

## A2. Exploratory Data Analysis on 2000 Presidential Election in FL

Tiffany Jaya

W209

[Hypothesis 1](#)

[Conclusion](#)

[Hypothesis 2](#)

[Conclusion](#)

[Hypothesis 3](#)

[#heading=h.9ak0eudt5dee](#)

### Hypothesis 1

We are under the impression that the “butterfly ballot” used in Palm Beach County during the 2000 Presidential Election has misled many voters into voting for Buchanan when they in fact intended to vote for Al Gore.

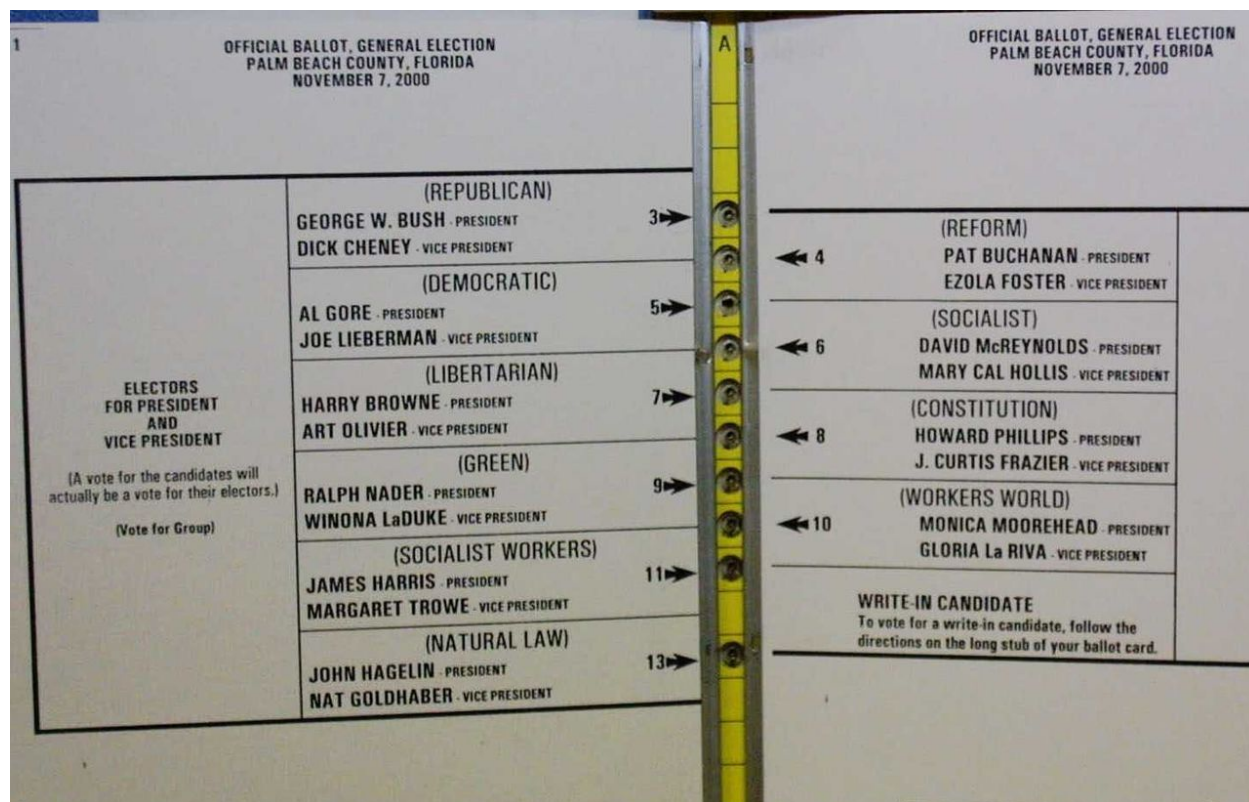


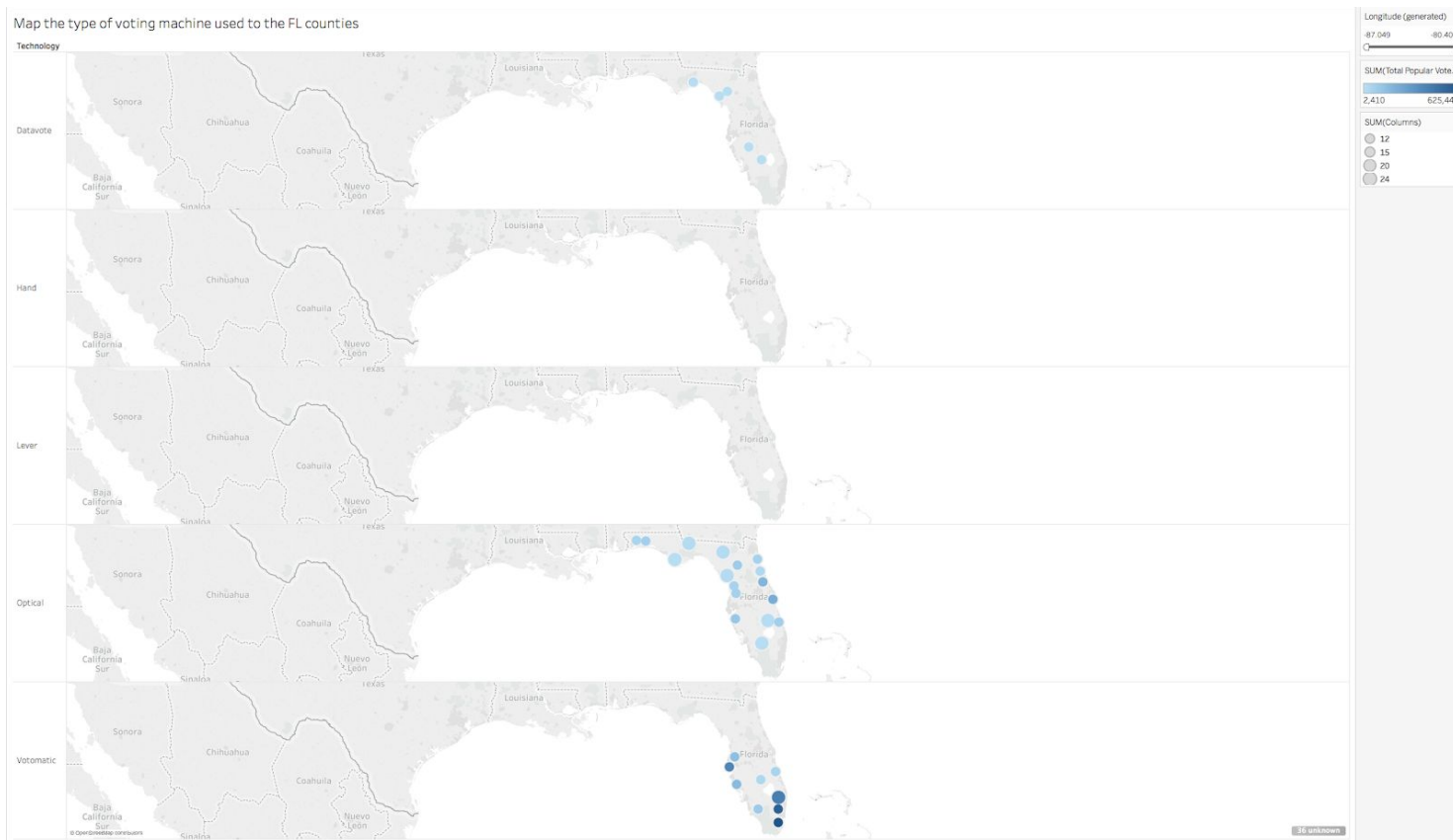
Figure 1. The “butterfly ballot” used in Palm Beach County during the November 7th 2000 Presidential Election

Type of voting machine and number of columns per county



**What's informative about this view:** Although this view did not explicitly state that Palm Beach County used a butterfly style ballot, it did verify that the ballot being used spreads the presidential candidates over two columns or two pages of the ballot. We can later use this information to compare similar counties with ballots that spread across two columns/pages and see if they are encountering the same issue found in Palm Beach County. This view also informs us that the type of voting machine being used is Votomatic.

**What could be improved about this view:** We do not have a feel of the geographic location where the types of voting machines are often used and the total vote counts of each county.

