CIRI Dataset Exercise

Lotem Bassan

2/13/2019

CIRI Human Rights Data Project

In this short exercise I will explore some of the values collected in the CIRI Human Rights Data Project. This dataset is one of the most prominent and commonly used datasets when testing human rights violations in IR. It contains standards-based quantitative information on government respect for 15 internationally recognized human rights for 202 countries, annually from 1981-2011

I will be looking at three different human rights violations, embodied in a number of International treaties: Political killings or unlawful deprivation of life, torture and abuse and freedom of speech. Variables are coded from 0 to 2, where 0 indicates that the norm is frequently violated (ie. frequent political killings and torture or complete government censorship of media), 1 indicates occasional violations and 2 indicates no reported violations.

I begin by cleaning the data:

```
#Political Killings:
CIRI$KILL[CIRI$KILL==-999] <- NA
CIRI$KILL[CIRI$KILL==-77] <- NA
CIRI$KILL[CIRI$KILL==-66] <- NA
#Torture:
CIRI$TORT[CIRI$TORT==-999] <- NA
CIRI$TORT[CIRI$TORT==-77] <- NA
CIRI$TORT[CIRI$TORT==-66] <- NA
#Freedom of speech:
CIRI$SPEECH[CIRI$SPEECH==-999] <- NA
CIRI$SPEECH[CIRI$SPEECH==-77] <- NA
CIRI$SPEECH[CIRI$SPEECH==-66] <- NA

# a little shorter with dplyr
CIRI %<>%
  mutate(KILL = ifelse(KILL %in% c(-999,-77,-66), NA, KILL)) %>%
  mutate(TORT = ifelse(TORT %in% c(-999,-77,-66), NA, TORT)) %>%
  mutate(SPEECH = ifelse(SPEECH %in% c(-999,-77,-66), NA, SPEECH))
```

Now, we can use the summarise() function to observe some of the variation in the variables. As mentioned in the previous paragraph, larger numbers will indicate less violations, while smaller numbers indicate frequent violations.

```
CIRI%>%
  filter(!is.na(KILL))%>%
  group_by(CTRY)%>%
  summarise(count_KILL=sum(KILL)) %>%
  complete(CTRY) ## complete() adds missing potential combinations of values. I don't think it is doing
```

```
## # A tibble: 202 x 2
##    CTRY                count_KILL
##    <fct>                    <int>
##  1 Afghanistan                  5
##  2 Albania                     32
```

```
##  3 Algeria                     31
##  4 Andorra                     20
##  5 Angola                      11
##  6 Antigua and Barbuda         20
##  7 Argentina                   30
##  8 Armenia                     33
##  9 Australia                   61
## 10 Austria                     59
## # ... with 192 more rows
```

Unsurprisingly, we learn that undemocratic countries (such as Afghanistan) receive low scores and do not conform with the international norms, while democratic countries such as Austria, receive higher scores.

We observe similar trends with torture and speech:

```
#Torture
CIRI%>%
  filter(!is.na(TORT))%>%
  group_by(CTRY)%>%
  summarise(count_TORT=sum(TORT))%>%
  complete(CTRY)
```

```
## # A tibble: 202 x 2
##    CTRY                count_TORT
##    <fct>                    <int>
##  1 Afghanistan                  2
##  2 Albania                     14
##  3 Algeria                     20
##  4 Andorra                     20
##  5 Angola                      12
##  6 Antigua and Barbuda         10
##  7 Argentina                   23
##  8 Armenia                      6
##  9 Australia                   38
## 10 Austria                     39
## # ... with 192 more rows
```

```
#Speech
CIRI%>%
  filter(!is.na(SPEECH))%>%
  group_by(CTRY)%>%
  summarise(count_SPEECH=sum(SPEECH))%>%
  complete(CTRY)
```

```
## # A tibble: 202 x 2
##    CTRY                count_SPEECH
##    <fct>                      <int>
##  1 Afghanistan                    5
##  2 Albania                       25
##  3 Algeria                       10
##  4 Andorra                       19
##  5 Angola                        10
##  6 Antigua and Barbuda           12
##  7 Argentina                     44
##  8 Armenia                       17
##  9 Australia                     60
## 10 Austria                       42
```

```
## # ... with 192 more rows
```

