R Packages and Publicly Sourced Data

New Member Presentation

Megha Patel
Economist
Acceptance Testing Team

Background

- Recent Master's Graduate in Economic Development - VU
  - Specialization: Economics of Poverty in Developed and Developing Countries, Microeconomics
- Undergraduate Degree: Economics - RU
- Skillset Coming in: R/RStudio, Python, STATA

What is a Package?

- **What is a Package?**
  - Packages in R are collections of functions and/or data that are bundled up for easy distribution. If an R user creates a package and thinks that it might be useful for others, that user can distribute it through a package repository.
  - **Packages can:**
    - Support Quality Control
    - Create visuals
    - Provide statistics

R Package: BLS APIs

- https://cran.r-project.org/web/packages/blscrapeR/blscrapeR.pdf

- Scrapes various data from The U.S. Bureau of Labor Statistics: the statistical branch of the United States Department of Labor. The package has additional functions to help parse, analyze and visualize the data.

`inflation_adjust` *Convert the Value of a US Dollar to a Given Year*

Description

Returns a data frame that uses data from the Consumer Price Index (All Goods) to convert the value of a US Dollar [$1.00 USD] over time.

Usage

```
inflation_adjust(base_year = NA, ...)
```

Arguments

`base_year`	= A string or integer argument to represent the base year that you would like dollar values converted to. For example, if you want to see the value of a 2007 dollar in 2015, you would select 2015 as a base year and find 2007 in the table.
`...`	additional arguments

Examples

```
## Get historical USD values based on a 2010 dollar.
values <- inflation_adjust(base_year = 2015)
```

Covid-19 R Maps File.R | Untitled1* | inflation_adjust

Source on Save

```r

#install and load package
install.packages("blscraper")
library("blscraper")

#demos use of API webscraper + inflation_adjust function for the CPI
base_year_2017<-inflation_adjust(base_year = 2017)
base_year_2017
base_year_1950<-inflation_adjust(base_year = 1950)
base_year_1950
```

```
> base_year_2017
# A tibble: 75 x 5
   year  avg_cpi adj_value base_year pct_increase
   <chr>   <dbl>     <dbl> <chr>            <dbl>
 1 1947     22.3      0.09 2017             -90.9
 2 1948     24.0      0.1  2017             -90.2
 3 1949     23.8      0.1  2017             -90.3
 4 1950     24.1      0.1  2017             -90.2
 5 1951     26.0      0.11 2017             -89.4
 6 1952     26.6      0.11 2017             -89.2
 7 1953     26.8      0.11 2017             -89.1
 8 1954     26.9      0.11 2017             -89.0
 9 1955     26.8      0.11 2017             -89.1
10 1956     27.2      0.11 2017             -88.9
# ... with 65 more rows
> base_year_1950<-inflation_adjust(base_year = 1950)
trying URL 'https://download.bls.gov/pub/time.series/cu/cu.data.1.
Content type 'text/plain' length 2460503 bytes (2.3 MB)
downloaded 2.3 MB

> base_year_1950
# A tibble: 75 x 5
   year  avg_cpi adj_value base_year pct_increase
   <chr>   <dbl>     <dbl> <chr>            <dbl>
 1 1947     22.3      0.93 1950            -7.19
 2 1948     24.0      1    1950            -0.0727
 3 1949     23.8      0.99 1950            -1.05
 4 1950     24.1      1    1950             0
 5 1951     26.0      1.08 1950             7.94
 6 1952     26.6      1.1  1950            10.4
 7 1953     26.8      1.11 1950            11.2
 8 1954     26.9      1.12 1950            11.6
 9 1955     26.8      1.11 1950            11.4
10 1956     27.2      1.13 1950            13.0
# ... with 65 more rows
```

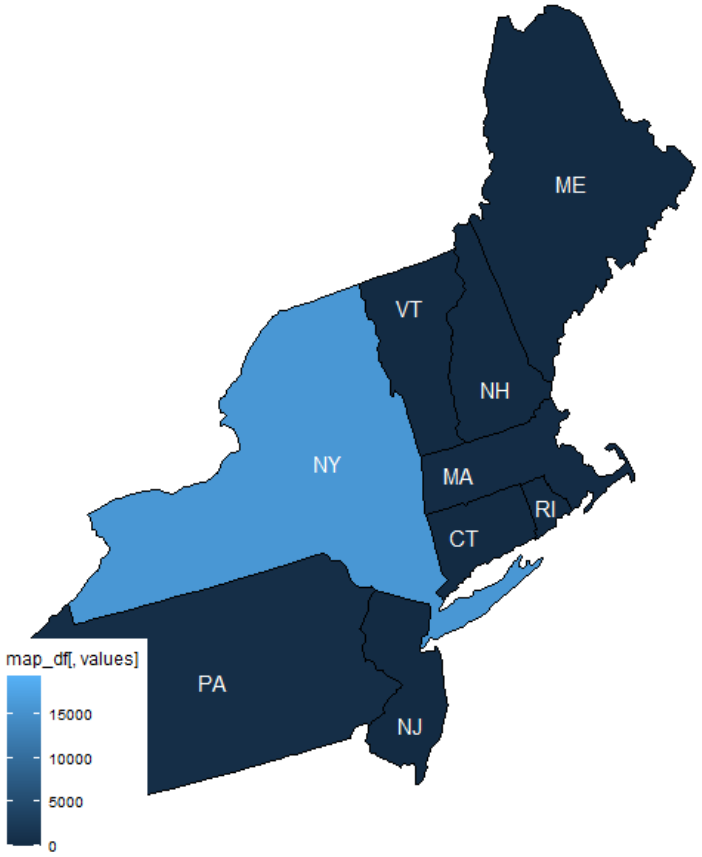
R and RStudio Resources

- https://www.rstudio.com/resources/
- https://cran.r-project.org/
- https://r4ds.had.co.nz/
- https://www.rstudio.com/resources/cheatsheets/
- https://r-graphics.org/index.html
- https://cran.r-project.org/web/packages/dlookr/vignettes/diagonosis.html

Ex. Simplifying Your Workload

R Project: Mapping Covid-19 Cases Jul-2020

- **Desc:** Takes New York Times Covid-19 data and creates maps based on region. Saves individual maps in file folder as images. Meant to be used weekly.
- Self Contained? – Yes
- Packages Used: 4
- Shareable to other R users? Yes
- **Data Source:**
  - https://github.com/nytimes/covid-19-data



Contact Information

Megha S Patel
732-762-8181
patel.megha@bls.gov