.0001
2024 D %	0.0152
2024 R Votes	<0.0001
2024 R %	0.0119
2024 Other Votes	<0.0001
2024 Other %	0.1458
2024 Total Votes	<0.0001

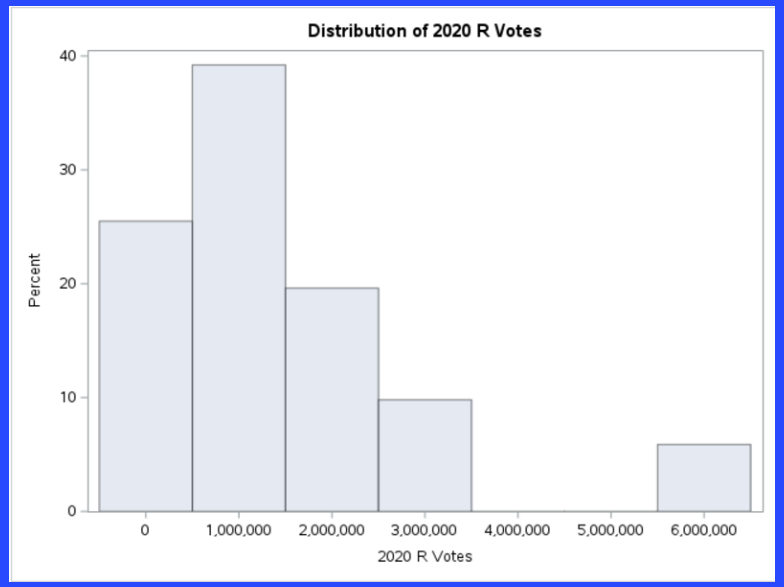
- All of our variables have a P-value of less than 0.05, except for the variable "*2024 Other %*".
- Histograms of the distributions of each variable are mostly normal.
- Only one variable is not normal, *"2024 Other %".*
- Perform our paired t-tests (7 times!)

