

Pennsylvania County Voting Anomaly Analysis

S. Stanley Young, PhD, FASA, FAAAS

This report looks at Pennsylvania county voting, 2008 to 2020. The data set has 67 row, with one row for every county. The first few rows are given here.

RowID	PA Counties	Obama 2008	Obama 2012	Clinton 2016	Biden 2020	Dif	Rank	Democratic	Republican	Other	Total	%Dem
1	Adams	17633	15091	14219	17919	2,271.3	20	19090	36862	10824	66776	28.6
2	Allegheny	373153	352687	367617	396767	32,281.3	3	555649	263952	138999	958600	58.0
3	Armstrong	11138	9045	7178	7130	-1,990.3	60	13903	22848	4848	41599	33.4
4	Beaver	40499	37055	32531	37389	694.0	24	53417	41701	14063	109181	48.9

This report is in the form of text describing a question of interest, a figure and discussion of the figure.

Item 1 —

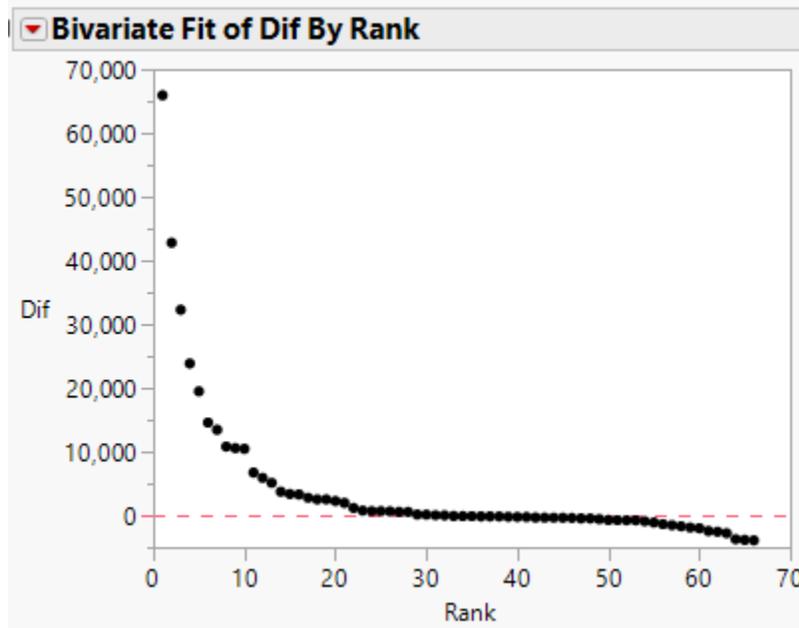
In the data set examined, we note that Philadelphia is ~80,000 votes below the average of the previous three presidential elections. It is an outlier and we eliminate that county from most of the analyses that follow.

RowID	PA Counties	Obama 2008	Obama 2012	Clinton 2016	Biden 2020	Dif
65	Crawford	16780	13883	10971	10052	-3,826.0
66	Cambria	32451	24249	18867	21293	-3,896.0
67	Philadelphia	595980	588806	584025	509491	-80,112.7

Note that Crawford and Cambria had about 3,800 fewer votes. Philadelphia is a clear outlier low.

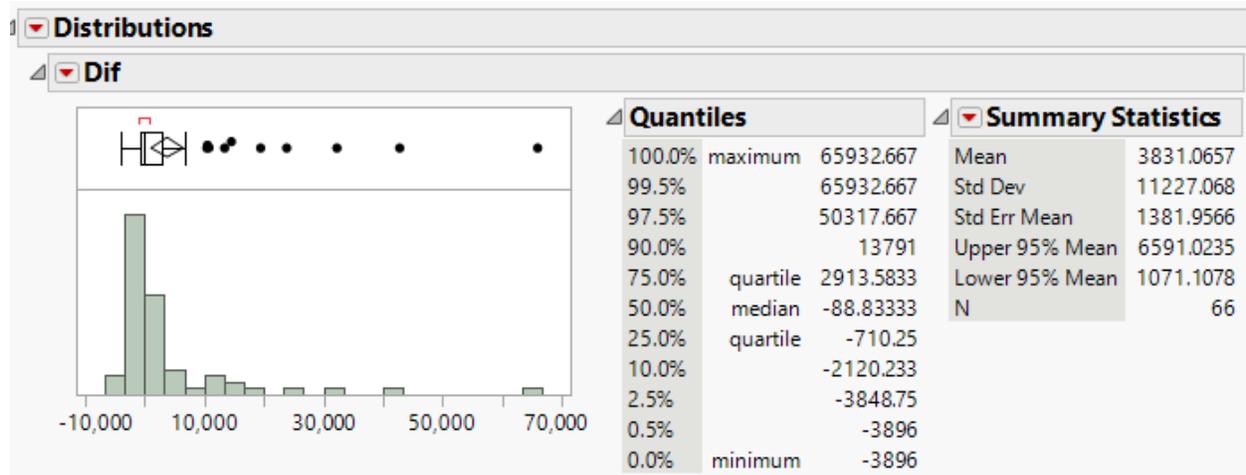
Item 2 —

Are there particular counties where voting does **not** follow previous voting patterns? We computed the difference in Biden 2020 versus the average of Obama 2008, Obama 2012 and Hillary 2016. The name of this variable is Dif. We plot Dif versus the rank of the difference. The largest Dif gets rank 1; the 2nd largest Dif gets rank 2, etc.



