UNvotes

Hanh Nguyen

6/12/2017

Erik Voeten "Data and Analyses of Voting in the UN General Assembly" Routledge Handbook of International Organization, edited by Bob Reinalda (published May 27, 2013) The package contains three datasets. 1. un_votes 2. un_roll_calls 3. un_roll_call_issues un_roll_calls: each row is a country-vote pair rcid = roll call id - one round of voting Details about the datasets can be found here https://github.com/dgrtwo/unvotes or by ??unvotes

We start with the first dataset **un_votes**, which is the history of each country's vote.
By running *unique(un_votes$vote)*, we see that the **vote** column has three values: *yes, no, & abstain.*

A summary of the votes within each year.

```
df = merge(x=un_votes, y=un_roll_calls, by="rcid", all.x=TRUE)
head(df)
```

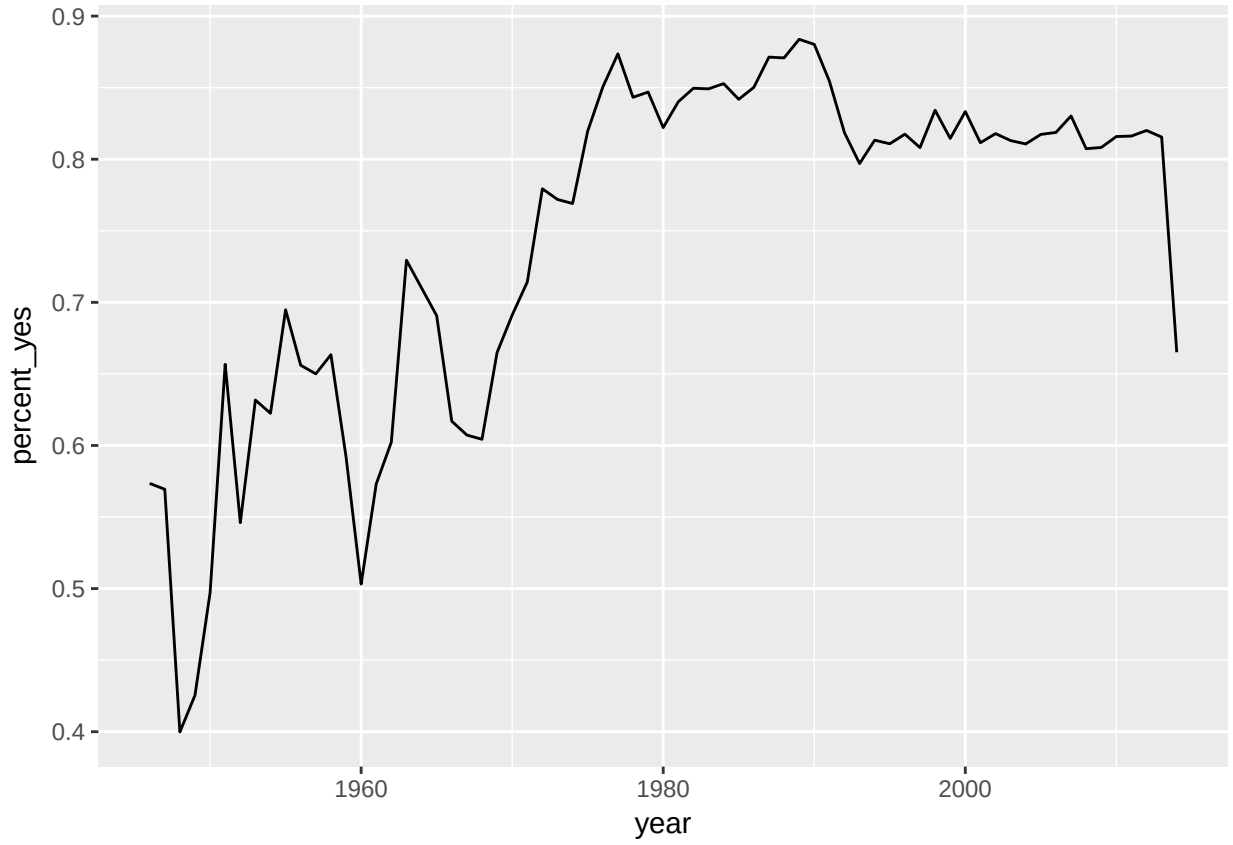
```
##   rcid     country    vote session importantvote       date  unres amend
## 1    3       Egypt abstain       1             0 1946-01-01 R/1/66     1
## 2    3    Honduras     yes       1             0 1946-01-01 R/1/66     1
## 3    3  Costa Rica     yes       1             0 1946-01-01 R/1/66     1
## 4    3 El Salvador     yes       1             0 1946-01-01 R/1/66     1
## 5    3      France      no       1             0 1946-01-01 R/1/66     1
## 6    3     Uruguay     yes       1             0 1946-01-01 R/1/66     1
##   para                          short
## 1    0 AMENDMENTS, RULES OF PROCEDURE
## 2    0 AMENDMENTS, RULES OF PROCEDURE
## 3    0 AMENDMENTS, RULES OF PROCEDURE
## 4    0 AMENDMENTS, RULES OF PROCEDURE
## 5    0 AMENDMENTS, RULES OF PROCEDURE
## 6    0 AMENDMENTS, RULES OF PROCEDURE
##
## 1 TO ADOPT A CUBAN AMENDMENT TO THE UK PROPOSAL REFERRING THE PROVISIONAL RULES OF PROCEDURE AND ANY
## 2 TO ADOPT A CUBAN AMENDMENT TO THE UK PROPOSAL REFERRING THE PROVISIONAL RULES OF PROCEDURE AND ANY
## 3 TO ADOPT A CUBAN AMENDMENT TO THE UK PROPOSAL REFERRING THE PROVISIONAL RULES OF PROCEDURE AND ANY
## 4 TO ADOPT A CUBAN AMENDMENT TO THE UK PROPOSAL REFERRING THE PROVISIONAL RULES OF PROCEDURE AND ANY
## 5 TO ADOPT A CUBAN AMENDMENT TO THE UK PROPOSAL REFERRING THE PROVISIONAL RULES OF PROCEDURE AND ANY
## 6 TO ADOPT A CUBAN AMENDMENT TO THE UK PROPOSAL REFERRING THE PROVISIONAL RULES OF PROCEDURE AND ANY
```

```
df$year <- as.numeric(format(df$date,"%Y"))
by_year = df %>%
  group_by(year) %>%
  summarize(total=n(), percent_yes = mean(vote=="yes"))
by_year
```

```
## # A tibble: 68 × 3
##     year total percent_yes
##    <dbl> <int>       <dbl>
## 1   1946  2143   0.5734951
## 2   1947  2039   0.5693968
## 3   1948  3454   0.3998263
## 4   1949  5700   0.4254386
```

