## Analyzing Texas Election Data Related to the 2020 U.S. Presidential Election

Matt Schroeder, Seungwoo Nam, Suyoung Park

12/17/2020

#### Introduction

- This election cycle, Texas saw more pre-election national political media interest than it had in our lifetimes
- The interest in Texas was partially due to record early vote turnout related to the COVID-19 pandemic
  - 108.3% of the total 2016 turnout was reached before election day
- Texas known for low voter turnout in Presidential elections
  - The state ranked 47th nationally for voter turnout in 2008, 2012, and 2016
- Democrats invested in last minute advertisements and a high profile visit by then VP Candidate Kamala Harris
- Rated as a toss up by NBC's Political Unit and Cook Political Report in late October
- The last time Democrats won Texas was in 1976.

#### **Data Wrangling - Overview**

- Texas Secretary of State html
  - Early vote data in Texas from 1996 to 2020
  - 2018 general election (Senator/Governor) data in Texas
  - 2020 election day data in Texas
- A Harvard Dataverse csv
  - Presidential election data from 1976 to 2016
- CENSUS csv
  - Demographic data in Texas

Red color: data sources

black color: types of our datasets blue color: the format of raw data. Data Wrangling - Texas Secertary of State



Figure 1: Raw Data from Texas Secretary of State



Figure 2: Corresponding HTML Code

### **Data Wrangling - Texas Secertary of State**

- 1 rvest
- 2 XML and XML2

#### **Data Wrangling - Texas Secertary of State**

```
##
       County Precints Registered Voters Suspense Voters Non Suspense Voters
## 340
         Wood
                     14
                                    32382
                                                      2872
                                                                           29510
## 341 Yoakum
                                     4407
                                                       383
                                                                           4024
## 342 Young
                     10
                                    11769
                                                      1099
                                                                          10670
## 343 Zapata
                      8
                                     8257
                                                       953
                                                                           7304
## 344 Zavala
                                      8066
                                                       620
                                                                           7446
       Cumulative In Person Cumulative In Person Percent Cumulative By Mail
## 340
                       16936
                                                     57.39
                                                                          2372
## 341
                        2011
                                                     49.98
                                                                           113
## 342
                        6215
                                                     58.25
                                                                           808
## 343
                        2238
                                                     30.64
                                                                           169
## 344
                        2477
                                                     33.27
                                                                           494
       Cumulative_By_Mail_Percent Cumulative_In_Person_By_Mail
## 340
                              8.04
                                                            19308
## 341
                              2.81
                                                            2124
## 342
                              7.57
                                                            7023
## 343
                              2.31
                                                            2407
## 344
                              6.63
                                                            2971
##
       Cumulative Percent Early Vote Year
## 340
                                65.43 2020
## 341
                                52.78 2020
## 342
                                65.82 2020
## 343
                                32.95 2020
## 344
                                39.90 2020
```

### Data Wrangling - 2020 Demographic data

Since CENSUS data contains the observations from 2010 to 2019, we projected the 2020 population based on 2018-2019 data.

 Assumption: the percent changes of population between 2018-2019 and 2019-2020 are the same.

$$POP_{2020,j} = (1 + \frac{POP_{2019,j} - POP_{2018,j}}{POP_{2018,j}}) * POP_{2019,j}$$

where j is the each variable.

#### **Data Wrangling - Result**

Total # of Observation	# of Variable			
Texas Secretary of State				
6858	9			
344	8			
1778	5			
516	7			
Harvard Dataverse				
3740	14			
CENSUS				
57912	80			
	of State  6858 344 1778 516 verse 3740			



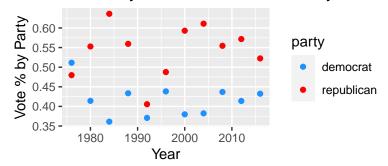
Name	Total # of Observation	# of Variable	Datasets
Early_Vote_Data_By_Day.csv	6858	7	(1) + (3)
Early_Vote_Data.csv	344	12	(2) + (3)
2018_Election_Data.csv	510	8	(4)
General_Election_Data.csv	22	6	(5)
2018_Demographic_Data_By_County_Total.csv	250	20	(6)
2020_Demographic_Data_By_County_Total.csv	250	20	(6)
2018_Demographic_Data_By_County.csv	3556	21	(6)
2020_Demographic_Data_By_County.csv	3556	21	(6)

 Table 1: Output of Data Wrangling Process

#### **Data Analysis Overview**

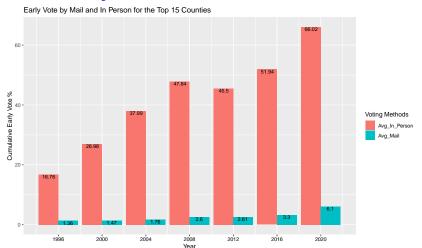
- Understanding the Historical Voting Patterns
- Understanding the 2020 Early Vote in Texas
- Showing the Swing Counties in '18
- Investigating the Whole Foods Counties
- Understanding Election Day Turnout

