

BLS - LAUS - Download

Datasets

François Geerolf

Introduction

```
rm(list = ls())
pklist <- c("curl", "tidyverse", "rvest")
source("https://fgeerolf.github.io/datasets/load-packages.R")
options(tibble.print_max = 100)
```

Introduction

An updated version of this note is available here: <https://fgeerolf.github.io/datasets/bls-la-download.pdf>
The documentation for the LAUS is available here: <https://www.bls.gov/lau/> The flat data files of the LAUS are: <https://download.bls.gov/pub/time.series/la/> Link: <https://download.bls.gov/pub/time.series/la/>

```
url <- "https://download.bls.gov/pub/time.series/la/"
```

Scrapping Data

Elements of the scrapping data:

```
read_html(url) %>%
  html_nodes("a") %>%
  html_text(trim = TRUE) %>%
  as.data.frame %>%
  rename(X0 = ".") %>%
  as.tibble
```

```
# # A tibble: 82 x 1
#   X0
#   <fct>
# 1 [To Parent Directory]
# 2 la.area
# 3 la.area_type
# 4 la.areamaps
# 5 la.contacts
# 6 la.data.0.CurrentU00-04
# 7 la.data.0.CurrentU05-09
# 8 la.data.0.CurrentU10-14
# 9 la.data.0.CurrentU15-19
# 10 la.data.0.CurrentU90-94
# 11 la.data.0.CurrentU95-99
# 12 la.data.1.CurrentS
# 13 la.data.10.Arkansas
# 14 la.data.11.California
```

```
# 15 la.data.12.Colorado
# 16 la.data.13.Connecticut
# 17 la.data.14.Delaware
# 18 la.data.15.DC
# 19 la.data.16.Florida
# 20 la.data.17.Georgia
# 21 la.data.18.Hawaii
# 22 la.data.19.Idaho
# 23 la.data.2.AllStatesU
# 24 la.data.20.Illinois
# 25 la.data.21.Indiana
# 26 la.data.22.Iowa
# 27 la.data.23.Kansas
# 28 la.data.24.Kentucky
# 29 la.data.25.Louisiana
# 30 la.data.26.Maine
# 31 la.data.27.Maryland
# 32 la.data.28.Massachusetts
# 33 la.data.29.Michigan
# 34 la.data.3.AllStatesS
# 35 la.data.30.Minnesota
# 36 la.data.31.Mississippi
# 37 la.data.32.Missouri
# 38 la.data.33.Montana
# 39 la.data.34.Nebraska
# 40 la.data.35.Nevada
# 41 la.data.36.NewHampshire
# 42 la.data.37.NewJersey
# 43 la.data.38.NewMexico
# 44 la.data.39.NewYork
# 45 la.data.4.RegionDivisionU
# 46 la.data.40.NorthCarolina
# 47 la.data.41.NorthDakota
# 48 la.data.42.Ohio
# 49 la.data.43.Oklahoma
# 50 la.data.44.Oregon
# 51 la.data.45.Pennsylvania
# 52 la.data.46.PuertoRico
# 53 la.data.47.RhodeIsland
# 54 la.data.48.SouthCarolina
# 55 la.data.49.SouthDakota
# 56 la.data.5.RegionDivisionS
# 57 la.data.50.Tennessee
# 58 la.data.51.Texas
# 59 la.data.52.Utah
# 60 la.data.53.Vermont
# 61 la.data.54.Virginia
# 62 la.data.56.Washington
# 63 la.data.57.WestVirginia
# 64 la.data.58.Wisconsin
# 65 la.data.59.Wyoming
# 66 la.data.60.Metro
# 67 la.data.61.Division
# 68 la.data.62.Micro
```

```

