Some {peacesciencer} Exercises and False Causes of War

Problem 1: Mountainous Piles of Data and War

In a classic study on the correlates of civil war, Fearon and Laitin (2003) showed that, among many factors, mountainous terrain is a positive predictor of conflict. The argument is that having a place to hide from government forces is a nice boost for rebels. Therefore, countries with rougher terrain (e.g., a lot of mountains) will have a bigger problem identifying and disbanding/eliminating insurgent groups. Let's test this claim using `{peacesciencer}` constructed data. Produce a state-year dataset then populate with UCDP armed conflict data and information about rugged terrain (there are functions for adding each set of variables to the data). Then visualize the relationship between rugged terrain and the likelihood of civil war onset.

Note: there are two measures of ruggedness in the data. You can write `?rugged` to learn more about them. Make sure to check whether you get the same answer to your research question with each measure.

Problem 2: Who's going to stop me?

`{peacesciencer}` gives you the ability to look within countries to the characteristics of individual leaders. Usually, when one country chooses to attack another, the chief executive is responsible for this decision. So it's worth asking whether leader characteristics, such as prior military experience, gender, or willingness to use force might play a role in conflict initiation in international politics.

To answer this question, create a leader-year dataset from 1875 to 2010 and set the option `standardize = "cow"` so that you can populate the data with certain country-level variables in addition to leader-specific variables. Notice that right off the bat, the leader-year dataset already includes some information about leaders like gender, age, and years in office.

Once you've made the base data, you'll want to use the following functions to populate it with the relevant variables for your analysis:

- `add_gml_mids()`: adds information about Militarized Interstate Disputes (MIDs) based on Gibler and Miller's extension to the original CoW MID dataset. You'll use the variable `gmlmidonset_init` as your outcome variable. It takes the value 1 if a leader started a dispute with another country in a given year. It takes the value 0 otherwise.
- `add_lead()`: adds a bunch of extra information about the leader of a country like education level, whether they have military service experience, actual combat experience, whether they were ever part of a rebel movement, whether they experienced wins or losses in past experiences in the military or as a rebel, and whether they're in good physical health and good mental health.
- `add_lwuf()`: adds a set of variables that some political scientists named Carter and Smith created using a Bayesian latent variable framework to quantify a leader's willingness to use force based on their characteristics and past experiences. There are a number of measures that you could use. The creators recommend using `theta2_mean` because they found it to be the best performing version of the measure.
- `add_nmc()`: adds data about national military capabilities. You'll want to use the `cinc` measure, which is an index constructed from other variables in the data to reflect a country's overall relative level of power compared to other countries in the international system.
- `add_democracy()`: adds three different measures of quality of democracy. I'd recommend using `xm_qudsest` since it has the best coverage for all countries in the data.

Once your data is constructed, try out some regression analyses using the `logitmfx()` function from the `{mfx}` package similar to the one I performed in section 4.3 in the lecture notes. Try out some different combinations of variables to see how including or dropping certain variables from the model changes the conclusions you'd draw about the role of leader characteristics in explaining why a country would choose to start a militarized dispute with another (**Note**: You might want to see what happens when you switch between using individual leader characteristics versus the single measure of leader willingness to use force).

Problem 3: Dangerous Dyads?

In section 4.3 of the notes, I showed you how to replicate the "dangerous dyads" study by Bremer. I've provided the code below for reference. I'd like you to see what happens when you replace the MID onset measure with, instead, a count of the yearly number of Militarized Interstate Events (MIEs) that take place between countries. To do this, you'll need to read in the MIE data I had you upload and save in your R project. Then you'll need to summarize it to the level of dyad-years from its current level of dyad-days. Once you do this, you can left-join it to the data you created below. When you perform your regression analysis (like the one I showed you in the notes), use the function `poissonmfx()` from the `{mfx}` package since

your outcome will be a count of the number of MIEs that take place between countries over the course of a year.

Here's the original code:

```
## create base data, and pipe (|>) to next function
create_dyadyears(directed = FALSE, subset_years = c(1816:2010)) |>
  ## subset data to politically relevant dyads (PRDs)
  ## pipe to next function
  filter_prd() |>
  ## add conflict information from GML-MID data
  ## pipe to next function
  add_gml_mids(keep = NULL) |>
  ## add peace years ("spells")
  ## pipe to next function
  add_spells() |>
  ## add capabilities data
  ## pipe to next function
  add_nmc() |>
  ## add some estimates about democracy for each state
  ## pipe to next function
  add_democracy() |>
  ## add information about alliance commitments in dyad-year
  add_cow_alliance() |>
  ## finish with information about population and GDP/SDP
  ## and then assign to object, called, minimally, 'Data'
  add_sdp_gdp() -> Data

## Update some variables for the analysis
Data <- Data |>
  ## Create dummies for contiguity and major power dyads
  mutate(landcontig = ifelse(conttype == 1, 1, 0)) |>
  mutate(cowmajdyad = ifelse(cowmaj1 == 1 | cowmaj2 == 1, 1, 0)) |>
  ## Create estimate of militarization as milper/tpop
  ## Then make military weak-link measure
  mutate(milit1 = milper1/tpop1,
         milit2 = milper2/tpop2,
         minmilit = ifelse(milit1 > milit2,
                           milit2, milit1)) |>
  ## Create CINC proportion (lower over higher)
  mutate(cincprop = ifelse(cinc1 > cinc2,
                           cinc2/cinc1, cinc1/cinc2)) |>
```

```r
  ## Create dem. weak-link specification using Quick UDS data
  mutate(mindemest = ifelse(xm_qudsest1 > xm_qudsest2,
                            xm_qudsest2, xm_qudsest1)) |>
  ## Create "weak-link" measure of jointly advanced economies
  mutate(minwbgdppc = ifelse(wbgdppc2011est1 > wbgdppc2011est2,
                             wbgdppc2011est2, wbgdppc2011est1))

## perform the analysis
 # estimate the model
model <- logitmfx(
  gmlmidonset ~ landcontig + cincprop + cowmajdyad + cow_defense +
                mindemest + minwbgdppc + minmilit +
                gmlmidspell + I(gmlmidspell^2) + I(gmlmidspell^3),
  data = Data |> mutate(dyad = paste0(ccode1, ccode2)),
  robust = T,
  clustervar1 = "dyad"
)

 # report the regression results with kbl()
library(kableExtra)
broom::tidy(model) |>
  dplyr::select(-atmean) |>
  kbl(
    caption = "Dangerous Dyads Analysis",
    digits = 3
  )
```