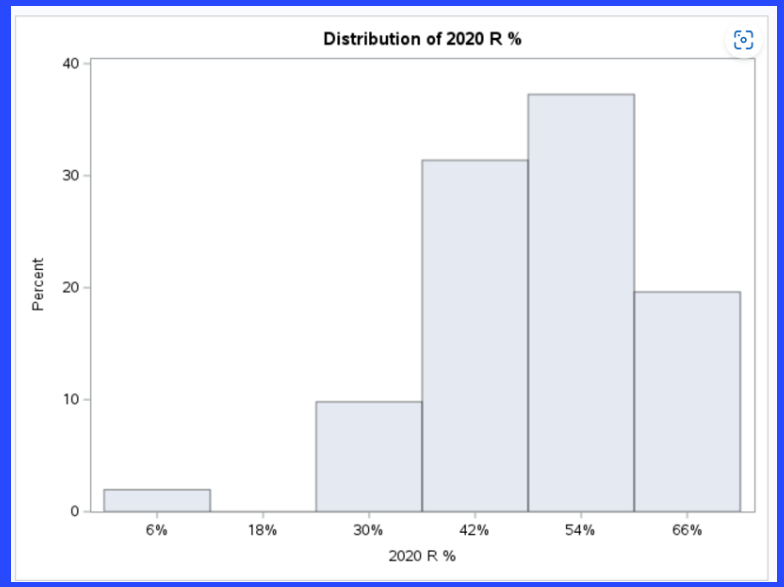
Histograms of Distributions 2020

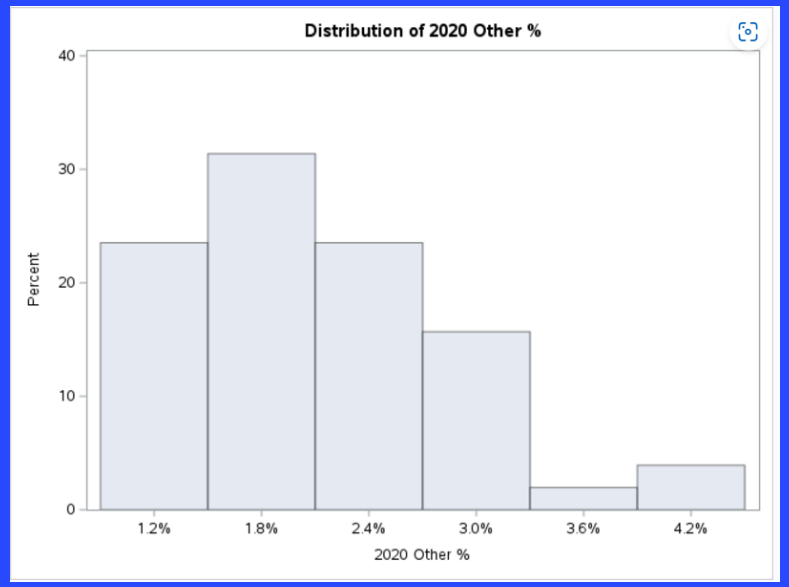




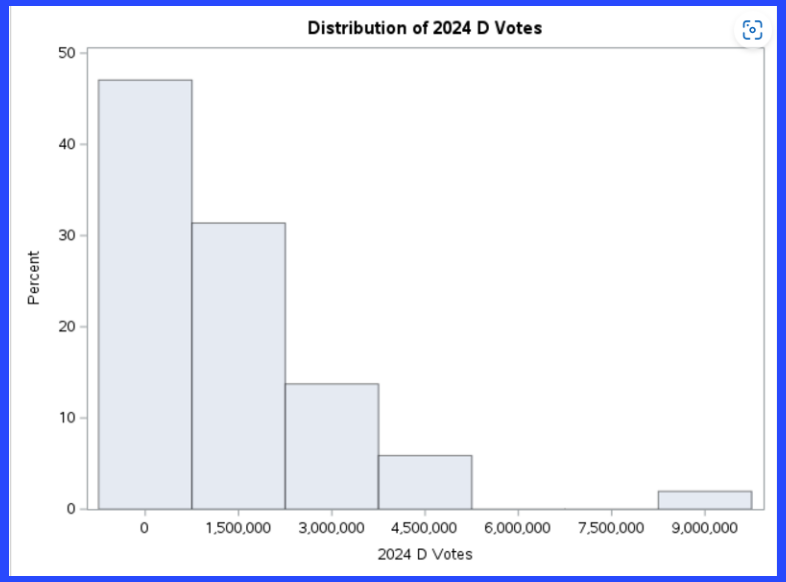


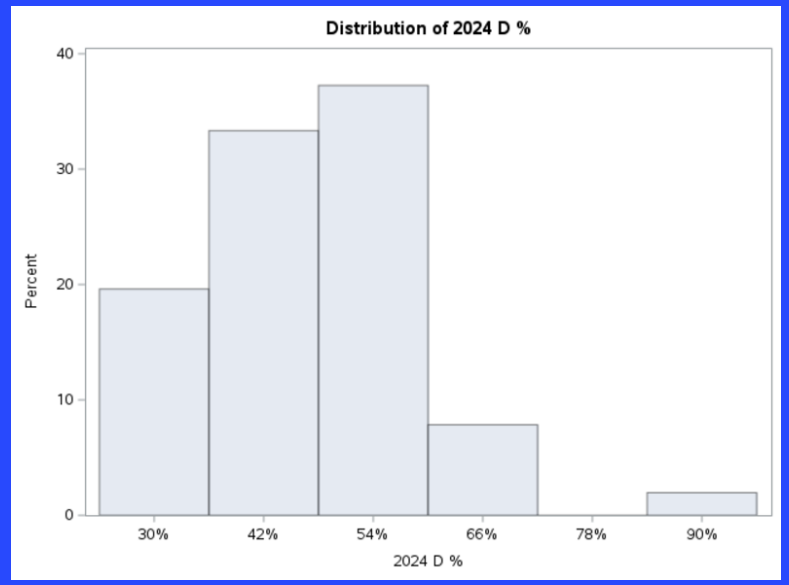


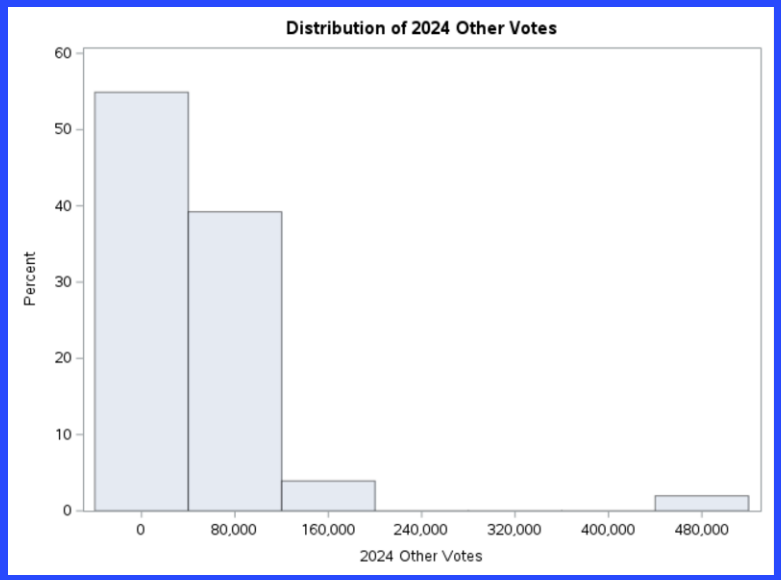


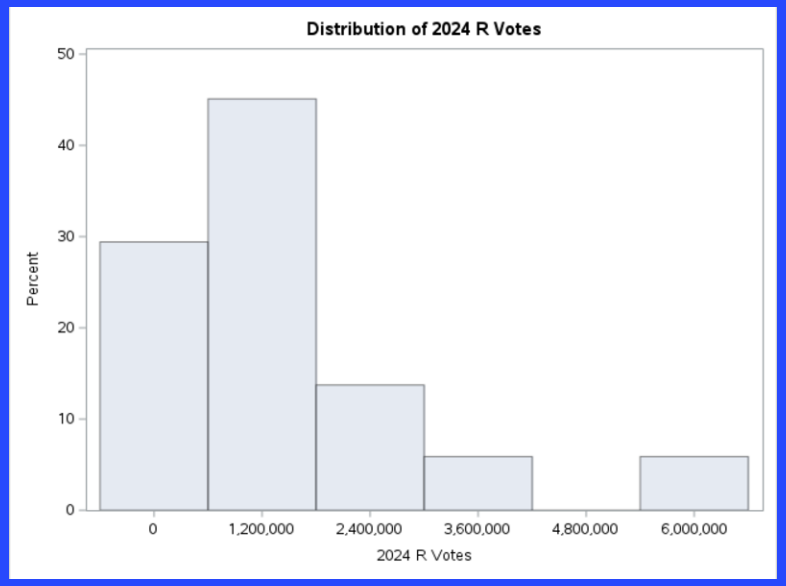


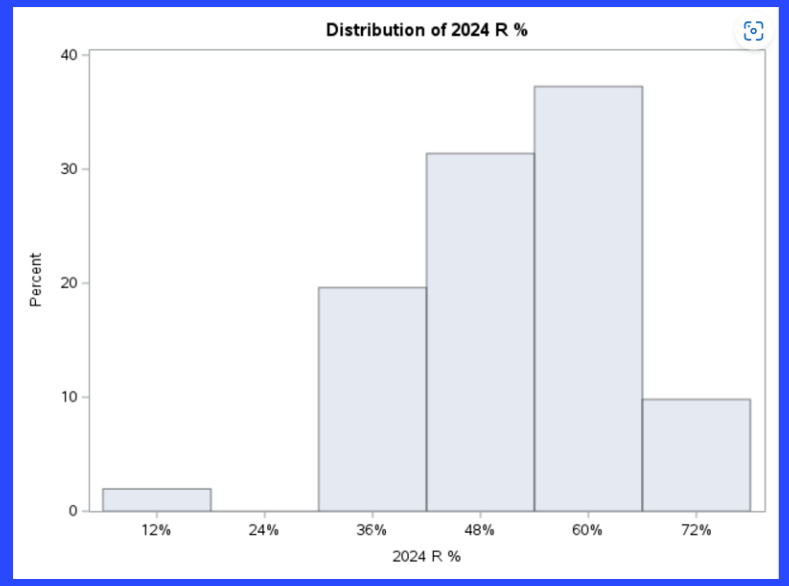
Histograms of Distributions 2024

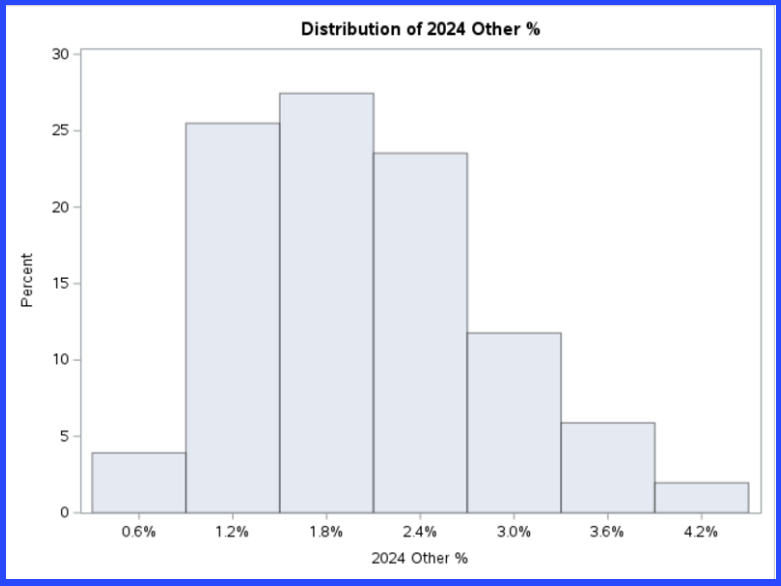












Paired T-tests results (Vote Count)

The TTEST Procedure

Difference: 2024 D Votes - 2020 D Votes

N	Mean	Std Dev	Std Err	Minimum	Maximum
51	-129746	297959	41722.5	-1856764	74385.0

Mean	95% CL Mean		Std Dev	95% CL Std Dev	
-129746	-213548	-45944.1	297959	249305	370386