# 69 la.data.63.Combined
# 70 la.data.64.County
# 71 la.data.65.City
# 72 la.data.7.Alabama
# 73 la.data.8.Alaska
# 74 la.data.9.Arizona
# 75 la.footnote
# 76 la.map_info
# 77 la.measure
# 78 la.period
# 79 la.series
# 80 la.state_region_division
# 81 la.txt
# 82 maps

read_html(url) %>%
  str_match_all("<a href=\"(.*)\"") %>%
  as.data.frame %>%
    mutate(X2 = paste0("https://download.bls.gov", X2)) %>%
  as.tibble

# Warning in stri_match_all_regex(string, pattern, omit_no_match = TRUE,
# opts_regex = opts(pattern)): argument is not an atomic vector; coercing

# # A tibble: 82 x 2
#   X1                                X2
#   <fct>                            <chr>
# 1 "<a href=\"/pub/time.series/\""  https://download.bls.gov/pub/time.se~
# 2 "<a href=\"/pub/time.series/la/l~ https://download.bls.gov/pub/time.se~
# 3 "<a href=\"/pub/time.series/la/l~ https://download.bls.gov/pub/time.se~
# 4 "<a href=\"/pub/time.series/la/l~ https://download.bls.gov/pub/time.se~
# 5 "<a href=\"/pub/time.series/la/l~ https://download.bls.gov/pub/time.se~
# 6 "<a href=\"/pub/time.series/la/l~ https://download.bls.gov/pub/time.se~
# 7 "<a href=\"/pub/time.series/la/l~ https://download.bls.gov/pub/time.se~
# 8 "<a href=\"/pub/time.series/la/l~ https://download.bls.gov/pub/time.se~
# 9 "<a href=\"/pub/time.series/la/l~ https://download.bls.gov/pub/time.se~
# 10 "<a href=\"/pub/time.series/la/l~ https://download.bls.gov/pub/time.se~
# 11 "<a href=\"/pub/time.series/la/l~ https://download.bls.gov/pub/time.se~
# 12 "<a href=\"/pub/time.series/la/l~ https://download.bls.gov/pub/time.se~
# 13 "<a href=\"/pub/time.series/la/l~ https://download.bls.gov/pub/time.se~
# 14 "<a href=\"/pub/time.series/la/l~ https://download.bls.gov/pub/time.se~
# 15 "<a href=\"/pub/time.series/la/l~ https://download.bls.gov/pub/time.se~
# 16 "<a href=\"/pub/time.series/la/l~ https://download.bls.gov/pub/time.se~
# 17 "<a href=\"/pub/time.series/la/l~ https://download.bls.gov/pub/time.se~
# 18 "<a href=\"/pub/time.series/la/l~ https://download.bls.gov/pub/time.se~
# 19 "<a href=\"/pub/time.series/la/l~ https://download.bls.gov/pub/time.se~
# 20 "<a href=\"/pub/time.series/la/l~ https://download.bls.gov/pub/time.se~
# 21 "<a href=\"/pub/time.series/la/l~ https://download.bls.gov/pub/time.se~
# 22 "<a href=\"/pub/time.series/la/l~ https://download.bls.gov/pub/time.se~
# 23 "<a href=\"/pub/time.series/la/l~ https://download.bls.gov/pub/time.se~
# 24 "<a href=\"/pub/time.series/la/l~ https://download.bls.gov/pub/time.se~
# 25 "<a href=\"/pub/time.series/la/l~ https://download.bls.gov/pub/time.se~
# 26 "<a href=\"/pub/time.series/la/l~ https://download.bls.gov/pub/time.se~
# 27 "<a href=\"/pub/time.series/la/l~ https://download.bls.gov/pub/time.se~
# 28 "<a href=\"/pub/time.series/la/l~ https://download.bls.gov/pub/time.se~

