

Intro

So, a bit of a dust-up in the State Senate primaries in New Mexico over the summer. Progressives took aim at five Democratic members of a conservative coalition that has controlled the chamber since 2009. This group of moderate Dems lost four among their ranks: Gabriel Ramos (D-28), Clemente Sanchez (D-30), John Arthur Smith (D-35), and Mary Papen (D-38). Only George Muñoz (D-04) survived. For better or worse, this faction repeatedly hamstrung Michele Lujan Grisham's more progressive agenda for the state; in response, folks came out to vote in primaries.

Here, then, some methods for investigating voting behaviors and political ideologies of lawmakers in US state legislatures via roll call data made available by [Open States](#) – using as an example, the 2019 session of the New Mexico State Senate.

The ultimate goal being to provide a reproducible work-flow – from scratch – for applying DW-NOMINATE scaling procedures to any/all state legislatures in the US. State politics matter, and state lawmakers need to be held accountable for their voting records.

NM State Senate Democrats

All 42 New Mexico State Senate seats are up for election in November 2020. Sitting Senators have held these seats for the last four years, spanning the 53rd and 54th state legislatures. Democrats held a 26 to 16 majority during this time period; however, as detailed in the intro, five Democrat turncoats have kept things interesting for the better part of the teens.

In addition to the four senators who lost primaries to left-wing challengers, the center-left will additionally be without Richard Martinez (D-05), who lost his primary as well, and John Sapien (D-09), who will not be seeking re-election in November. Sitting Senate Democrats who will not be on ballots come November, then, are summarized in the table below.

Open-States Data

I have previously scraped legislative activity from the New Mexico State Legislature website for the 53rd legislature, and made it available as an R data package. I will not do this again, for any number of reasons. While most of the code I have written does scale to new legislatures, there is enough idiosyncrasy (across thousands of PDFs per session) to make the process less than pleasant. For some more details on this front, see the [Git Hub readme](#).

Instead: [Open States](#) !!! The folks at Open States have streamlined these scraping processes across all fifty state legislatures (!), and make data available in a uniform format. Here, we walk through (some of) the tables made available by Open States, specifically those relevant to **quantifying political ideology within the NOMINATE framework**. We also consider important relationships among these tables (from a database organization perspective), and some potential limitations.

/Quick notes:

Our focus here is on the New Mexico State Senate. The 54th State Legislature concluded in January 2020, at the end of the second session. (The legislature actually ended with a small special session held in June 2020). However, the state has yet to make available details of the second session. As a result, analyses here are limited to the legislative activity of the first session of the 54th, which convened January-March 2019.

/Legislators

Details on the composition each state legislature are made available by Open States [here](#). In the case of New Mexico, this list only includes legislators from the most recent session of the most recent congress, ie, the second session of the 54th legislature, which convened Jan-Feb 2020.

```
os_legislators <- read.csv(paste0(open_states_dir, '/legislators/nm.csv'))
```

Here, we make some small changes to the legislator data set to help align identifiers with those used in subsequent data sets.

```
legs2020 <- os_legislators %>%
  mutate(family_name = gsub('Lara Cadena', 'Cadena', family_name)) %>%
  arrange(current_chamber, family_name, given_name) %>%
  group_by(current_chamber, family_name) %>%
  mutate(n = n()) %>%
  ungroup() %>%
  mutate(family_name = ifelse(n>1,
                              paste0(family_name, ', ', given_name),
                              family_name)) %>%
  select(current_chamber, name, family_name, given_name, current_party,
         current_district)
```

