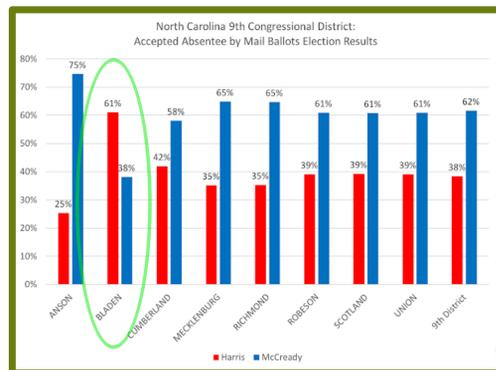


# Data Science for Election Verification



S. Singer and B. Yesilada, Hatfield School of Government, Portland State University



Data analysis

can protect elections



## HOW CAN DATA SCIENCE BE USED SYSTEMATICALLY AND BROADLY TO PROTECT ELECTIONS?

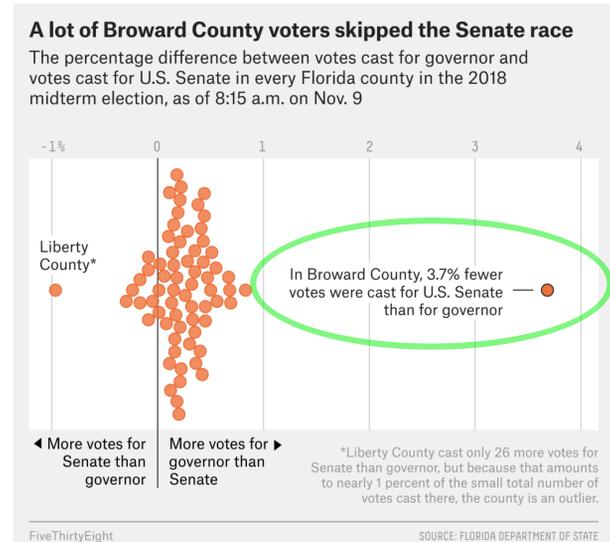
### Ad Hoc Example: North Carolina, 2018 General Election, 9<sup>th</sup> Congressional district

Blogger posts graph of absentee ballot results showing anomaly in Bladen County. Subsequent investigation reveals a “coordinated, unlawful and substantially resourced” operation, possibly involving over 1,000 ballots in a contest won by under 1,000 votes. North Carolina State Board of Elections declines to certify the results. No winner is declared, and a new election is conducted several months later.

### Ad Hoc Example: Florida 2018 General Election, Broward County ballot design

FiveThirtyEight.com publishes scatterplot a few days after the polls close, showing two outliers. One outlier is inconclusive because of small sample size. The other one draws attention to a design flaw in the Broward County ballot.

Our goal is to create systematic, sophisticated tools to analyze election results in real time.



**Technical approach:** start simple, build automated tools, use common data formats

**Key innovation:** don't try to assess likelihood of fraud, but rather use data science to identify anomalies for investigation by candidates and election administrators with boots on the ground and skin in the game

**Key innovation:** let timing of election events – from end of voting to the certification of final results weeks later – drive the tool development

## DIRECT BENEFIT TO SOCIETY:

Increases resilience of election systems, which are “Critical Infrastructure” essential to national security.

## SCIENTIFIC IMPACT:

Applies data science in an atypical setting: relatively sparse data and high incidence of unpredictable, often legitimate anomalies.

## BROADER IMPACT:

Brings data science to the attention of the diverse, broad population of election stakeholders

Award ID#: 1936809

Bladen County chart: <http://www.oldnorthstatepolitics.com/2018/11/ncs-closest-congressional-contest-gets.html#more>, Washington Post screenshot: [https://www.washingtonpost.com/politics/north-carolina-sets-new-primaries-general-election-in-scandal-plagued-9th-congressional-district/2019/03/04/ccfec0a8-3ea1-11e9-9361-301ffb5bd5e6\\_story.html](https://www.washingtonpost.com/politics/north-carolina-sets-new-primaries-general-election-in-scandal-plagued-9th-congressional-district/2019/03/04/ccfec0a8-3ea1-11e9-9361-301ffb5bd5e6_story.html), Florida scatterplot: <https://fivethirtyeight.com/features/something-looks-weird-in-broward-county-heres-what-we-know-about-a-possible-florida-recount>, “... coordinated, unlawful...” quote: <https://www.nytimes.com/2019/02/18/us/north-carolina-election-fraud.html>

