CO 2017 Variable Analysis

Theodore Dounias

5/10/2018

Dimensions

```r
#Voter History Records Dimensions
dim(CO_2017_VHist_full)
```

```
## [1] 34182666        8
```

```r
#Voter Registration Records Dimensions
#dim(CO_2017_VRF_full)
```

Voter History Variables

Voter ID

```r
#Voter ID NA's
sum(is.na(CO_2017_VHist_full$VOTER_ID))
```

```
## [1] 0
```

```r
#Unique Voter ID Values
length(unique(CO_2017_VHist_full$VOTER_ID))
```

```
## [1] 3270401
```

This last number is smaller than the total amount of entries, which makes sense since multiple records of voting exist for single people.

Election Type

```r
#Number of NA's
sum(is.na(CO_2017_VHist_full$ELECTION_TYPE))
```

```
## [1] 0
```

```r
#Summary of election types
summary(as.factor(CO_2017_VHist_full$ELECTION_TYPE))
```

```
##       Coordinated             General    Municipal Municipal Run-off
##           8688825            19190198       590774            133694
##           Primary              Recall       School           Special
##           4873770              148576        83518            416715
##   Special District
##             56596
```

There are 9 types of elections listed: Coordinated, General, Municipal, Municipal Run-Off, Primary, Recall, School, Special, and Special District.