Not taken into account by the folks at Open States is the fact that the compositions of the both state houses in New Mexico have changed a bit from the 1st (2019) to 2nd (2020) session. The table below includes state lawmakers who served in the first session but not the second. *Note that while our focus here will be the Senate, we present methods for sorting out data for both chambers.*

```
current_chamber = c('upper', 'upper', 'lower', 'lower')
name <- c('Carlos Cisneros', 'John Pinto',
         'William Praat', 'Roberto "Bobby" J. Gonzales')
family_name <- c('Cisneros', 'Pinto', 'Praat', 'Gonzales')
current_party <- rep('Democratic', 4)
current_district <- c(6, 3, 27, 42)
## session_identifier <- '2019'
```

```
adds <- data.frame(current_chamber,
                  name,
                  family_name,
                  current_party,
                  current_district)
```

```
adds %>% knitr::kable()
```

current_chamber	name	family_name	current_party	current_district
upper	Carlos Cisneros	Cisneros	Democratic	6
upper	John Pinto	Pinto	Democratic	3
lower	William Praat	Praat	Democratic	27
lower	Roberto "Bobby" J. Gonzales	Gonzales	Democratic	42

To account for these changes, then, we build an independent table for first session lawmakers of the 54th Congress.

```
legs2019 <- legs2020 %>%
  filter(!paste0(current_chamber, current_district) %in%
```

```

      paste0(adds$current_chamber, adds$current_district)) %>%
bind_rows(adds)%>%
arrange(current_chamber, family_name, given_name)

```

/Bills

Details about bills - title, language, sponsors, etc - can be accessed via Open States [here](#). Bill information is made available by legislative session; files for the 2019 session include:

```

list.files(path = open_states_dir, pattern = "NM_2019_bill_", recursive = T)
## [1] "bills/2019/NM_2019_bill_actions.csv"
## [2] "bills/2019/NM_2019_bill_document_links.csv"
## [3] "bills/2019/NM_2019_bill_documents.csv"
## [4] "bills/2019/NM_2019_bill_sources.csv"
## [5] "bills/2019/NM_2019_bill_sponsorships.csv"
## [6] "bills/2019/NM_2019_bill_version_links.csv"
## [7] "bills/2019/NM_2019_bill_versions.csv"

```

Load files:

```

setwd(open_states_dir)
bill_files <- list.files(path = open_states_dir, pattern = "bills.csv",
recursive = T)
os_bills <- lapply(bill_files, read.csv) %>% data.table::rbindlist() %>%
  mutate(session_identifier = as.character(session_identifier))

```

/Roll-calls

Next, we load roll call data. These data are scattered across multiple tables, and to get a full picture of vote results and bill details, we need to do some table joining. Below, we load three tables: (1) the `votes` table, which contains meta data about the vote/bill, and helps cross votes to bill; (2) the `votes_people` table, which contains actual legislator-level roll calls; and (3) `vote_counts`, which summarizes vote results.

```

setwd(open_states_dir)
vdetails_files <- list.files(path = open_states_dir, pattern = "votes.csv",
recursive = T)
os_vdetails <- lapply(vdetails_files, read.csv) %>% data.table::rbindlist()
%>%
  mutate(session_identifier = as.character(session_identifier))
vote_files <- list.files(path = open_states_dir, pattern = "vote_people.csv",
recursive = T)
os_votes <- lapply(vote_files, read.csv) %>% data.table::rbindlist()

##
roll <- read.csv('bills/2019/NM_2019_vote_counts.csv') %>%
  select(-id) %>%
  spread(option, value)

```

Here, we piece these three tables together to get a clearer perspective on things.

```

bill_votes <- os_vdetails %>% select(-identifier) %>%
  #mutate(session_identifier = as.character(session_identifier)) %>%
  left_join(os_bills, by = c('bill_id' = 'id',

```

```

      'jurisdiction' = 'jurisdiction',
      'session_identifier' = 'session_identifier'))
%>%
  left_join(roll, by = c('id' = 'vote_event_id'))

```

A sample of votes on the Senate side for which there is some contention:

```

bill_votes %>%
  filter(no > 10 & motion_text == 'senate passage') %>%
  select(identifier, title, yes, no) %>%
  DT::datatable(rownames = FALSE,
    options = list(dom = 't',
      pageLength = 16,
      scrollX = TRUE))

```

/Legislator votes

Lastly, we combine tables with (1) legislator details and (2) roll call information.

```

leg_votes <- os_votes %>%
  select(-id) %>%
  left_join(bill_votes, by = c('vote_event_id' = 'id'))