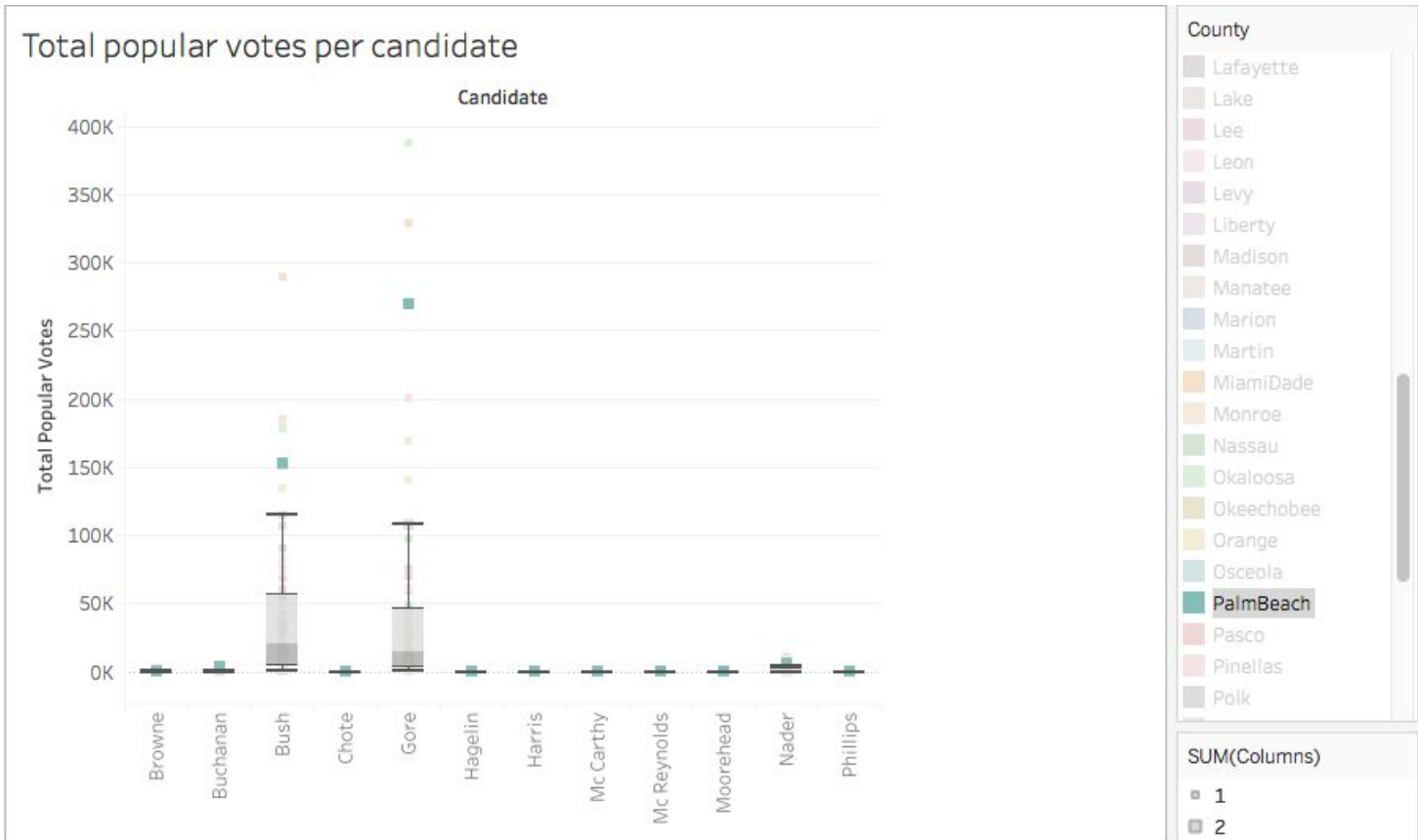


**What's informative about this view:** This view incorporates two elements that the previous view lacks: geographic location of each voting machines and the total number of popular votes for each county. The size of the circle represents the number of columns/pages in the ballot (1 or 2), and the color represents the total number of popular votes (the darker it is, the higher the number of votes are).

From this view, we can surmise that Optical and Votomatic are the most popular types of voting machines used in Florida. Optical is used in the northern and middle region of Florida where there are fewer voters or fewer voters who turn out to vote. Votomatic, on the other hand, is popularly used in south Florida where there are more voters per county in general.

