National Economies

Stat 341 – Spring 2017

GDP and ruggedness

> Bad geography tends to be related to bad economies outside of Africa, but African economies seem immune to bad geography. (McElreath 211)

How do we fit models to assess this claim?

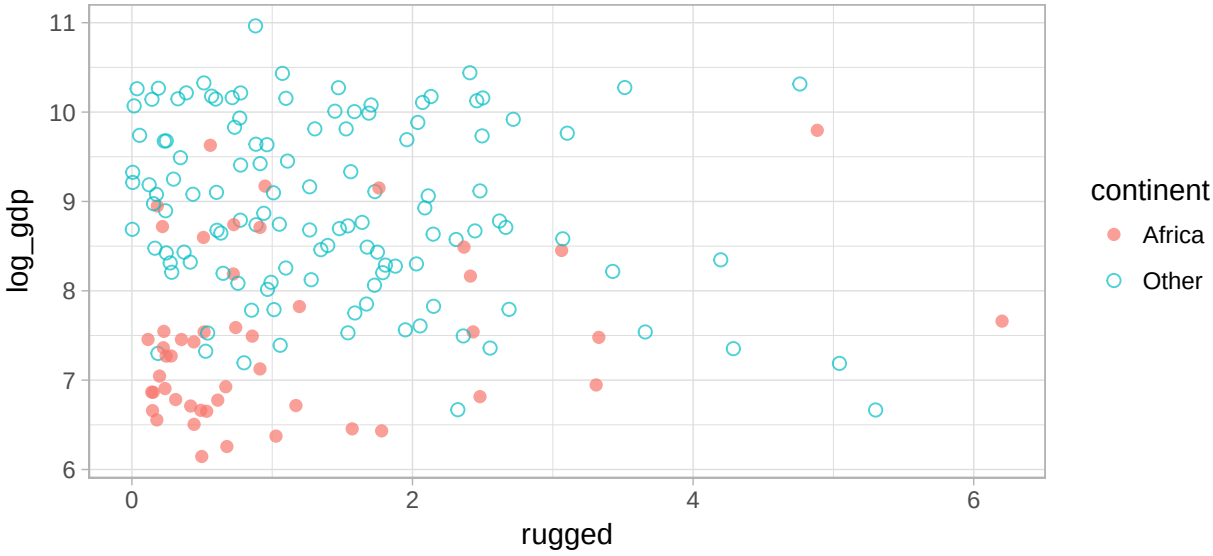
The Data

The data used for this are in the `rugged` data set. Naming a data frame after one of its variables is not a very good idea. Let's rename the data and

- add in log of 2000 GDP,
- add `continent`, which is handy for plotting later
- remove countries that don't have 2000 GDP data (since they won't be used in the analysis below)

```
library(rethinking)
data(rugged)

Nations <-
  rugged %>%
  mutate(
    log_gdp = log(rgdppc_2000),
    continent = ifelse(cont_africa, "Africa", "Other")
  ) %>%
  filter(! is.na(log_gdp))
```



Approach 1: A bad approach

Here's a bad idea that we will improve upon:

- split the data into two portions (African and non-African countries)
- fit a model with log GDP as outcome and ruggedness as predictor in each portion

What's bad about this idea?

Approach 2: This also isn't quite right

What is wrong with this model?

```
m7.4 <- map(
  alist(
    log_gdp ~ dnorm(mu, sigma),
    mu <- a + bR * rugged + bA * cont_africa,
    a ~ dnorm(8, 100),
    bR ~ dnorm(0, 1),
    bA ~ dnorm(0, 1),
    sigma ~ dunif(0, 10)
  ),
  data = Nations
)
```

Approach 3: Fixing the problem

Describe a single model that allows each country group to have its own intercept *and slope*.

- Do this as many ways as you can think of.
- Are your different ways just descriptions of the same thing or are they actually different models?
- Which formulation(s) do you like better? Why?

Comparing Appraoches

What things can you do to compare the approaches?

- You don't have to actually do them, just make a list of what you would do and what you would look for/at.
- For each one, jot down the names of the R funcitons you would use.
- Once your list is complete, if you have time, go ahead and do the things on your list.