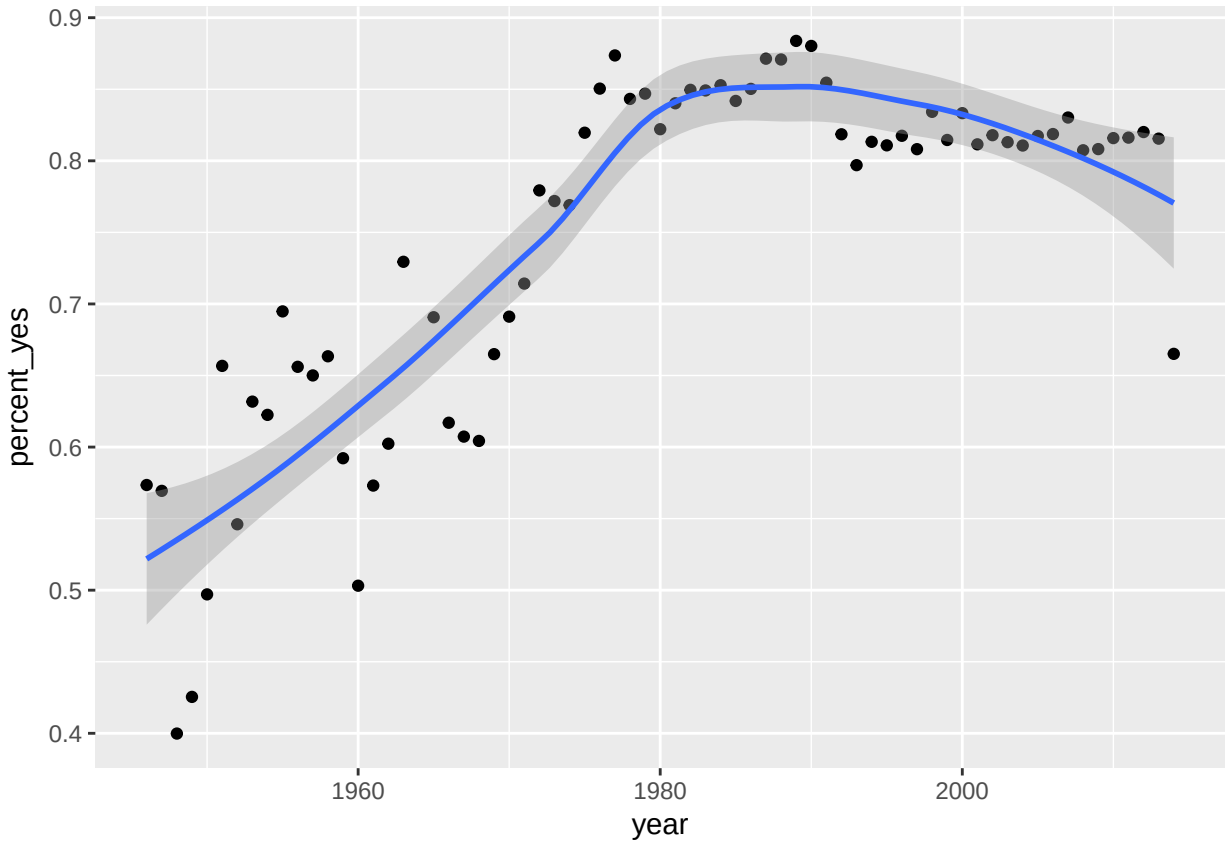
```
## 5   1950  2911   0.4970800
## 6   1951   402   0.6567164
## 7   1952  4082   0.5460559
## 8   1953  1537   0.6317502
## 9   1954  1788   0.6224832
## 10  1955  2169   0.6947902
## # ... with 58 more rows
```

```
ggplot(by_year, aes(year, percent_yes)) +
  geom_line()
```




```
ggplot(by_year, aes(year, percent_yes)) +
  geom_point() +
  geom_smooth()
```

```
## `geom_smooth()` using method = 'loess'
```