We see that for most counties there is very little difference in this election relative to previous elections. In some counties the Biden count is lower than previous Democratic counts. The same data can be displayed as a histogram.

Item 3 —



On the left of the figure, from -5,000 to about 10,000 we see bars that resemble a normal distribution. The values above 10,000 appear to be outliers. It is unusual to see a gain of 10,000 votes or more.

Item 4 —

Here are the outlier counties:

PA Counties	Obama 2008	Obama 2012	Clinton 2016	Biden 2020	Dif
Montgomery	253393	233356	256082	313543	65,932.7
Chester	137833	124311	141682	177408	42,799.3
Allegheny	373153	352687	367617	396767	32,281.3
Bucks	179031	160521	167060	192719	23,848.3
Lancaster	99586	88481	91093	112536	19,482.7
Cumberland	48306	44367	47085	61168	14,582.0
Northampt...	75255	67606	66272	83163	13,452.0
Delaware	178870	171792	177402	186832	10,810.7
Dauphin	69975	64965	64706	77125	10,576.3
York	82839	73191	68524	85323	10,471.7

Consider Montgomery County. Obama/Hillary vote counts ranged from 233,000 to 256,000. Biden received 313,000. The ten outlier counties together provide 244,237 excess votes.

Item 5 —

We examine the Biden votes against Democratic registered voters.

	County	Biden 2020	Registered D	Dif2
1	Montgomery	313543	274955	38588
2	Chester	177408	142423	34985
3	Allegheny	396767	555649	-158882
4	Bucks	192719	195772	-3053
5	Lancaster	112536	106762	5774
6	Cumberland	61168	59656	1512
7	Northampton	83163	95710	-12547
8	Delaware	186832	190702	-3870
9	Dauphin	77125	83635	-6510
10	York	85323	95027	-9704

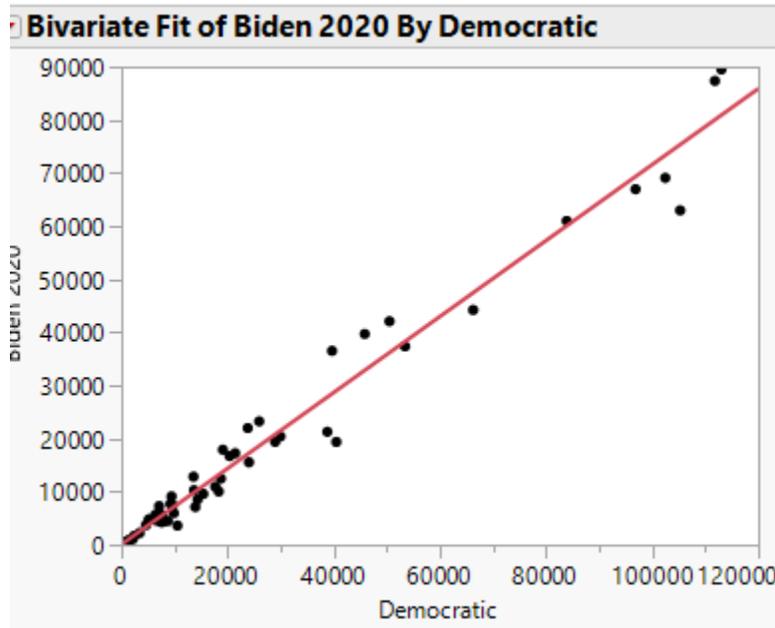
Note: in Montgomery and Chester counties there are more Biden voters than registered Democrats...

Item 6 —

Allegheny is anomalous. There are 158,000 fewer Biden votes than registered voters. So, in this data set Allegheny and Philadelphia counties are short 158,000 and 509,000 voters. Curious.