DF	t Value	Pr > |t|
50	-3.11	0.0031

The TTEST Procedure

Difference: 2024 R Votes - 2020 R Votes

N	Mean	Std Dev	Std Err	Minimum	Maximum
51	56394.6	104843	14680.9	-53728.0	502689

Mean	95% CL Mean		Std Dev	95% CL Std Dev	
56394.6	26907.0	85882.1	104843	87723.0	130328

DF	t Value	Pr > |t|
50	3.84	0.0003

The TTEST Procedure

Difference: 2024 Other Votes - 2020 Other Votes

N	Mean	Std Dev	Std Err	Minimum	Maximum
51	-2145.9	23600.7	3304.8	-79883.0	119463

Mean	95% CL Mean		Std Dev	95% CL Std Dev	
-2145.9	-8783.7	4491.9	23600.7	19746.9	29337.5

DF	t Value	Pr > |t|
50	-0.65	0.5191

The TTEST Procedure

Difference: 2024 Total Votes - 2020 Total Votes

N	Mean	Std Dev	Std Err	Minimum	Maximum
51	-75497.7	271991	38086.4	-1683262	251177

Mean	95% CL Mean		Std Dev	95% CL Std Dev	
-75497.7	-151996	1001.0	271991	227577	338106

DF	t Value	Pr > |t|
50	-1.98	0.0530

Paired T-tests results (Percent%)

The TTEST Procedure

Difference: 2024 D % - 2020 D %

N	Mean	Std Dev	Std Err	Minimum	Maximum
51	-0.0216	0.0135	0.00189	-0.0556	0.00158

Mean	95% CL Mean		Std Dev	95% CL Std Dev	
-0.0216	-0.0254	-0.0178	0.0135	0.0113	0.0168

DF	t Value	Pr > |t|
50	-11.44	<.0001

The TTEST Procedure

Difference: 2024 Other % - 2020 Other %

N	Mean	Std Dev	Std Err	Minimum	Maximum
51	-0.00107	0.00529	0.000741	-0.0143	0.0106

Mean	95% CL Mean		Std Dev	95% CL Std Dev	
-0.00107	-0.00256	0.000416	0.00529	0.00443	0.00658