```

An issue with Open States data – at least in the case of New Mexico – is that legislators are not assigned a unique `voter_id` (in the `votes_people` data set) (the column exists, but is mostly empty - ?), and there are differences in how legislators are referred to in the legislator meta data and in the roll call data – so we can't join the two data sets. For now, we rely on alphabetical order and some hacks to relate the two data sets, but surnames in New Mexico can present challenges on this front.

```

legs_in_rolls <- leg_votes %>%
  group_by(organization_classification, voter_name) %>%
  count() %>%
  filter(voter_name != 'LT. GOVERNOR' & n > 100) %>%
  select(-n) %>%
  ungroup() %>%
  bind_cols(legs2019)

```

Then we restructure data in a wider format to ready things for subsequent NOMINATE-based analyses.

```

dups <- c('ocd-vote/77a697c2-c53b-4856-8631-0773e72f9f06',
  'ocd-vote/2b8128f5-768e-4630-9af5-21e986dc2fa8',
  'ocd-vote/c7a04b77-f55d-47f0-a8f1-860ed4b3a3d6')

wide_rolls <- leg_votes %>%
  filter(!vote_event_id %in% dups) %>%
  mutate(tid = paste0(session_identifier, '_',
    gsub(' ', '-', identifier))) %>%
  select(voter_name, tid, option) %>%
  mutate(vote = case_when(option == "yes" ~ 1,
    option == "no" ~ 6,
    !option %in% c(1,6) ~ 9)) %>%
  dplyr::select(-option) %>%
  spread(key = tid, value = vote) %>%
  arrange(voter_name)

```

DW-NOMINATE procedure

Next, we investigate political ideologies in the (first session of the) 54th New Mexico State Legislature using the R package `wnominate`. I have discussed these methods previously in the context of [New Mexico's 53rd State Legislature](#). Here, then, a quick/simple run through of the code.

```
chamber <- 'upper'
leg_sub <- legs_in_rolls %>% filter(current_chamber == chamber)
```

/Build roll call object

```
roll_obj <- wide_rolls %>%
  filter(voter_name %in% leg_sub$voter_name) %>%
  select(-voter_name) %>%
  pscl::rollcall(yea = 1,
                 nay = 6,
                 missing = 9,
                 notInLegis = NA,
                 vote.names = grep('-', colnames(wide_rolls), value = T),
                 legis.names = leg_sub$voter_name)
```

/Build DW-NOMINATE model

```
if(chamber == 'lower') { pol <- c('Townsend', 'Rehm')} else
  {pol <- c('INGLE', 'INGLE')}

ideal_2d <- roll_obj %>%
  wnominate::wnominate (dims = 2,
                        minvotes = 20,
                        lop = 0.025,
                        polarity = pol,
                        verbose = FALSE)
```

/Visualize two-dimensional solution

A two-dimensional model was specified. Per the multidimensional scaling procedure, the plot below represents legislators in two-dimensional political space based on roll call voting records for the first session of the 54th legislature. Dimensions:

- 1D: Right-Left → Conservative – Liberal;
- 2D: North-South → socially conservative – socially liberal.

The latter: a distinction that has become less useful (in accounting for variation in voting behavior at the federal level) over the past 20 years.

```
chamber_data <- ideal_2d$legislators %>%
  bind_cols(leg_sub)

chamber_data %>%
  ggplot(aes(x=coord1D,
             y=coord2D,
             label = family_name)) +
  geom_point(aes(color = current_party),
```