Election Date

```r
#Number of NA's
sum(is.na(CO_2017_VHist_full$ELECTION_DATE))
```

```
## [1] 0
```

```r
#Earliest and Latest Date Listed
dates_vhist <- mdy(CO_2017_VHist_full$ELECTION_DATE)

#Earliest Date
min(dates_vhist)
```

```
## [1] "1900-01-01"
```

```r
#Latest Date
max(dates_vhist)
```

```
## [1] "2017-11-07"
```

```r
#Number of Dates--Not necessarily number of elections!
length(unique(dates_vhist))
```
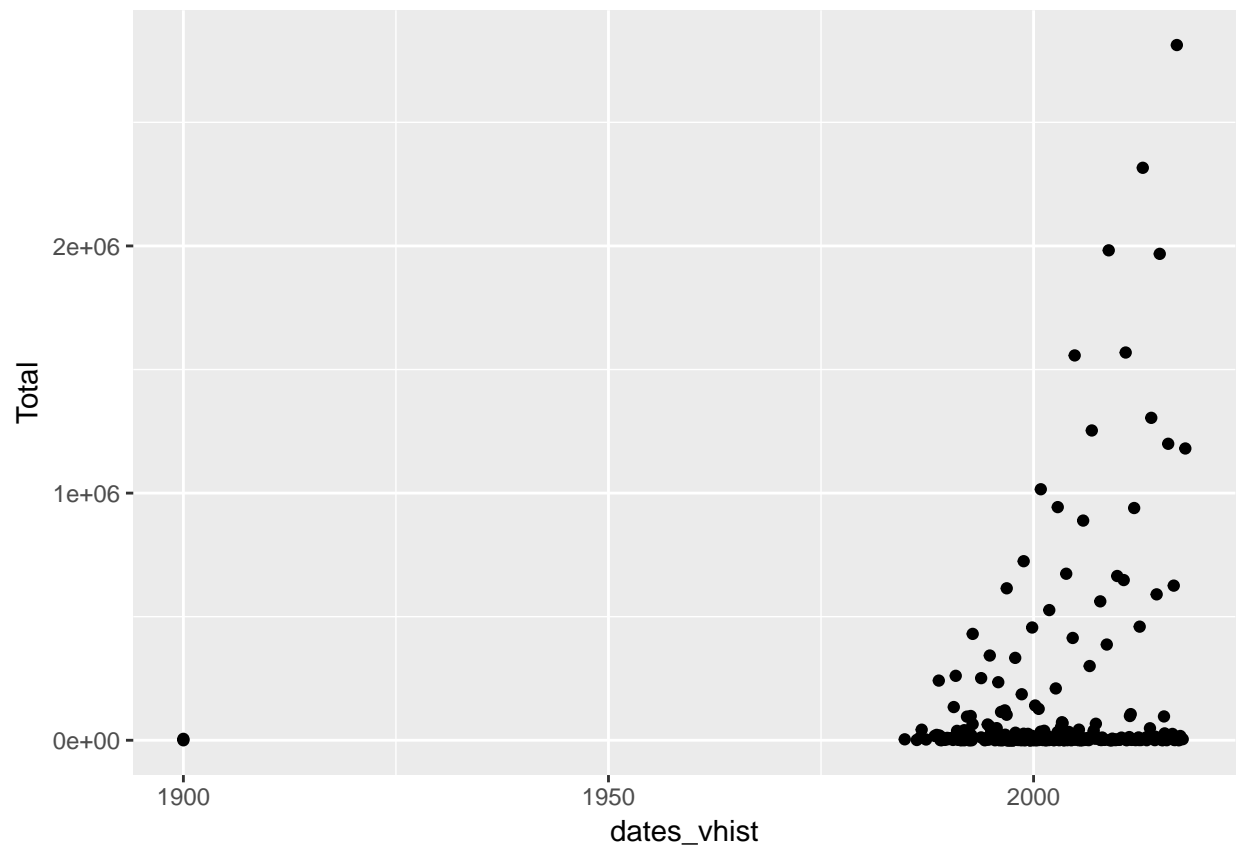
```
## [1] 246
```

```r
dates_vhist <- data.frame(dates_vhist)

#Distribution of Dates
dist_dates <- dates_vhist %>%
  mutate("count" = 1) %>%
  group_by(dates_vhist) %>%
  summarise("Total" = sum(count))

#Plot of all dates with frequencies
ggplot(dist_dates, aes(y = Total)) +
  geom_point(aes(x = dates_vhist))
```
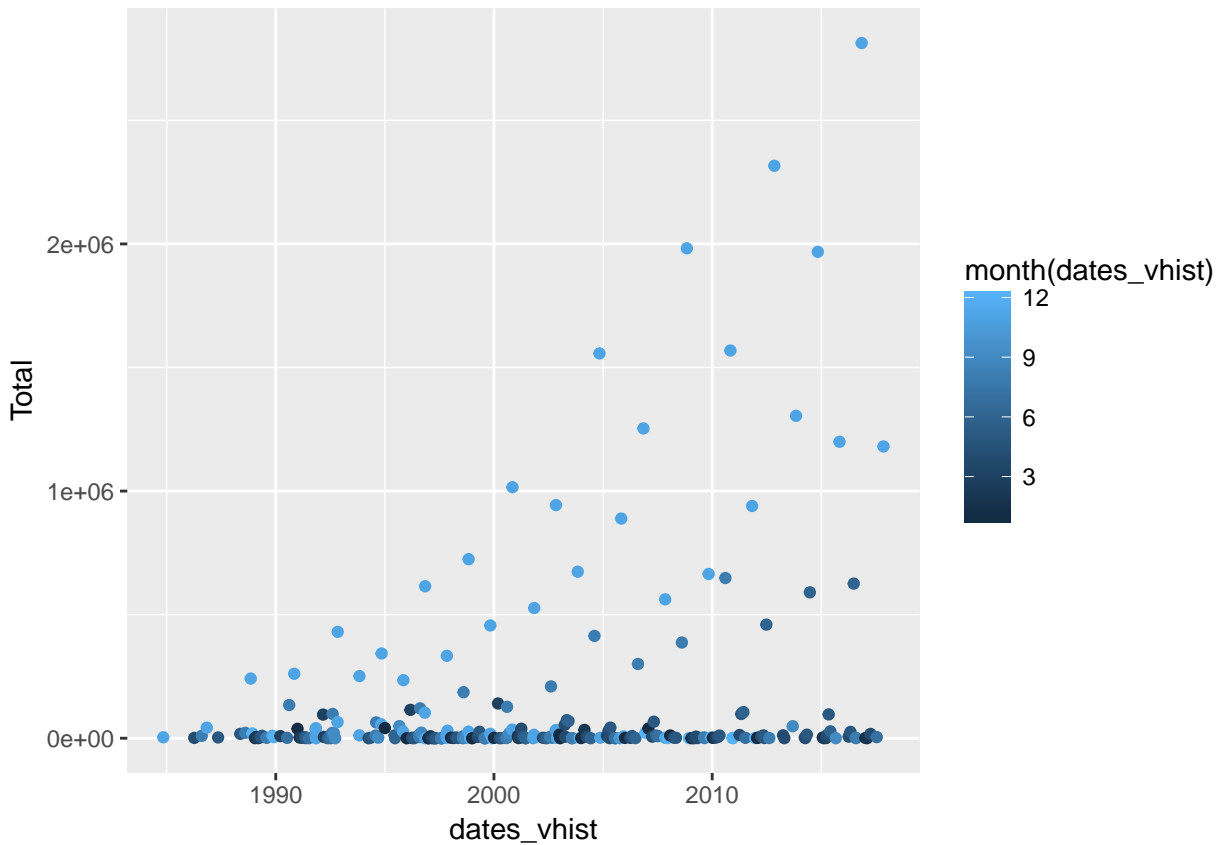
```
#Number of invalid dates
kable(dist_dates[1:2,])
```