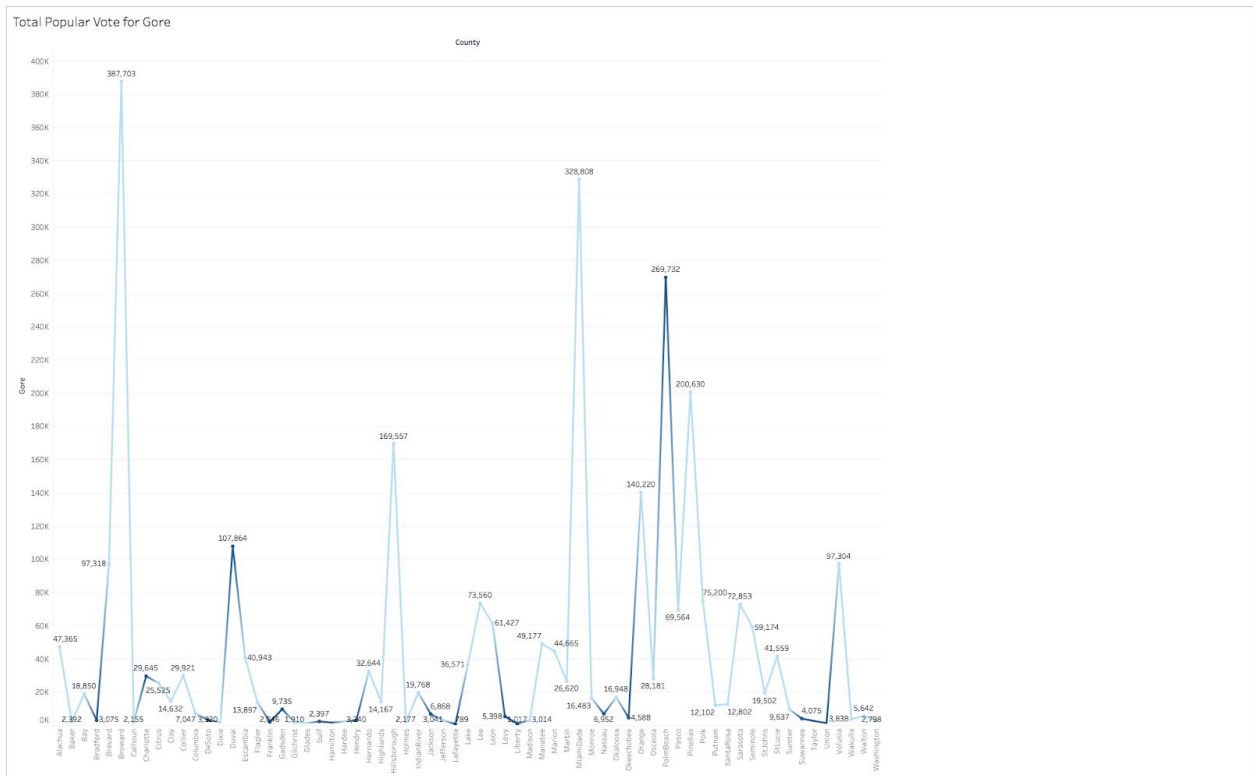
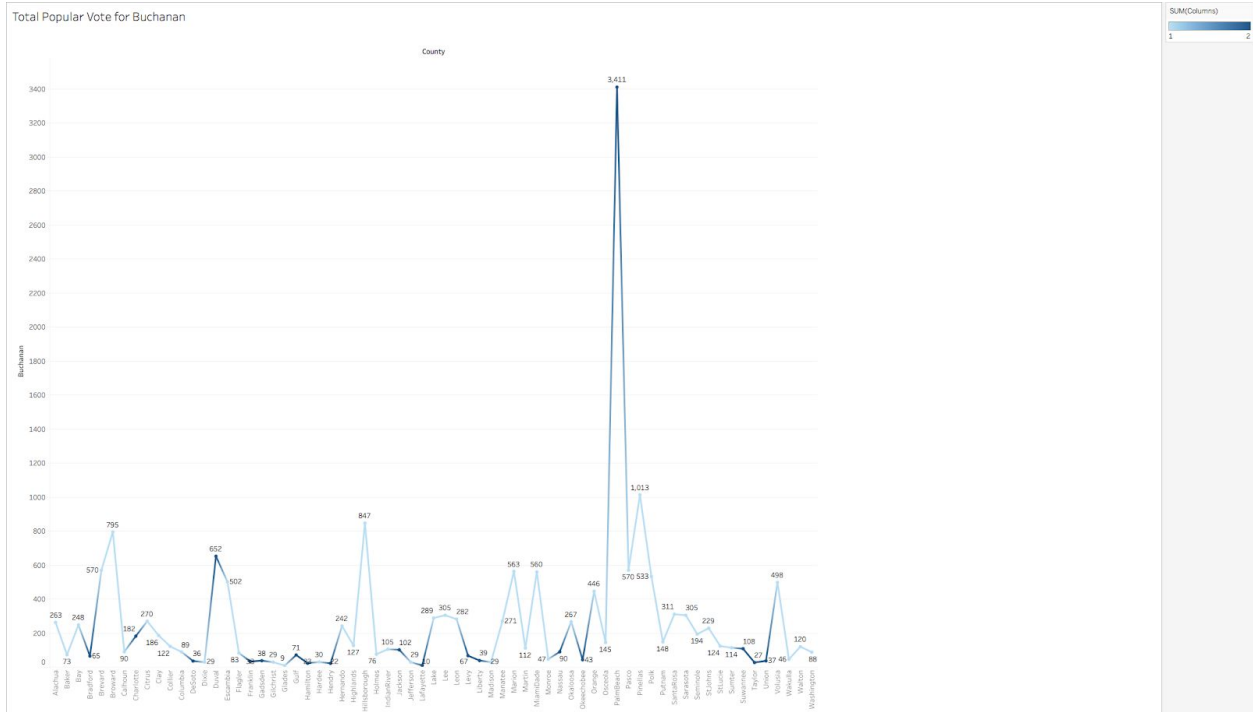
What is interesting to note is that the use of one or two columns/pages in a ballot is not limited by region.

**What could be improved about this view:** To make the view more condensed, we might not want to include the voting types Hand or Lever since it did not contribute much to the view. In addition, circle might not be a good design choice since it is challenging to differentiate the number of columns/pages used, especially if in the future counties decide to use more than two columns/pages in their ballots.



**What's informative about this view:** From this view, we notice that the butterfly ballot used in Palm Beach County are the only two columns/pages ballot with outliers among candidates Buchanan, Bush, Gore and Naber.

**What could be improved about this view:** Whether the outlier is caused by a large voter turnout or a larger population size or simply the butterfly ballot, we cannot make this conclusion. But what we can improve on is to place Bush and Gore in a separate chart so that we can better see the outliers in candidates with fewer votes than Bush and Gore.



**What's informative about this view:** It is clear from this view that Buchanan received much more votes in Palm Beach County than in any other counties. Since the total number of votes

he received from each county is smaller than Gore's, the difference is apparent. It is not as apparent in Gore who receives more votes in general.