Indeed, we observe a similar trend. We may want to test in the future, using other datasets, whether there are other variations worth exploring (for example, why do some non-democracies or new democracies violate these rights more frequently than others?, one explanation is the treaties and organizations they are members of, others is whether or not they are frequently targeted by NGOs).

```
d  <- CIRI

# When we group by country, we get groups of years 1981-2011, but not all countries are measured every

d %<>%
  group_by(CTRY)%>%
  mutate(mean_KILL = mean(KILL, na.rm = T),
         mean_TORT = mean(TORT, na.rm = T),
         mean_SPEECH = mean(SPEECH, na.rm = T) )

glimpse(d)
```

```
## Observations: 6,262
## Variables: 31
## Groups: CTRY [202]
## $ CTRY        <fct> Afghanistan, Afghanistan, Afghanistan, Afghanistan...
## $ YEAR        <int> 1981, 1982, 1983, 1984, 1985, 1986, 1987, 1988, 19...
## $ CIRI        <int> 101, 101, 101, 101, 101, 101, 101, 101, 101, 101, ...
## $ COW         <int> 700, 700, 700, 700, 700, 700, 700, 700, 700, 700, ...
## $ POLITY      <int> 700, 700, 700, 700, 700, 700, 700, 700, 700, 700, ...
## $ UNCTRY      <int> 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4,...
## $ UNREG       <int> 142, 142, 142, 142, 142, 142, 142, 142, 142, 142, ...
## $ UNSUBREG    <int> 62, 62, 62, 62, 62, 62, 62, 62, 62, 62, 62, 62, 62...
## $ PHYSINT     <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, NA, NA, NA, NA, 0...
## $ DISAP       <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, -77, -77, -77, -7...
## $ KILL        <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, NA, NA, NA, NA, 0...
## $ POLPRIS     <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, -77, -77, -77, -7...
## $ TORT        <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, NA, NA, NA, NA, 0...
## $ OLD_EMPINX  <int> 0, 2, 0, 1, 0, 0, 1, 2, 2, 2, 0, NA, NA, NA, NA, 3...
## $ NEW_EMPINX  <int> 2, 1, 0, 1, 0, 1, 3, 2, 3, 2, 3, NA, NA, NA, NA, 3...
## $ ASSN        <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, -77, -77, -77, -7...
## $ FORMOV      <int> 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 2, -77, -77, -77, -7...
## $ DOMMOV      <int> 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, -77, -77, -77, -7...
## $ OLD_MOVE    <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, -77, -77, -77, -7...
## $ SPEECH      <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, NA, NA, NA, NA, 1...
## $ ELECSD      <int> 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, -77, -77, -77, -7...
## $ OLD_RELFRE  <int> 0, 1, 0, 0, 0, 0, 0, 1, 1, 1, 0, -77, -77, -77, -7...
## $ NEW_RELFRE  <int> 1, 1, 0, 0, 0, 1, 2, 2, 2, 1, 1, -77, -77, -77, -7...
## $ WORKER      <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, -77, -77, -77, -7...
## $ WECON       <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, -77, -77, -77, -7...
## $ WOPOL       <int> 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, -77, -77, -77, -7...
## $ WOSOC       <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, -77, -77, -77, -7...
## $ INJUD       <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, -77, -77, -77, -7...
## $ mean_KILL   <dbl> 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, ...
## $ mean_TORT   <dbl> 0.08, 0.08, 0.08, 0.08, 0.08, 0.08, 0.08, 0.08, 0....
## $ mean_SPEECH <dbl> 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, ...
```

```
# make a smaller data frame with one observation per country
d1 <- d %>%
```

```
  select(CTRY, POLITY, mean_KILL, mean_TORT, mean_SPEECH) %>%
  distinct()

glimpse(d1)
```

```
## Observations: 204
## Variables: 5
## Groups: CTRY [202]
## $ CTRY        <fct> Afghanistan, Albania, Algeria, Andorra, Angola, An...
## $ POLITY      <int> 700, 339, 615, 232, 540, 58, 160, 371, 900, 305, 3...
## $ mean_KILL   <dbl> 0.2000000, 1.2307692, 1.0000000, 2.0000000, 0.3666...
## $ mean_TORT   <dbl> 0.08000000, 0.45161290, 0.64516129, 2.00000000, 0....
## $ mean_SPEECH <dbl> 0.2000000, 0.8064516, 0.3225806, 1.9000000, 0.3333...
```

The correlation between Polity score and freedom from speech violations is -0.325773. The correlation between Polity score and freedom from torture violations is 0.0055049 The correlation between Polity score and freedom from political killings is -0.0454884

```
d1 %>%
  ggplot() +
  aes(x = mean_SPEECH,
      y = mean_TORT,
      color = mean_KILL,
      size = POLITY,
      label = CTRY) +
  geom_point(alpha = .3) +
  geom_text(check_overlap = T,
            size = 4)+
  labs(x = "Freedom of Speech",
       y = "Freedom from Torture",
       color = "Freedom from Political Killings")
```