dates_vhist	Total
1900-01-01	5747
1900-01-02	1

```
#Plot with invalid dates removed
dist_dates <- dist_dates[-(1:2),]

#Plot without invalid dates
ggplot(dist_dates, aes(y = Total)) +
  geom_jitter(aes(x = dates_vhist, col = month(dates_vhist)))
```

The plots in this section show how many entries are present for each date. Each point represents a date on the x axis, and the y axis is the number of observations/votes for that date.

Election Description

```
#Different Election Descriptions
summary(as.factor(CO_2017_VHist_full$ELECTION_DESCRIPTION))
```

```
##                             GENERAL ELECTION
##                                       443458
##                        2008 GENERAL ELECTION
##                                       357169
##            2016 Denver County General Election
##                                       335537
##         2016 Jefferson County General Election
##                                       332926
##           2016 El Paso County General Election
##                                       321312
##          2016 Arapahoe County General Election
##                                       307152
##                             General Election
##                                       295828
##                               2004 - GENERAL
##                                       288779
##                           11/02/2004 GENERAL
##                                       287151
```

4

```
##          2012 Jefferson County General Election
##                                          283701
##          2012 Denver County General Election
##                                          270693
##                        General Election 2008
##                                          261652
##          2012 Arapahoe County General Election
##                                          258419
##          2012 El Paso County General Election
##                                          254927
##                      General Election - 2008
##                                          253754
##          2014 Jefferson County General Election
##                                          247407
##                              2006 - GENERAL
##                                          239727
##                          11/07/2006 GENERAL
##                                          234399
##          2014 Denver County General Election
##                                          222452
##                          11/04/2008 General
##                                          221418
##          2014 El Paso County General Election
##                                          217576
##          2014 Arapahoe County General Election
##                                          216170
##                        2008 General Election
##                                          211250
##                              2000 - GENERAL
##                                          203082
##          2010 Jefferson County General Election
##                                          202143
##                          11/07/2000 GENERAL
##                                          198085
##          2016 Larimer County General Election
##                                          197453
##                          COORDINATED ELECTION
##                                          197215
##          2016 Adams County General Election
##                                          195870
##          2016 Douglas County General Election
##                                          189905
##          2016 Boulder County General Election
##                                          189835
##                              2002 - GENERAL
##                                          185522
##  2015 Jefferson County Coordinated Election
##                                          179574
##                          11/05/2002 GENERAL
##                                          177925
##                          2005 - COORDINATED
##                                          176926
##          2010 Arapahoe County General Election
##                                          174009
```