```

4

```

datasets <- read_html(url) %>%
  html_nodes("a") %>%
  html_text(trim = TRUE) %>%
  as.data.frame %>%
  rename(X0 = ".") %>%
  cbind(read_html(url) %>%
    str_match_all("<a href=\"(.*)\"") %>%
    as.data.frame %>%
    mutate(X2 = paste0("https://download.bls.gov", X2))) %>%
  mutate_all(paste)

```

```

# Warning in stri_match_all_regex(string, pattern, omit_no_match = TRUE,
# opts_regex = opts(pattern)): argument is not an atomic vector; coercing

```

```

datasets %>%
  as.tibble

```

```

# # A tibble: 82 x 3
#   X0          X1          X2
#   <chr>      <chr>      <chr>
# 1 [To Parent D~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 2 la.area      "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 3 la.area_type "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 4 la.areamaps  "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 5 la.contacts  "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 6 la.data.0.Cu~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 7 la.data.0.Cu~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 8 la.data.0.Cu~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 9 la.data.0.Cu~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 10 la.data.0.Cu~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 11 la.data.0.Cu~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 12 la.data.1.Cu~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 13 la.data.10.A~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 14 la.data.11.C~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 15 la.data.12.C~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 16 la.data.13.C~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 17 la.data.14.D~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 18 la.data.15.DC "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 19 la.data.16.F~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 20 la.data.17.G~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 21 la.data.18.H~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 22 la.data.19.I~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 23 la.data.2.Al~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 24 la.data.20.I~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 25 la.data.21.I~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 26 la.data.22.I~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 27 la.data.23.K~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 28 la.data.24.K~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 29 la.data.25.L~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 30 la.data.26.M~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 31 la.data.27.M~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 32 la.data.28.M~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 33 la.data.29.M~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 34 la.data.3.Al~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 35 la.data.30.M~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~

```

```

# 36 la.data.31.M~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 37 la.data.32.M~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 38 la.data.33.M~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 39 la.data.34.N~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 40 la.data.35.N~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 41 la.data.36.N~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 42 la.data.37.N~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 43 la.data.38.N~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 44 la.data.39.N~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 45 la.data.4.Re~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 46 la.data.40.N~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 47 la.data.41.N~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 48 la.data.42.O~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 49 la.data.43.O~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 50 la.data.44.O~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 51 la.data.45.P~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 52 la.data.46.P~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 53 la.data.47.R~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 54 la.data.48.S~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 55 la.data.49.S~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 56 la.data.5.Re~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 57 la.data.50.T~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 58 la.data.51.T~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 59 la.data.52.U~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 60 la.data.53.V~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 61 la.data.54.V~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 62 la.data.56.W~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 63 la.data.57.W~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 64 la.data.58.W~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 65 la.data.59.W~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 66 la.data.60.M~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 67 la.data.61.D~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 68 la.data.62.M~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 69 la.data.63.C~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 70 la.data.64.C~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 71 la.data.65.C~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 72 la.data.7.Al~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 73 la.data.8.Al~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 74 la.data.9.Ar~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 75 la.footnote  "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 76 la.map_info  "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 77 la.measure   "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 78 la.period    "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 79 la.series    "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 80 la.state_reg~ "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 81 la.txt       "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~
# 82 maps        "<a href=\"/pub/time.seri~ https://download.bls.gov/pub/~

```

Downloading all data

```

for (i in 2:80){
  file <- datasets[i, "X0"]
}

```

```

cat("\nDownloading from BLS Website LA:", file)
assign(file, read.csv(datasets[i, "X2"], sep = "\t", row.names = NULL))
do.call(save, list(file, file = paste0(file, ".RData")))
}

```

```

#
# Downloading from BLS Website LA: la.area
# Downloading from BLS Website LA: la.area_type
# Downloading from BLS Website LA: la.areamaps
# Downloading from BLS Website LA: la.contacts
# Downloading from BLS Website LA: la.data.0.CurrentU00-04
# Downloading from BLS Website LA: la.data.0.CurrentU05-09
# Downloading from BLS Website LA: la.data.0.CurrentU10-14
# Downloading from BLS Website LA: la.data.0.CurrentU15-19
# Downloading from BLS Website LA: la.data.0.CurrentU90-94
# Downloading from BLS Website LA: la.data.0.CurrentU95-99
# Downloading from BLS Website LA: la.data.1.CurrentS
# Downloading from BLS Website LA: la.data.10.Arkansas
# Downloading from BLS Website LA: la.data.11.California
# Downloading from BLS Website LA: la.data.12.Colorado
# Downloading from BLS Website LA: la.data.13.Connecticut
# Downloading from BLS Website LA: la.data.14.Delaware
# Downloading from BLS Website LA: la.data.15.DC
# Downloading from BLS Website LA: la.data.16.Florida
# Downloading from BLS Website LA: la.data.17.Georgia
# Downloading from BLS Website LA: la.data.18.Hawaii
# Downloading from BLS Website LA: la.data.19.Idaho
# Downloading from BLS Website LA: la.data.2.AllStatesU
# Downloading from BLS Website LA: la.data.20.Illinois
# Downloading from BLS Website LA: la.data.21.Indiana
# Downloading from BLS Website LA: la.data.22.Iowa
# Downloading from BLS Website LA: la.data.23.Kansas
# Downloading from BLS Website LA: la.data.24.Kentucky
# Downloading from BLS Website LA: la.data.25.Louisiana
# Downloading from BLS Website LA: la.data.26.Maine
# Downloading from BLS Website LA: la.data.27.Maryland
# Downloading from BLS Website LA: la.data.28.Massachusetts
# Downloading from BLS Website LA: la.data.29.Michigan
# Downloading from BLS Website LA: la.data.3.AllStatesS
# Downloading from BLS Website LA: la.data.30.Minnesota
# Downloading from BLS Website LA: la.data.31.Mississippi
# Downloading from BLS Website LA: la.data.32.Missouri
# Downloading from BLS Website LA: la.data.33.Montana
# Downloading from BLS Website LA: la.data.34.Nebraska
# Downloading from BLS Website LA: la.data.35.Nevada
# Downloading from BLS Website LA: la.data.36.NewHampshire
# Downloading from BLS Website LA: la.data.37.NewJersey
# Downloading from BLS Website LA: la.data.38.NewMexico
# Downloading from BLS Website LA: la.data.39.NewYork
# Downloading from BLS Website LA: la.data.4.RegionDivisionU
# Downloading from BLS Website LA: la.data.40.NorthCarolina
# Downloading from BLS Website LA: la.data.41.NorthDakota
# Downloading from BLS Website LA: la.data.42.Ohio
# Downloading from BLS Website LA: la.data.43.Oklahoma