A summary of the votes within each country, which shows voting patterns among countries.

```
by_country = df %>%
  group_by(country) %>%
  summarize(total=n(), percent_yes = mean(vote=="yes"))
by_country
```

```
## # A tibble: 200 × 3
##                country total percent_yes
##                  <chr> <int>       <dbl>
## 1          Afghanistan  4824   0.8381012
## 2              Albania  3363   0.7204877
## 3              Algeria  4374   0.8978052
## 4              Andorra  1410   0.6510638
## 5               Angola  2950   0.9223729
## 6  Antigua and Barbuda  2521   0.9170964
## 7            Argentina  5207   0.7743422
## 8              Armenia  1479   0.7592968
## 9            Australia  5245   0.5542421
## 10             Austria  4786   0.6320518
## # ... with 190 more rows
```

Looking at countries that voted yes most and least

```
arrange(by_country, percent_yes)
```

```
## # A tibble: 200 × 3
##                            country total percent_yes
##                              <chr> <int>       <dbl>
```

```
## 1                         Zanzibar     2   0.0000000
## 2                   United States  5237   0.2850869
## 3                           Palau   777   0.3063063
## 4                          Israel  4790   0.3503132
## 5      Federal Republic of Germany  2151   0.3984193
## 6  Micronesia, Federated States of  1341   0.4131245
## 7                  United Kingdom  5218   0.4269835
## 8                          France  5171   0.4320248
## 9                Marshall Islands  1468   0.4788828
## 10                         Belgium  5238   0.4925544
## # ... with 190 more rows
```

```
arrange(by_country, desc(percent_yes))
```

```
## # A tibble: 200 × 3
##                   country total percent_yes
##                     <chr> <int>       <dbl>
## 1              Seychelles  1698   0.9770318
## 2             Timor-Leste   697   0.9670014
## 3  Sao Tome and Principe  2329   0.9665092
## 4                Djibouti  3193   0.9564673
## 5           Guinea-Bissau  2933   0.9546539
## 6                 Comoros  2435   0.9462012
## 7              Cabo Verde  3153   0.9454488
## 8              Mozambique  3306   0.9431337
## 9                   Yemen  1527   0.9423707
## 10                Zimbabwe  2766   0.9421547
## # ... with 190 more rows
```

The country that voted least frequently, Zanzibar, had only 2 votes in the entire dataset, thus it's hard to conclude anything. We will exclude countries with fewer than 100 votes.

```
by_country %>%
  arrange(percent_yes) %>%
  filter(total >= 100)
```

```
## # A tibble: 197 × 3
##                            country total percent_yes
##                              <chr> <int>       <dbl>
## 1                    United States  5237   0.2850869
## 2                            Palau   777   0.3063063
## 3                           Israel  4790   0.3503132
## 4      Federal Republic of Germany  2151   0.3984193
## 5  Micronesia, Federated States of  1341   0.4131245
## 6                    United Kingdom  5218   0.4269835
## 7                           France  5171   0.4320248
## 8                 Marshall Islands  1468   0.4788828
## 9                           Belgium  5238   0.4925544
## 10                       Luxembourg  5169   0.5105436
## # ... with 187 more rows
```

summarize by both year and country, constructing a dataset that shows what fraction of the time each country votes "yes" in each year.