5

```
##                    11/02/2004 GENERAL ELECTION
##                                          171874
##           2010 El Paso County General Election
##                                          168950
##   2013 Jefferson County Coordinated Election
##                                          164925
##             2010 Denver County General Election
##                                          164899
##   2004 GENERAL & SPECIAL MUNICIPAL ELECTIONS
##                                          161790
##           2012 Larimer County General Election
##                                          161598
##             2012 Adams County General Election
##                                          160566
##           2012 Boulder County General Election
##                                          159926
##                                    2008-GENERAL
##                                          159337
##     2017 El Paso County Coordinated Election
##                                          153113
##           2012 Douglas County General Election
##                                          153030
##                               1996 - GENERAL
##                                          150902
##                               1998 - GENERAL
##                                          150476
##                         11/03/1998 GENERAL
##                                          142839
##     2017 Denver County Coordinated Election
##                                          142664
##           2014 Larimer County General Election
##                                          140224
##     2015 El Paso County Coordinated Election
##                                          139473
##                    11/07/2006 GENERAL ELECTION
##                                          138525
##     2013 El Paso County Coordinated Election
##                                          138118
##           2014 Boulder County General Election
##                                          137564
##                         GENERAL ELECTION 2008
##                                          136246
##             2016 Weld County General Election
##                                          135947
##                         11/05/1996 GENERAL
##                                          135430
##                                    2004-GENERAL
##                                          134620
##           2014 Douglas County General Election
##                                          133590
##   2017 Jefferson County Coordinated Election
##                                          133165
##   2013 Arapahoe County Coordinated Election
##                                          130304
```

```
##       2013 Denver County Coordinated Election
##                                         129755
##          2014 Adams County General Election
##                                         126180
##                              1992 - GENERAL
##                                         125805
##              11/07/2000 GENERAL ELECTION
##                                         123794
##           2006 GENERAL & SPECIAL MUNICIPAL
##                                         123526
##                          2003 COORDINATED
##                                         122650
##     2015 Denver County Coordinated Election
##                                         119261
##  2011 Jefferson County Coordinated Election
##                                         118461
##                      2004 GENERAL ELECTION
##                                         114541
##                        11/03/1992 GENERAL
##                                         113383
##   2015 Arapahoe County Coordinated Election
##                                         112131
##         2010 Boulder County General Election
##                                         110434
##                              1996 GENERAL
##                                         110262
##                          2007 - COORDINATED
##                                         109640
##         2010 Larimer County General Election
##                                         109216
##      2011 Denver Municipal Run-off Election
##                                         105877
##                11/05/2002 GENERAL ELECTION
##                                         105827
##                              2004 GENERAL
##                                         105580
##          2012 Weld County General Election
##                                         105573
## 2000 GENERAL AND SPECIAL MUNICIPAL ELECTION
##                                         105010
##                               2006-GENERAL
##                                         104570
##         2010 Douglas County General Election
##                                         103481
##                              1994 - GENERAL
##                                         102494
##    2013 Douglas County Coordinated Election
##                                         100572
##    2011 El Paso County Coordinated Election
##                                         100546
##                          1999 - COORDINATED
##                                         100404
##          2010 Adams County General Election
##                                          98357
```

```
##                2011 Denver Municipal Election
##                                        98198
##                11/01/2005 COORDINATED/COMBINED
##                                        98176
##       2013 Boulder County Coordinated Election
##                                        97376
##                           2001 - COORDINATED
##                                        97054
##       2009 El Paso County Coordinated Election
##                                        96846
##                         2006 GENERAL ELECTION
##                                        96786
##         2015 Denver Municipal General Election
##                                        96475
##     2002 GENERAL & SPECIAL MUNICIPAL ELECTIONS
##                                        95550
##        2011 Denver County Coordinated Election
##                                        94893
##                                       (Other)
##                                      17473486
```

```r
#Why is the previous function cutting off?
length(unique(CO_2017_VHist_full$ELECTION_DESCRIPTION))
```

```
## [1] 1595
```

```r
#Too many values!

#Number of NA's
sum(is.na(CO_2017_VHist_full$ELECTION_DESCRIPTION))
```

```
## [1] 0
```

Voting Method

```r
#Number of NA's
sum(is.na(CO_2017_VHist_full$VOTING_METHOD))
```

```
## [1] 1269812
```

```r
#Percentage NA
sum(is.na(CO_2017_VHist_full$VOTING_METHOD))/length(CO_2017_VHist_full$VOTING_METHOD)
```

```
## [1] 0.03714783
```

```r
#Types of voting methods
summary(as.factor(CO_2017_VHist_full$VOTING_METHOD))
```

```
##      Absentee Carry        Absentee Mail        Early Voting
##              106196             13258152             1224179
## Early Voting - DRE            In Person      In Person - DRE
##              418364               187168              160042
##         Mail Ballot   Mail Ballot - DRE         Polling Place
##             8397030                 2840              8746351
##         Vote Center   Vote Center - DRE                 NA's
##              231354               181178              1269812
```

There is a large number of NA's in this variable, but it is only about 3% of the data; which seems reasonable.