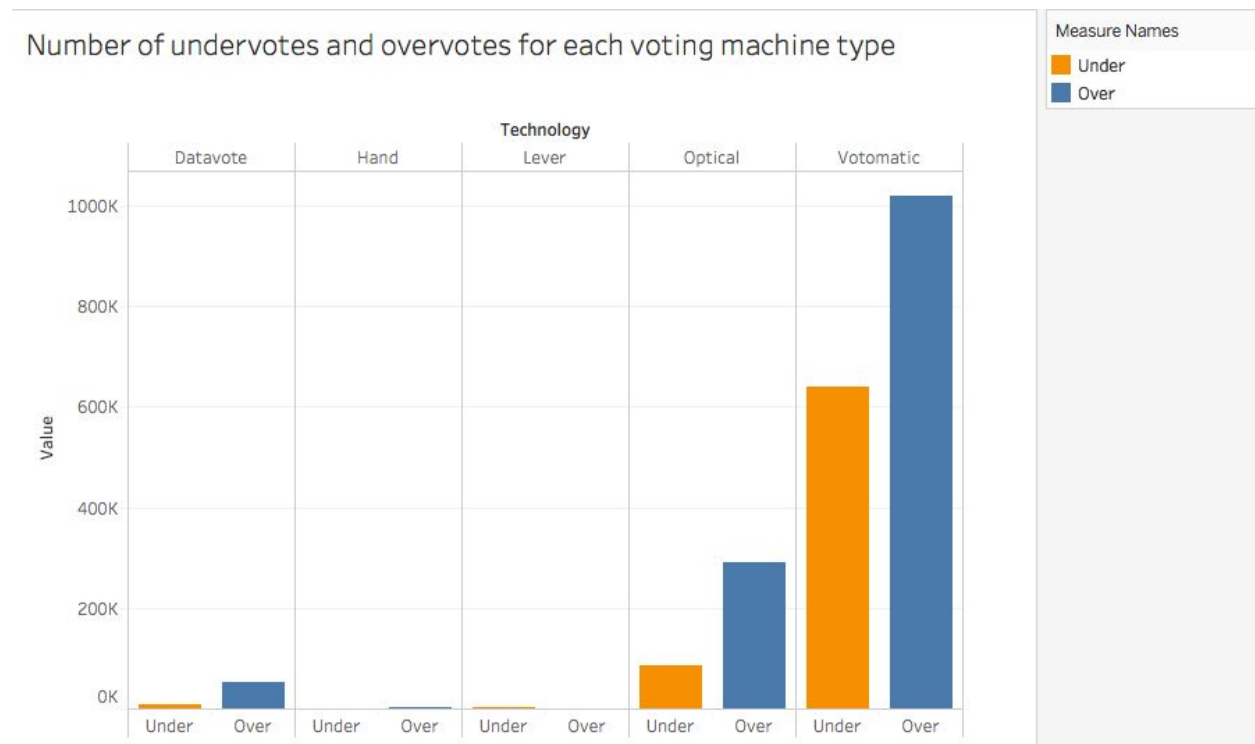
**What could be improved about this view:** It would have been nice to superimpose the total number of votes for Buchanan and Gore for comparison.

## Conclusion

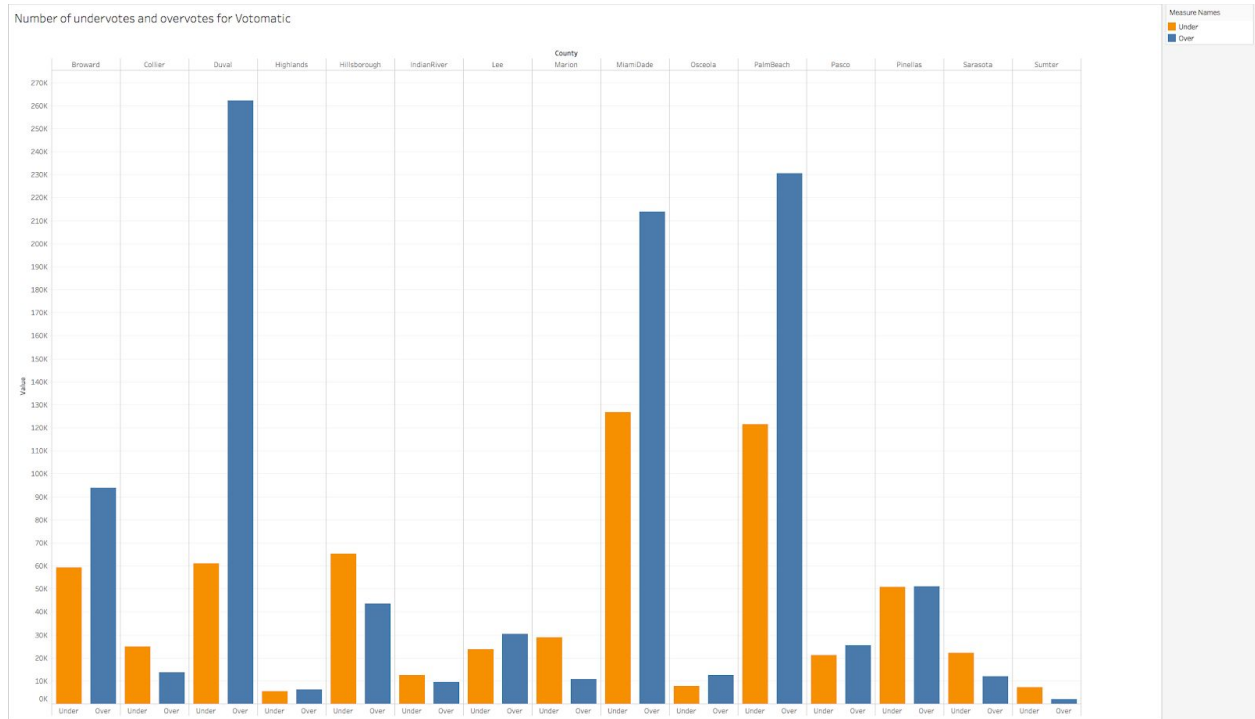
These data views suggest that the number of total votes for Buchanan in Palm Beach County is an outlier, and the “butterfly ballot” may play a role in misleading many voters into voting for Buchanan when they in fact intended to vote for Al Gore.