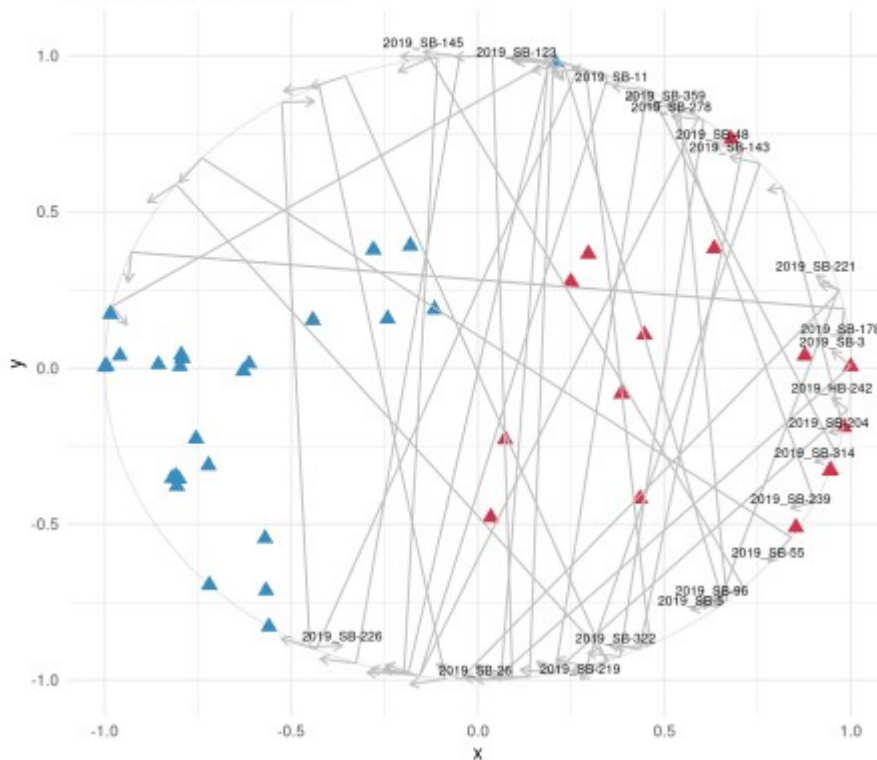


```

                                arrow_length = 0.05)
ggplot () +
  wnomadds::scale_color_party() +
  theme_minimal() +
  theme(legend.position = 'none') +
  annotate("path",
    x = cos(seq(0,2*pi, length.out = 300)),
    y = sin(seq(0,2*pi, length.out = 300)),
    color = 'lightgray',
    size = .25) +
  geom_point(data=chamber_data,
    aes(x=coord1D, y=coord2D, color = current_party),
    size= 3,
    shape= 17) +
  geom_segment(data=with_cuts,
    aes(x = x_1, y = y_1, xend = x_2, yend = y_2), color='gray',)
+
  geom_segment(data=with_cuts,
    aes(x = x_2, y = y_2, xend = x_2a, yend = y_2a), color='gray',
    arrow = arrow(length = unit(0.2,"cm")))) +
  geom_segment(data=with_cuts,
    aes(x = x_1, y = y_1, xend = x_1a, yend = y_1a), color='gray',
    arrow = arrow(length = unit(0.2,"cm")))+
  geom_text(data=with_cuts,
    aes(x = x_1a, y = y_1a, label = Bill_Code),
    size=2.5,
    nudge_y = 0.03,
    check_overlap = TRUE) +
  #coord_fixed(ratio=1) +
  labs(title = "Cutting lines & legislator coordinates",
    subtitle = "New Mexico State Senate 2019")