```



```
# Downloading from BLS Website LA: la.data.44.Oregon
# Downloading from BLS Website LA: la.data.45.Pennsylvania
# Downloading from BLS Website LA: la.data.46.PuertoRico
# Downloading from BLS Website LA: la.data.47.RhodeIsland
# Downloading from BLS Website LA: la.data.48.SouthCarolina
# Downloading from BLS Website LA: la.data.49.SouthDakota
# Downloading from BLS Website LA: la.data.5.RegionDivisionS
# Downloading from BLS Website LA: la.data.50.Tennessee
# Downloading from BLS Website LA: la.data.51.Texas
# Downloading from BLS Website LA: la.data.52.Utah
# Downloading from BLS Website LA: la.data.53.Vermont
# Downloading from BLS Website LA: la.data.54.Virginia
# Downloading from BLS Website LA: la.data.56.Washington
# Downloading from BLS Website LA: la.data.57.WestVirginia
# Downloading from BLS Website LA: la.data.58.Wisconsin
# Downloading from BLS Website LA: la.data.59.Wyoming
# Downloading from BLS Website LA: la.data.60.Metro
# Downloading from BLS Website LA: la.data.61.Division
# Downloading from BLS Website LA: la.data.62.Micro
# Downloading from BLS Website LA: la.data.63.Combined
# Downloading from BLS Website LA: la.data.64.County
# Downloading from BLS Website LA: la.data.65.City
# Downloading from BLS Website LA: la.data.7.Alabama
# Downloading from BLS Website LA: la.data.8.Alaska
# Downloading from BLS Website LA: la.data.9.Arizona
# Downloading from BLS Website LA: la.footnote
# Downloading from BLS Website LA: la.map_info
# Downloading from BLS Website LA: la.measure
# Downloading from BLS Website LA: la.period
# Downloading from BLS Website LA: la.series
# Downloading from BLS Website LA: la.state_region_division
```

Computing Environment

```
Sys.time()
```

```
## [1] "2018-09-23 21:56:59 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/libRlapack.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 rvest_0.3.2      xml2_1.2.0      forcats_0.3.0
## [5] stringr_1.3.1  dplyr_0.7.6      purrr_0.2.5     readr_1.1.1
## [9] tidyr_0.8.1    tibble_1.4.2     ggplot2_3.0.0   tidyverse_1.2.1
## [13] curl_3.2
##
## 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     httr_1.3.1       knitr_1.20      hms_0.4.2
## [25] rprojroot_1.3-2 grid_3.5.1       tidyselect_0.2.4 glue_1.3.0
## [29] R6_2.2.2        fansi_0.3.0      readxl_1.1.0    rmarkdown_1.10
## [33] selectr_0.4-1   modelr_0.1.2     magrittr_1.5     backports_1.1.2
## [37] scales_1.0.0    htmltools_0.3.6  assertthat_0.2.0 colorspace_1.3-2
## [41] utf8_1.1.4      stringi_1.2.4    lazyeval_0.2.1  munsell_0.5.0
## [45] broom_0.5.0     crayon_1.3.4
```