## Hypothesis 2

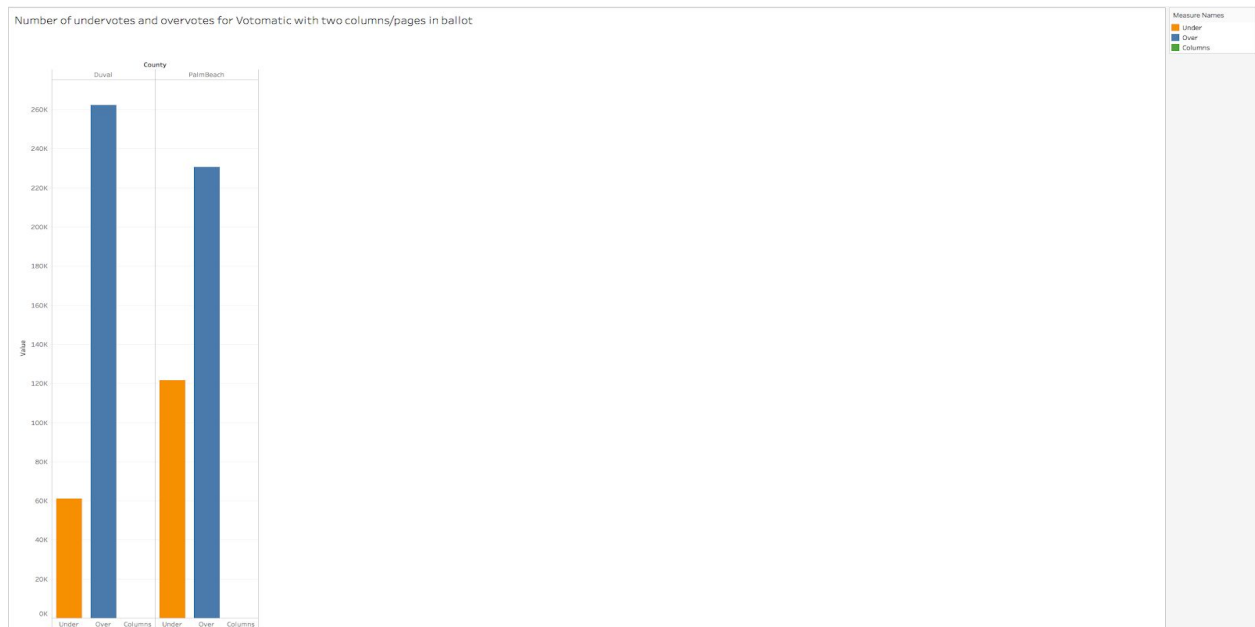
As a result of the “butterfly ballot” used in Palm Beach County, people tend to overvote than undervote during the election.



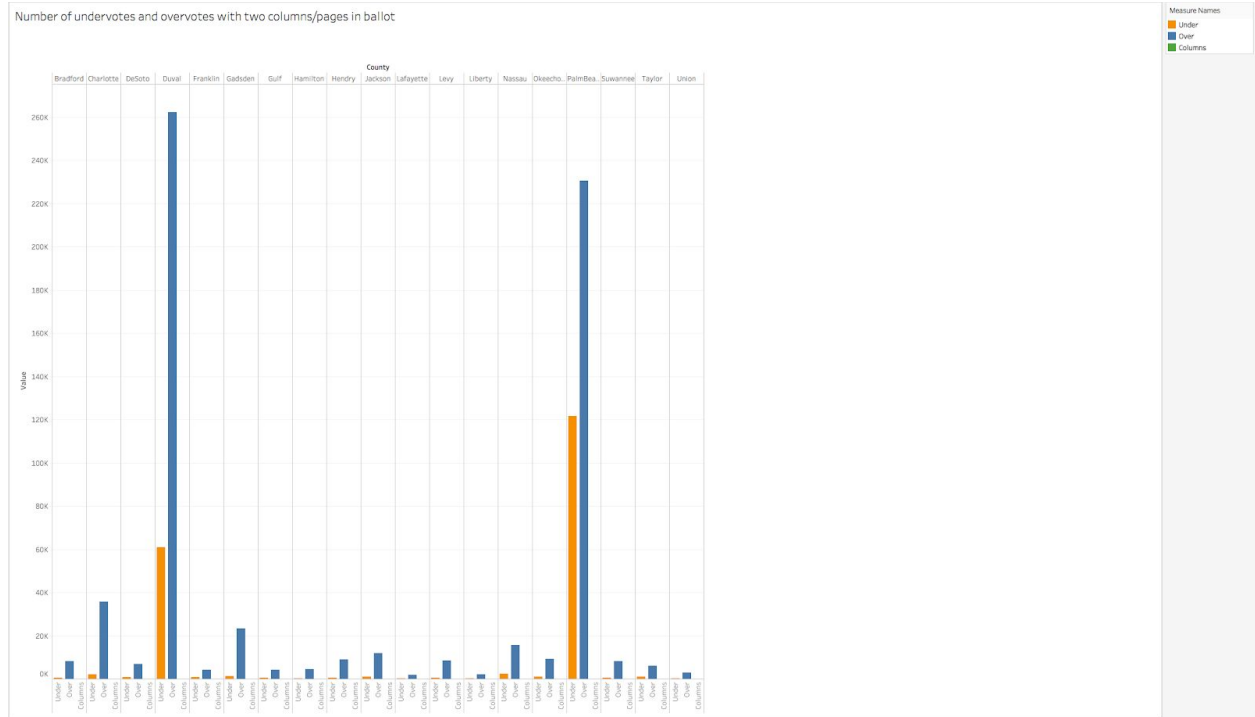
Among the popular voting machines, voters tend to overvote, meaning that voters tend to vote more than one candidate. Compare to overvote, there are fewer undervote incidents, meaning there are fewer instances when a ballot is not counted because of unclear marking by the voter.



