

WDI - Download

Datasets

François Geerolf

Contents

Preamble	1
Steps	1
Download	1
Unzip	1
Main CSV	1
Auxiliary CSV	2
Saving as RData	2
Delete Downloaded Files	2
Computing Environment	2

Preamble

```
rm(list = ls())
pklist <- c("tidyverse", "curl")
source("https://fgeerolf.github.io/datasets/load-packages.R")
```

Steps

The World Development Indicators are available interactively here: <http://databank.worldbank.org/data/source/world-development-indicators>.

Download

```
"http://databank.worldbank.org/data/download/WDI_csv.zip" %>%
  curl_download(url = ., destfile = "WDI_csv.zip", quiet = FALSE, mode = "wb")
```

Unzip

```
unzip("WDI_csv.zip")
unlink("WDI_csv.zip")
```

Main CSV

```
WDI <- read.csv("WDIData.csv") %>%
  gather(year, value, -Country.Name, -Country.Code, -Indicator.Name, -Indicator.Code) %>%
  mutate(year = year %>% substr(2, 5) %>% as.numeric %>% as.factor) %>%
  filter(!is.na(value))
```

Auxiliary CSV

```
WDI.series <- read.csv('WDISeries.csv')

WDI.country <- read.csv('WDICountry.csv') %>%
  select(-X)

WDI.variable.nobs <- WDI %>%
  group_by(Indicator.Name, Indicator.Code) %>%
  summarise(nobs = sum(!is.na(value))) %>%
  ungroup %>%
  select(Indicator.Code, Indicator.Name, nobs) %>%
  mutate(Indicator.Code = Indicator.Code %>% paste %>% as.character,
         Indicator.Name = Indicator.Name %>% paste %>% as.character) %>%
  arrange(Indicator.Code) %>%
  unique
```

Saving as RData

```
save(WDI, file = "WDI.RData")
save(WDI.series, file = "WDI.series.RData")
save(WDI.country, file = "WDI.country.RData")
save(WDI.variable.nobs, file = "WDI.variable.nobs.RData")
```

Delete Downloaded Files

```
for (file in list.files(include.dirs = FALSE, recursive = TRUE, pattern = "\\\\.csv$")) unlink(file)
```

Computing Environment

```
Sys.time()
```

```
## [1] "2018-09-24 15:46:19 PDT"
```

```
sessionInfo()
```

```
## R version 3.5.1 (2018-07-02)
```

```
## Platform: x86_64-apple-darwin15.6.0 (64-bit)
```

```
## Running under: macOS High Sierra 10.13.6
```

```
##
```

```
## Matrix products: default
```

```
## BLAS: /Library/Frameworks/R.framework/Versions/3.5/Resources/lib/libRblas.0.dylib
```

```
## LAPACK: /Library/Frameworks/R.framework/Versions/3.5/Resources/lib/libLapack.dylib
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## attached base packages:
## [1] stats      graphics  grDevices  utils      datasets  methods   base
##
## other attached packages:
## [1] bindrcpp_0.2.2  curl_3.2      forcats_0.3.0  stringr_1.3.1
## [5] dplyr_0.7.6     purrr_0.2.5   readr_1.1.1    tidyr_0.8.1
## [9] tibble_1.4.2    ggplot2_3.0.0  tidyverse_1.2.1
##
## loaded via a namespace (and not attached):
## [1] Rcpp_0.12.18    cellranger_1.1.0 pillar_1.3.0    compiler_3.5.1
## [5] plyr_1.8.4      bindr_0.1.1     tools_3.5.1     digest_0.6.15
## [9] lubridate_1.7.4 jsonlite_1.5     evaluate_0.11    nlme_3.1-137
## [13] gtable_0.2.0    lattice_0.20-35 pkgconfig_2.0.2  rlang_0.2.2
## [17] cli_1.0.0       rstudioapi_0.7  yaml_2.2.0      haven_1.1.2
## [21] withr_2.1.2     xml2_1.2.0      httr_1.3.1      knitr_1.20
## [25] hms_0.4.2       rprojroot_1.3-2 grid_3.5.1       tidyselect_0.2.4
## [29] glue_1.3.0      R6_2.2.2        readxl_1.1.0    rmarkdown_1.10
## [33] modelr_0.1.2    magrittr_1.5     backports_1.1.2  scales_1.0.0
## [37] htmltools_0.3.6 rvest_0.3.2     assertthat_0.2.0 colorspace_1.3-2
## [41] stringi_1.2.4   lazyeval_0.2.1  munsell_0.5.0   broom_0.5.0
## [45] crayon_1.3.4
```