DF	t Value	Pr > |t|
50	-1.45	0.1542

The TTEST Procedure

Difference: 2024 R % - 2020 R %

N	Mean	Std Dev	Std Err	Minimum	Maximum
51	0.0227	0.0119	0.00166	0.00495	0.0616

Mean	95% CL Mean		Std Dev	95% CL Std Dev	
0.0227	0.0193	0.0260	0.0119	0.00994	0.0148

DF	t Value	Pr > |t|
50	13.64	<.0001

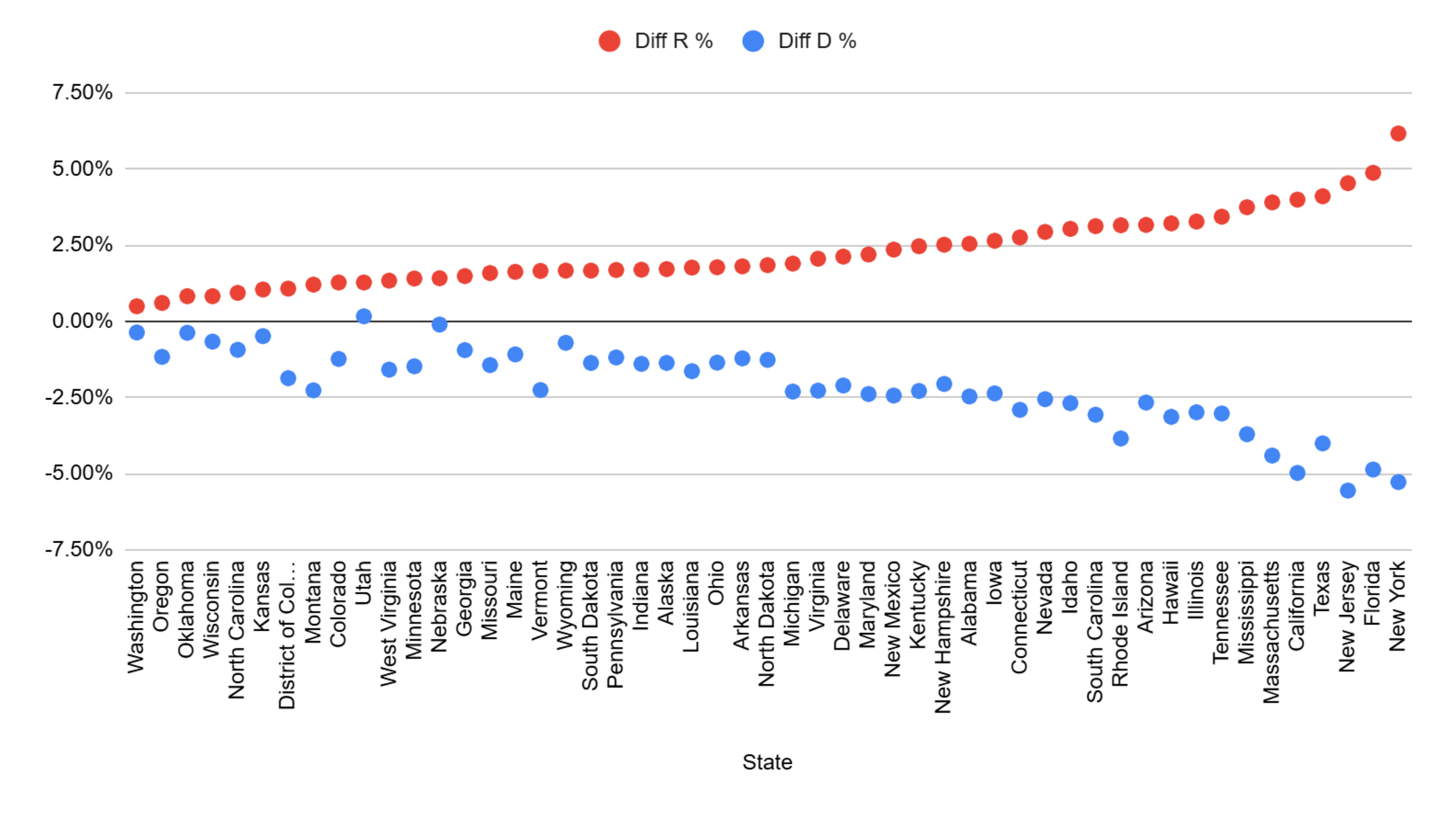
	Diff D	*Diff D %*	*Diff R*	*Diff R %*	*Diff Other*	*Diff Other %*	*Diff Total*
Mean	-129,746	-2.16%	56,394	2.27%	-2,145	-0.11%	-75,498
95% Min	-213,548	-2.54%	26,907	1.93%	-8,784	-0.26%	-151,996
95% Max	-45,944	-1.78%	85,882	2.60%	4,492	0.04%	1,001
P-Value	0.0031	0.0001	0.0003	0.0001	0.5191	0.1542	0.053
Null Hyp	Reject	Reject	Reject	Reject	Fail To Reject	Fail To Reject	Fail to Reject, close to cutoff
Alt Hyp	Evidence For	Evidence For	Evidence For	Evidence For			