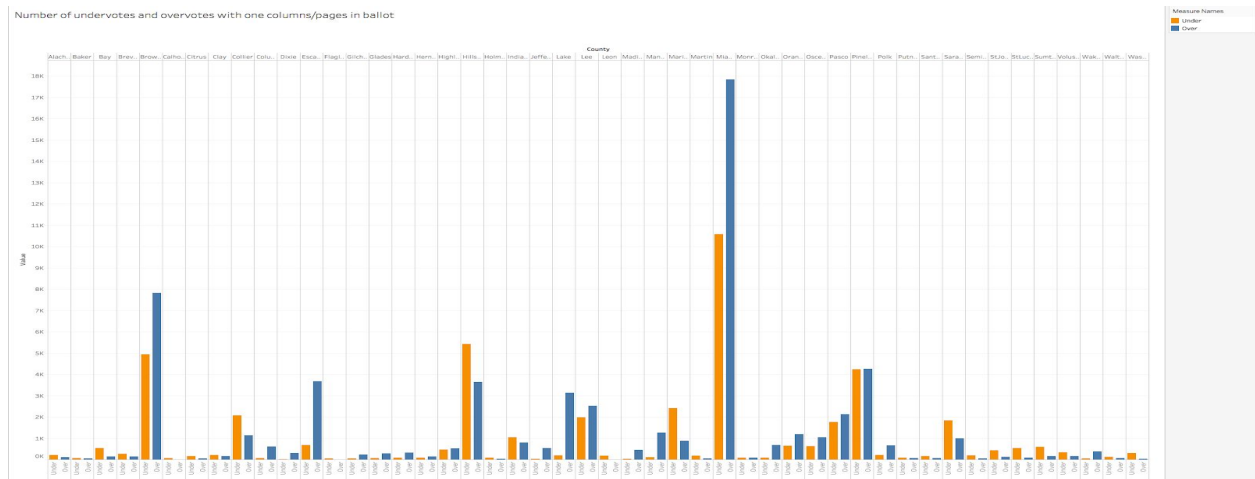
I filtered the number of overvotes and undervotes by Votomatic type to gain a better overview. The number of undervotes can sometime overtakes the number of overvotes.



But for ballots with two columns/pages, the number of overvotes is always greater than the number of undervotes for Votomatic voting machine type. Also, it is interesting to note that the number of undertypes is higher in Palm Beach County than in Duval County.



This is also true among all counties with two columns/pages ballots. The number of undervotes is fewer than the number of overvotes. Also, there are more undervotes in Palm Beach County than in other comparable counties with two columns/page ballots.



I want to see if the reverse holds true for one column/page ballot. It does not seem like it.

## Conclusion

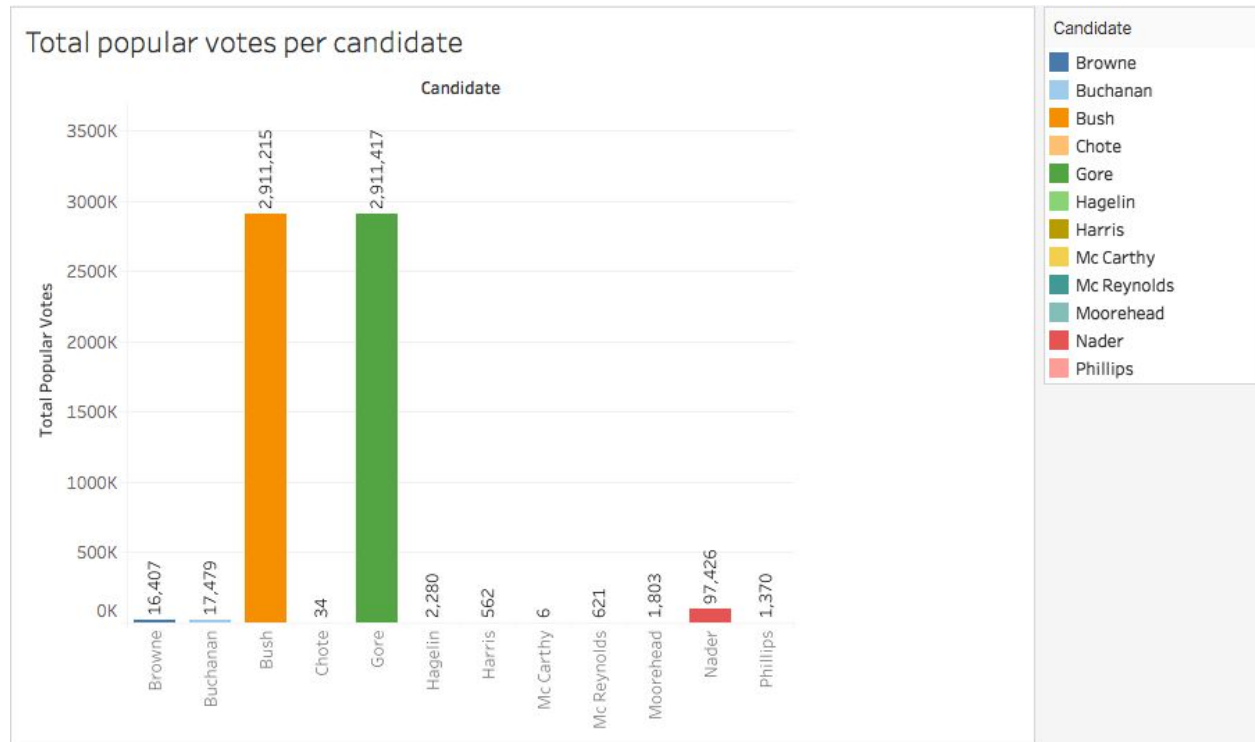
It is true that there is a higher number of people who overvote than undervote in Palm Beach County, but that is true in general for counties with two columns/pages ballots. What is more