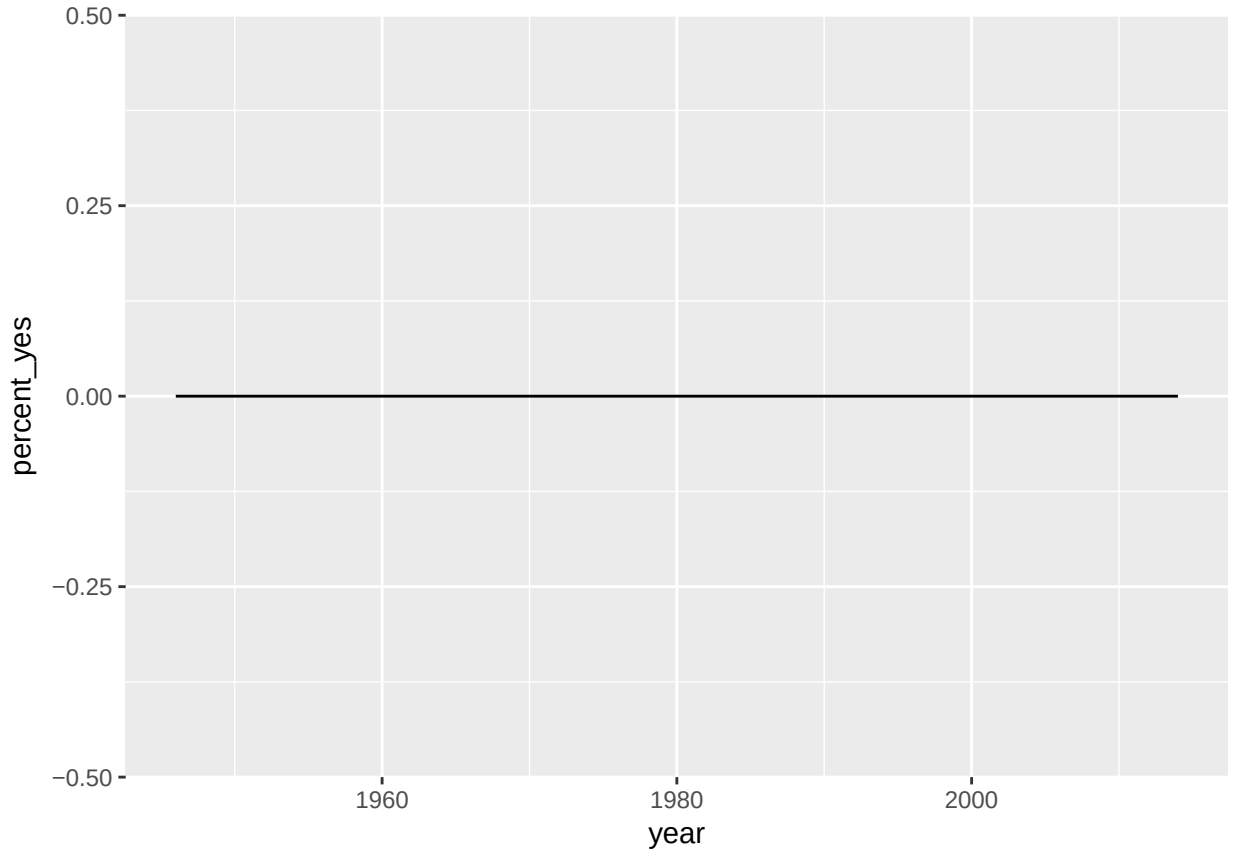
```
by_year_country = df %>%
  group_by(year, country) %>%
  summarize(total = n(),
```

```
            percent_yes = mean(vote == 1))
by_year_country
```

```
## Source: local data frame [9,496 x 4]
## Groups: year [?]
##
##     year                         country total percent_yes
##    <dbl>                           <chr> <int>       <dbl>
## 1   1946                     Afghanistan    17           0
## 2   1946                       Argentina    43           0
## 3   1946                       Australia    43           0
## 4   1946                         Belarus    43           0
## 5   1946                         Belgium    43           0
## 6   1946 Bolivia, Plurinational State of    43           0
## 7   1946                          Brazil    43           0
## 8   1946                          Canada    42           0
## 9   1946                           Chile    43           0
## 10  1946                        Colombia    42           0
## # ... with 9,486 more rows
```

Create a filtered version: US_by_year

```
US_by_year = by_year_country %>%
  filter(country=="United States")
ggplot(US_by_year, aes(x=year,y=percent_yes)) +
  geom_line()
```



Plotting just one country at a time is interesting, but you really want to compare trends between countries.

For example, suppose you want to compare voting trends for the United States, the UK, France, and China. The %in% operator!

```
countries <- c("United States", "United Kingdom",
               "France", "China")
filtered_4_countries = by_year_country %>%
  filter(country %in% countries)
ggplot(filtered_4_countries, aes(x=year,y=percent_yes,color=country)) +geom_line()
```