Problem Set 1 Solutions

Sean Westwood

Fraud and Turnout in the 2020 Georgia Presidential Election

This assignment is built on a controversial working paper: Lott, John R., "A Simple Test for the Extent of Voter Fraud with Absentee Ballots in the 2020 Presidential Election: Georgia and Pennsylvania Data," Unpublished Manuscript, https://papers.ssrn.com/sol3/papers.cfm?abstractid=3756988, 2021. The author attempts to identify fraud by comparing precinct in Georgia. For this assignment we will explore turnout in these precinct.

The following variables are in the georgia.csv file:

Variable	Description
CountyWithFraud	County with accused fraud
countyName	County Name
precinct	Precinct name
registeredVoters	Number of registered voters
trumpElectionDayVotes	Number of votes for Trump on election day
trumpAdvancedVotingVotes	Number of advance votes for Trump
trumpAbsenteeVotes	Number of absentee votes for Trump
trumpProvisionalVotes	Number of provisional votes for Trump
trumpTotalVotes	Total Trump votes
bidenElectionDayVotes	Number of votes for Biden on election day
bidenAdvancedVotingVotes	Number of advance votes for Biden
bidenAbsenteeVotes	Number of absentee votes for Biden
bidenProvisionalVotes	Number of provisional votes for Biden
bidenTotalVotes	Total Biden votes
totalOfAllVotes	Total votes for all candidates
percentWhite	Percent of the precinct that is White
percentBlack	Percent of the precinct that is Black
percentHispanic	Percent of the precinct that is Hispanic

Loading the data

```
georgia <- read.csv("georgia.csv")

georgia <- unique(georgia)
georgia <- georgia[georgia$percentHispanic != 0.006395614,
    ]
```

Questions

Question 1

1.1. How many precincts are there in the data set (ignore multiple entries)?

```
nrow(georgia)
```

```
[1] 35
```

Answer: 35

1.2. What is the range of total votes for Trump?

```
range(georgia$trumpTotalVotes)
```

```
[1]    180 17581
```

Answer: 180, 17,581

1.3. What is the range of total votes for Biden?

```
range(georgia$bidenTotalVotes)
```

```
[1]    64 7648
```

Answer: 64, 7,648

Question 2

2.1. Trump and Biden were not the only presidential candidates on the ballot in Georgia. Calculate the proportion of votes for a third party in each precinct (this is commonly called a vote share).

$$Third\ party\ vote\ share = \frac{Total\ votes - (Total\ Biden\ votes + TotalTrump\ Votes)}{Total\ votes}$$

```
(georgia$totalOfAllVotes - (georgia$trumpTotalVotes +
    georgia$bidenTotal))/georgia$totalOfAllVotes
```

```
 [1] 0.011261261 0.008660352 0.012539185 0.012829651 0.014610390 0.009942439
 [7] 0.007657658 0.007798960 0.013307985 0.012522361 0.013258542 0.018617021
[13] 0.016631130 0.015064562 0.008746356 0.009708738 0.009009009 0.010401189
[19] 0.014207650 0.006657019 0.012320329 0.014496439 0.017971368 0.013858337
[25] 0.011788977 0.011128127 0.015608818 0.018744906 0.016576251 0.015533656
[31] 0.014877001 0.017400205 0.011767420 0.013450292 0.012143018
```

2.2. What was the largest vote share for a third party observed in the data? (round to two decimal places)

```
round(max((georgia$totalOfAllVotes - (georgia$trumpTotalVotes +
    georgia$bidenTotal))/georgia$totalOfAllVotes),
    2)
```

```
[1] 0.02
```

Answer: 0.02

Question 3

3.1. Compute the percentage of Trump votes that were cast on election day and save this to a new column called `percentTrumpElectionDay`. Do the same for Biden to create a `percentBidenElectionDay`

```
georgia$percentTrumpElectionDay <- (georgia$trumpElectionDayVotes/georgia$trumpTotalVotes) *
    100
georgia$percentBidenElectionDay <- (georgia$bidenElectionDayVotes/georgia$bidenTotalVotes) *
    100
```

3.2. In how many precincts was election day turnout larger for Trump than for Biden? (hint: use the > operator to test an inequality; hint: the table() function could be helpful)

```
table(georgia$trumpElectionDayVotes > georgia$bidenElectionDayVotes)
```

```
FALSE   TRUE
    4     31
```

Answer: 31

3.3. Compute the percentage of Trump votes that were cast in advance of election day and save this to a new column called `percentTrumpAdvance`. Do the same for Biden to create a `percentBidenAdvance`. In how many precincts was the percent of advance turnout larger for Biden than Trump?

```
georgia$percentTrumpAdvance <- georgia$trumpAdvancedVotingVotes/georgia$trumpTotalVotes
georgia$percentBidenAdvance <- georgia$bidenAdvancedVotingVotes/georgia$bidenTotalVotes

table(georgia$percentBidenAdvance > georgia$percentTrumpAdvance)
```

```
FALSE   TRUE
   28      7
```

Answer: 10

Question 4

4.1. In how many precincts did Trump beat Biden?

```
table(georgia$trumpTotalVotes > georgia$bidenTotalVotes)[2]
```

```
TRUE
  23
```

Answer: 23

4.2. Create a new variable called `advanceAbsenteeProvisional` that records how many votes Biden earned form advance voting, absentee voting and provisional.

```
georgia$advanceAbsenteeProvisional <- georgia$bidenAbsenteeVotes +
    georgia$bidenAdvancedVotingVotes + georgia$bidenProvisionalVotes
```

4.3. Compute Biden's margin of victory (Biden's total votes - Trump's total votes) in each precinct and save to a new variable called `bidenMargin`. If the result of this calculation is negative, then Trump won. It if is positive them Biden won.

```
georgia$bidenMargin <- georgia$bidenTotalVotes - georgia$trumpTotalVotes
```

4.4. In how many precincts did Biden defeat Trump because of advance, absentee and provisional voting? From Question 4.3. we know how many precincts Biden won. Now assume that all advance, absentee and provisional voting for Biden didn't happen by subtracting these votes from Biden's total. Without advance, absentee and provisional voting how many precincts did Biden win? (hint: it might be useful to use the approach from 4.3 and look at the sign of the margin).

```
table(sign((georgia$bidenTotalVotes - georgia$advanceAbsenteeProvisional) -
    georgia$trumpTotalVotes))
```

```
-1
35
```

Answer: 0

4.5. Repeat 4.4, but only remove absentee voters from Biden's total. In how many precincts did Biden win?

```
table(sign((georgia$bidenTotalVotes - georgia$bidenAbsenteeVotes) -
    georgia$trumpTotalVotes))[2]
```

```
1
6
```

Answer: 6