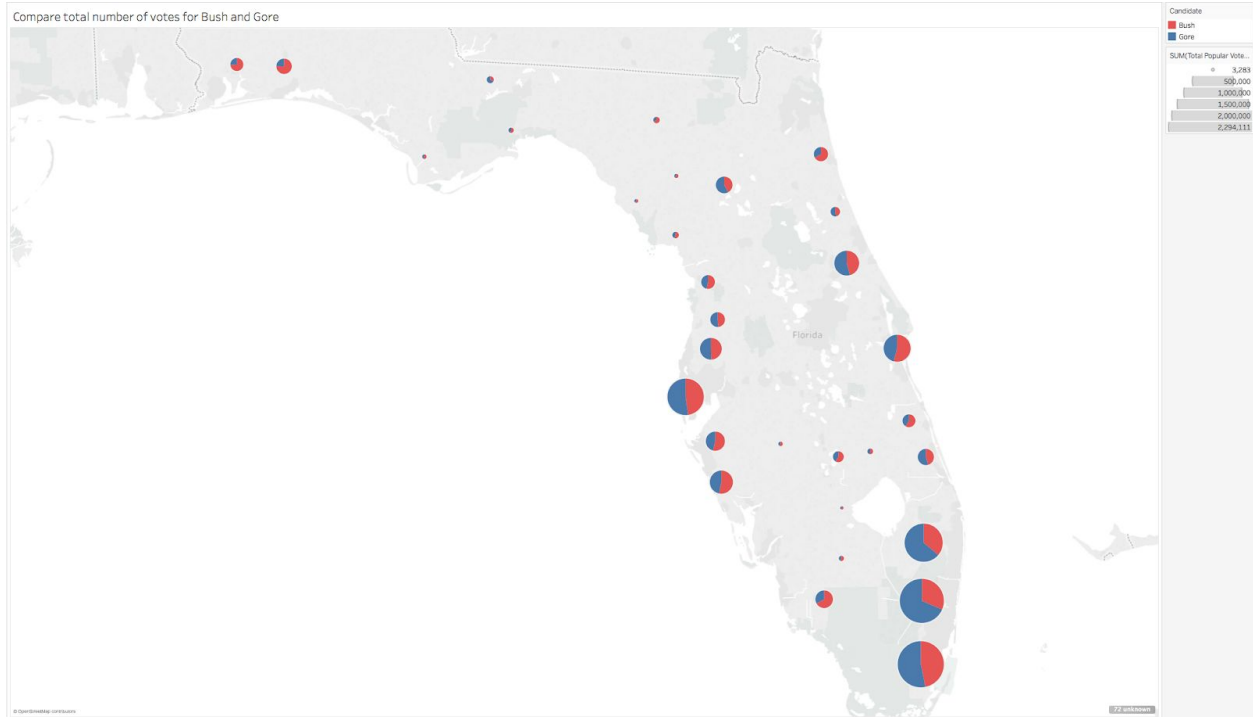
interesting to note is that the number of undervotes in Palm Beach County is higher than all other counties' undervotes relative to the number of overvotes.

### Hypothesis 3

For the 2000 Presidential Election, Al Gore is more popular in southern Florida with its more liberal constituents while George Bush is more popular in northern Florida with its more conservative constituents.



Before comparing the two candidates, I would like to verify that Gore's total vote is actually higher than Bush's by a slight margin, not including the federal absentee votes. The graph confirms it to be true.



I then map the total number of votes with red representing the Republican Bush and blue representing the Democrat Gore. From the map with pie chart representing the total number of votes for each candidate, the north seems to support Bush more compare to the south which seems to have a stronger support for Gore.

### Conclusion

These data views seem to suggest that Al Gore is more popular in southern Florida with its more liberal constituents while George Bush is more popular in northern Florida with its more conservative constituents.