```

Cutting lines & legislator coordinates
New Mexico State Senate 2019



Loss of the center

The plot below summarizes ideology scores for New Mexico's State Senate (based on roll calls for the first session of the 54th congress). Again, the 1D scores reflect the traditional left-right ideological distinction. Sitting senators not running in the 2020 general election are denoted with a triangle.

```
centers <- outs %>%
  #filter(Party == 'Democrat') %>%
  mutate(label = gsub('^.* )([A-Za-z]*$)', '\\2', Senator))

chamber_data %>%
  mutate(size1 = ifelse(family_name %in% centers$label, 'Y', 'N')) %>%

  ggplot(aes(x=reorder(family_name, coord1D),
    y = coord1D,
    group = size1,
    label = family_name)) +
  geom_point(stat='identity',
    aes(col = current_party,
      shape = size1,
      size = size1)) + #, size = 3

  scale_shape_manual(values=c(16, 17))+
  scale_size_manual(values=c(2, 4.5))+

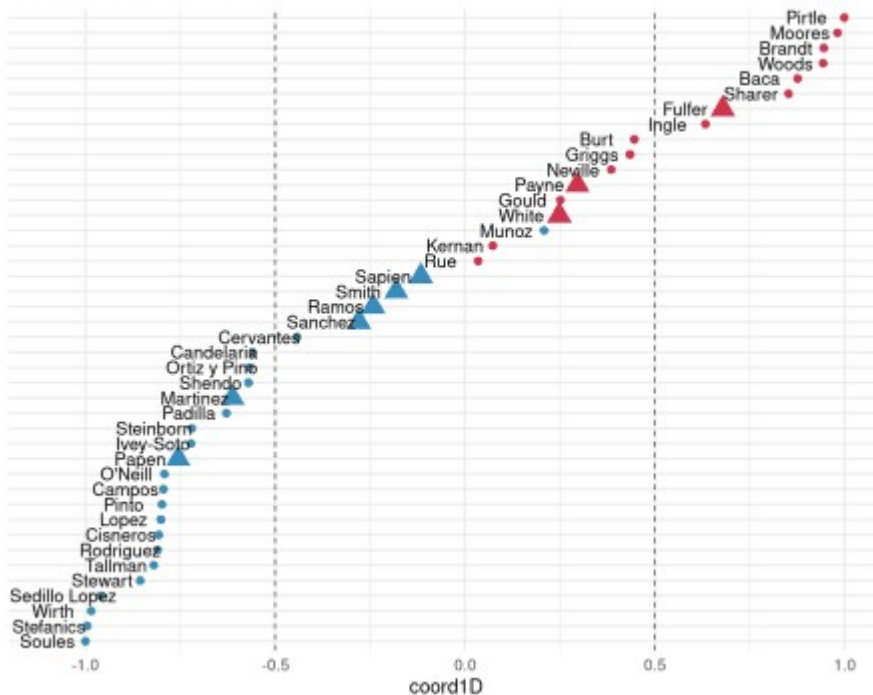
  geom_hline(yintercept = c(-0.5, 0.5),
    linetype = 2,
    color = 'black',
    size = .25) +
  wnomadds::scale_color_party()+
```



```
geom_text(size=3.5, nudge_y = -0.1) +
labs(title = "1D Ideal Point Estimates",
      subtitle = "New Mexico State Senate 2019") +

theme_minimal() +
theme(axis.title.y=element_blank(),
      axis.text.y=element_blank(),
      axis.ticks = element_blank(),
      legend.position = 'none') +
coord_flip()
```

1D Ideal Point Estimates
New Mexico State Senate 2019



So, Soules (D-37), Las Cruces, was the most liberal voter during this session; **Pirtle (R-32), Roswell**, the most conservative. [Per a previous post](#), the ideological space is fairly comparable to that of New Mexico's 53rd Senate. As can be noted per the space demarcated by dashed lines, lots of moderate voters on the way out.

Using the R package `nmlegisdatr`, the table below highlights **2016 general election margins-of-victory** for the six Democrats vacating seats in 2020. Of these six, four won uncontested races in 2016 – exceptions being the seats vacated by Papen and Sapien. Perhaps progressive nominees may inspire a contest or two among Republicans in November.

```
nmelectiondatr::nmel_results_summary %>%
  mutate(District = as.integer(Type_Sub),
         Percent = round(Percent * 100, 1)) %>%
  filter(Type == 'State Senator' & Winner == 'Winner' &
         District %in% outs$District & Party == 'DEM') %>%
  arrange(District) %>%
  select(District, Candidate:Percent) %>%

DT::datatable(rownames = FALSE,
              options = list(dom = 't',
                             pageLength = nrow(outs),
```

```
scrollX = TRUE))
```

Summary

So, as to whether the loss of the center will result in Republican seats, or a clearer path for more progressive politics with an MLG-led New Mexico – we shall see. Regardless, some new faces come 2021 in the State Senate.

And, hopefully some useful methods presented here, reproducible to investigating political ideology in other state legislatures....