

R Notebook

```
library(rvest)
```

```
## Loading required package: xml2
```

```
library(magrittr)
```

Retrieve population of countries

The last piece of information I want for the analysis is the population size of each country. There are obviously many sources for such information. I could have used Jenny Bryan's Gapminder package, for example, but a better source I found was from GeoNames since they have the 2-letter code for each country, so it'll be easier to merge with our data.

Once again, I will use Hadley Wickham's rvest package, this time to scrape a table of country populations from GeoNames. I'm still excited by how easy this is!

```
countryPopUrl <- "http://www.geonames.org/countries/"
countryPopulationPage <- read_html(countryPopUrl)
countryPop <-
  countryPopulationPage %>%
    html_nodes("table") %>%
    extract2(2) %>%
    html_table

knitr::kable(head(countryPop))
```

ISO-3166alpha2	ISO-3166alpha3	ISO-3166numeric	fips	Country	Capital	Area in km ²	P
AD	AND	20	AN	Andorra	Andorra la Vella	468.0	8
AE	ARE	784	AE	United Arab Emirates	Abu Dhabi	82,880.0	4
AF	AFG	4	AF	Afghanistan	Kabul	647,500.0	29
AG	ATG	28	AC	Antigua and Barbuda	St. John's	443.0	80
AI	AIA	660	AV	Anguilla	The Valley	102.0	13
AL	ALB	8	AL	Albania	Tirana	28,748.0	2