Interesting outliers

	Diff D	Diff D %	Diff R	Diff R %	Diff Other	Diff Other %	Diff Total
Total	-6,617,062		2,876,124		-109,443		-3,850,381
Average	-129,746	-2.16%	56,395	2.27%	-2,146	-0.11%	-75,498

State	Diff D	State	Diff D %	State	Diff R	State	Diff R %
California	-1,856,764	New Jersey	-5.56%	Washington	-53,728	Washington	0.49%
New York	-847,000	New York	-5.28%	Louisiana	-47,271	Oregon	0.60%
Florida	-614,007	California	-4.98%	Oregon	-43,703	Oklahoma	0.82%
Illinois	-433,190	Florida	-4.87%	Mississippi	-40,714	Wisconsin	0.82%
Texas	-424,518	Massachusetts	-4.41%	Kansas	-29,942	North Carolina	0.93%
Only states that saw increase							
Nevada	1,711	Kansas	-0.49%	Michigan	166,784	California	3.99%
Utah	2,284	Oklahoma	-0.38%	Georgia	201,263	Texas	4.10%
North Carolina	31,090	Washington	-0.37%	New York	220,801	New Jersey	4.53%
Wisconsin	37,363	Nebraska	-0.11%	Florida	441,394	Florida	4.87%
Georgia	74,385	Utah	0.16%	Texas	502,689	New York	6.16%

State	Diff Total
California	-1,683,262
New York	-706,082
Illinois	-463,221
New Jersey	-275,414
Washington	-188,638
Pennsylvania	98,320
Wisconsin	122,125
Michigan	124,884
North Carolina	173,795
Georgia	251,177


Diff R %
Diff D %
7.50%
5.00%
2.50%
0.00%
-2.50%
-5.00%
-7.50%
Washington
Oregon
Oklahoma
Wisconsin
North Carolina
Kansas
District of Col…
Montana
Colorado
Utah
West Virginia
Minnesota
Nebraska
Georgia
Missouri
Maine
Vermont
Wyoming
South Dakota
Pennsylvania
Indiana
Alaska
Louisiana
Ohio
Arkansas
North Dakota
Michigan
Virginia
Delaware
Maryland
New Mexico
Kentucky
New Hampshire
Alabama
Iowa
Connecticut
Nevada
Idaho
South Carolina
Rhode Island
Arizona
Hawaii
Illinois
Tennessee
Mississippi
Massachusetts
California
Texas
New Jersey
Florida
New York
State

Conclusions

- The means for Republican votes and Republican % went up an average 56,395 and 2.27%. The means are not equal.
- Democrat votes went down by an average of 129,746 and 2.16%. The means are not equal.
- Other votes did not see a significant change. The means are close enough to be considered equal.
- Total votes went down by 3,850,381. The means are close enough to be considered equal.
- Republicans received more votes in most states and Democrats received less votes in most states. The Total Vote count went down, but it still stayed relatively the same for each state.
- All states had an increase in % Republican vote
- All, except 1 state (Utah), had a decrease in % Democrat Vote.

Thank you

References:

https://www.nbcnews.com/politics/2024-elections/president-results

https://www.nbcnews.com/politics/2020-elections/president-results/