Party

```
#Number of NA's
sum(is.na(CO_2017_VHist_full$PARTY))
```

```
## [1] 624887
```

```
#Percentage of NA's
sum(is.na(CO_2017_VHist_full$PARTY))/length(CO_2017_VHist_full$PARTY)
```

```
## [1] 0.01828082
```

```
#Counts of Party Affiliations
summary(as.factor(CO_2017_VHist_full$PARTY))
```

```
##                 ACN AMERICAN CONSTITUTION                    DEM
##                9764                  5468                1625968
##            DEMOCRAT            DEMOCRATIC                  GREEN
##               16910               1463588                    349
##                 GRN     GUN OWNERS RIGHTS                    LBR
##               12377                     5                  41604
##         LIBERTARIAN               NO DATA               PRO-LIFE
##                2661              25123227                      2
##              REFORM                   REP             REPUBLICAN
##                  38               1733438                2173946
##                 UAF          UNAFFILIATED                    UNI
##             1252621                 95067                    746
##                NA's
##              624887
```

I assume here that the three letter variables are just concise forms of the whole part names.

County Name

```
#Number of NA's
sum(is.na(CO_2017_VHist_full$COUNTY_NAME))
```

```
## [1] 0
```

```
#Counts of county votes across all elections
summary(as.factor(CO_2017_VHist_full$COUNTY_NAME))
```

```
##        Adams      Alamosa     Arapahoe    Archuleta         Baca         Bent
##      2053966        96137      3709094        86775        36137        31366
##      Boulder   Broomfield      Chaffee     Cheyenne  Clear Creek      Conejos
##      2689444       326236       133980        23503        82810        68632
##     Costilla      Crowley       Custer        Delta       Denver      Dolores
##        38063        27871        45622       229703      4287463        18342
##      Douglas        Eagle       Elbert      El Paso      Fremont    Garfield
##      1940459       248526       192444      3567938       291829       313742
##       Gilpin        Grand     Gunnison     Hinsdale     Huerfano      Jackson
##        40622       110281        99259        10511        57919        15082
##    Jefferson        Kiowa    Kit Carson         Lake     La Plata      Larimer
```

```
##     4558186      17634      51812      43923     346228    2287320
## Las Animas    Lincoln      Logan       Mesa    Mineral     Moffat
##      106573      44428     142804    1142624       9986     102418
##   Montezuma   Montrose     Morgan      Otero      Ouray       Park
##      156604     288434     143654     114665      49152     122798
##    Phillips     Pitkin     Prowers     Pueblo Rio Blanco Rio Grande
##       44681     130499      81201    1038130      51655      75302
##       Routt   Saguache   San Juan San Miguel   Sedgwick     Summit
##      168492      40928       7771      63924      27954     170536
##      Teller Washington       Weld       Yuma
##      150301      39491    1411792      77010
```

Switching Voters? A fun result if it works (IN PROGRESS)

```r
#Counting voters that voted in more than one county
switches <- 0
per_voter <- data.frame(unique(CO_2017_VHist_full$VOTER_ID), 0)
names(per_voter) <- c("ID", "Switches")

vote_data <- CO_2017_VHist_full
vote_data <- vote_data %>%
  arrange(VOTER_ID, mdy(ELECTION_DATE)) %>%
  select(2, 8)
##FIX INDEXING!
a <- 1
for(i in 1:15100){
  if(vote_data[i, 1] == vote_data[i+1, 1] && vote_data[i, 2] != vote_data[i+1, 2]) {
    switches <- switches + 1
    per_voter$Switches[a] <- per_voter$Switches[a] + 1
  }
  if(vote_data[i, 1] != vote_data[i+1, 1]) a <- a+1
}
```