Item 7 —

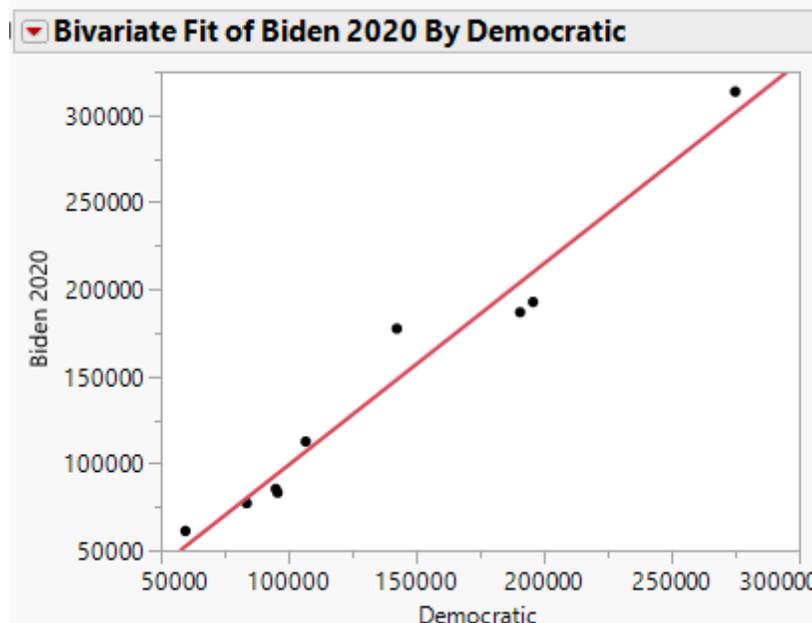
We seek to estimate the fraction of registered Democratic voters that voted for Biden. We want an unbiased estimate so the 10 outlier counties and Philadelphia were removed from the analysis. 55 counties were used for simple linear regression.



The data are fit well with a simple line. $Biden\ 2020 = 111.10562 + 0.7150147 * Democratic$
We expect 70%± of registered Democratic voters to vote for Biden.

Item 8 —

We seek to estimate the fraction of registered Democratic voters that voted for Biden among the 10 outlier counties. We remove Allegheny County as so many voters appear not to have voted. We want an unbiased estimate. 9 counties were used for simple linear regression.



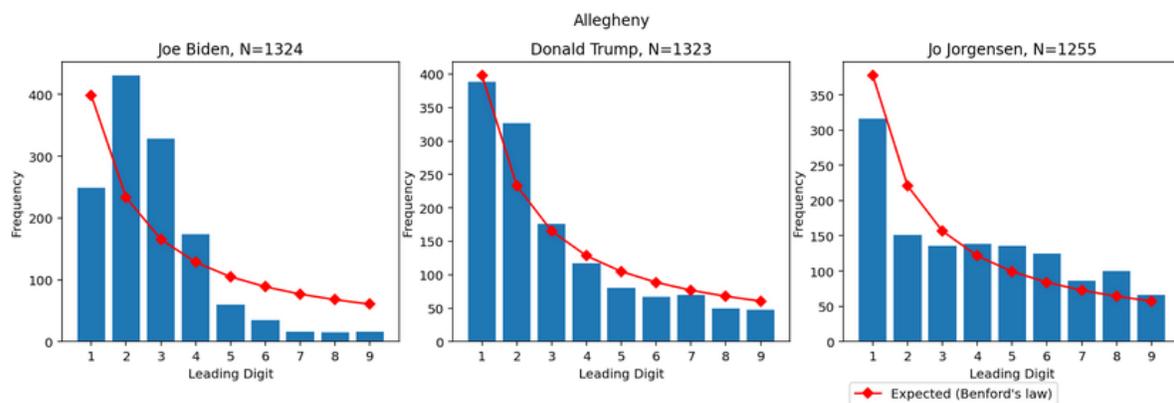
The data (on the prior page graph) fit well with a simple line. Biden 2020 = -16654.26 + 1.1567224*Democratic

The estimate is that 115% of registered Democratic voters voted for Biden. We have a contradiction. Did many Trump voters cross over and all registered Democratic voters vote for Biden? $115 - 71 = 44$. So, 44% excess voters voted for Biden in the 10-1=9 outlier counties. That is not reasonable. An alternative explanation is that excess votes were added to the Biden total that did not come from voters.

Item 9 —

Allegheny County is quite odd. One technique to detect fraud from numerical data is to use the so called Benford's Law*. That law asserts that the number 1 should appear more frequently as the lead digit in a precinct voter count. The number 2 should be less frequent, etc. The distribution, with no fraud, should decline exponentially. That method has been applied to Allegheny County's voting precincts. See ciph8914/2020_benfords.

Allegheny, PA



The shape of the voter count first digits follows Benford's Law rather closely for Trump, but does not for Biden. This result lends support for the county analysis that Allegheny County is anomalous.

SUMMARY:

1. Both Philadelphia and Allegheny Counties report suspiciously low numbers of votes in this data set. The question is: are these the final actual totals of votes? If it is something *earlier* than final, that might indicate that the actual numbers of voters was relatively low, and Dems then used prior turnout numbers to hide the fact that they were manufacturing votes.
2. The votes for Biden are unusually high for ten counties, reporting an excess of ~244,000 votes in excess of expectation. These deviations are legitimate reasons to insist on closely monitored recounts.
3. Normally 70%± of PA registered voters vote in a national election. After elimination of Allegheny County, 115% of PA registered Democratic voters voted for Biden. That is a serious statistical aberration.
4. Another way to track down fraudulent votes is to look closely at how many of the votes did little or no down-ticket voting. When manufacturing votes, it is too time consuming to vote for other office holders.
5. A standard voter fraud method, *Benford's Law*, flags Allegheny County as suspect.

*"[The Theory and Applications of Benford's Law](#)" (Steven J. Miller, editor)