## Historic Voting Patterns Two Party Candidate Vote Share by Year



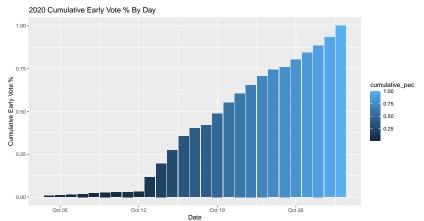
- Last time a Democrat won the state of Texas was in 1976 (Jimmy Carter's Victory)
- Lower two party vote share in 1992 and 1996
  - Third party candidate (and Texan) Ross Perot received 22.01% and then 6.75% of the vote
- Historically, Texas hasn't been close enough to be a toss up

#### Historical Early Vote and Mail Data



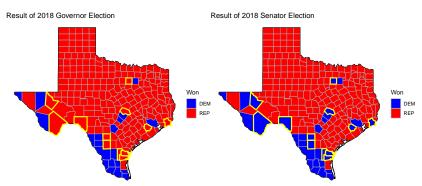
- 2020 Voting pattern deviates from the historic voting pattern
  - Record setting early in person and mail vote in 2020

### 2020 Early Vote Percentage By Day



 When looking at the state's cumulative early votes by day, we see an incredible surge in early voting on October 13th, the first day of early voting. The votes before then represent mail in votes.

### Swing Counties from the 2018 Election

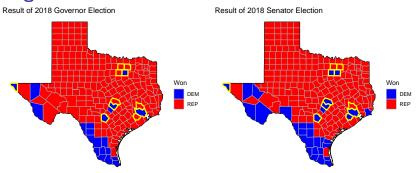


- A small minority of Texas counties are competitive
  - 12 Counties voted Republican for the Governor Election, but voted Democrat in the Senate Election: Brewster, Fort Bend, Frio, Jefferson, Jim Wells, Kleberg, La Salle, Nueces, Reeves, Tarrant, Val Verde, Williamson
- This group of counties is to the left of the average Texas county

#### **Texas Counties with Whole Foods**

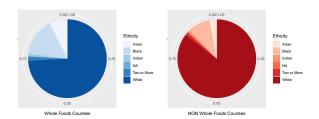
- Split the state up into urban and rural areas based on whether or not they have a Whole Foods store
  - Popularized by prominent political analyst David Wasserman
- Only 11 counties have a Whole Foods store:
  - El Paso, Bexar, Travis, Williamson, Montgomery, Harris, Fort Bend, Tarrant, Dallas, Collin, Denton
- 54.83% of Texas's population lives in one of these 11 counties

#### **Voting Habits of Citizens in Whole Food Counties**



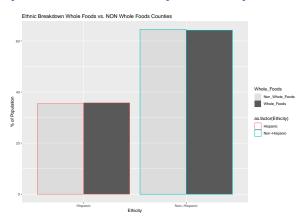
- 5 out of 11 counties voted for a Democratic Candidate in the 2018 Governor Election
- 8 out of 11 counties voted for a Democratic Candidate in the 2018 Senator Election
- Urban areas leaning Democratic is not unique to Texas, rather a nationwide phenomena

## Comparing the Racial Demographics of Whole Foods vs. Non-Whole Foods Counties



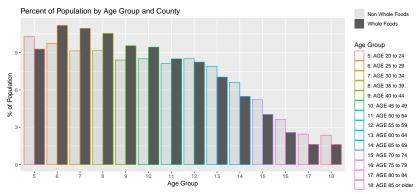
 The clearest difference is that the Whole Foods counties have a lower percentage of White residents and a higher percentage of African and Asian Americans compared to non-Whole Foods counties.

#### Hispanic vs. Non-Hispanic Populations



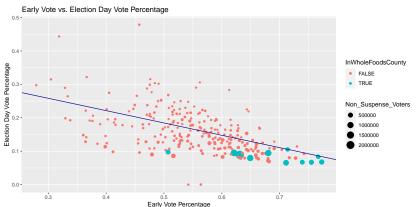
- The Hispanic population of the two sets of counties is within one percent.
  - On the previous slide Hispanics are grouped with Whites, this shows Whites grouped with other races as Non-Hispanic

# Age Distribution of Whole Food Counties vs Non Whole Food Counties.



- Non Whole Food Counties have a larger population of citizens of age 20-24 and age groups 55-59 and higher.
- Whole Food Counties has a larger population of citizens from age groups 25-29 to 50-54.

### Early Vote vs. Election Day Turnout



- Early vote turnout was negatively correlated with Election Day turnout
- Whole Foods counties had higher early vote turnout and lower election day turnout

#### **Conclusion**

- Ultimately, Republicans took the state again by 5.6% of the votes making it their 11th Presidential win in the state in a row
- Our research gave some context to why many considered the state to be a toss up state due to the early vote.
- We showed how the voting patterns deviated from historic norms this election cycle.
- We invesigated the swing counties from 2018
- Provided an extensive look at how the constituents in the Whole Foods counties and non-Whole Foods counties differ by race, age, and political preferences.
- We examined election day vs. early vote turnout.