DATA607 - Assignment Week 9
Web APIs

Donald Butler

10/24/2021

Introduction

The New York Times web site provides a rich set of APIs, as described here: https://developer.nytimes.com/apis. Your task is to choose one of the New York Times APIs, construct an interface in R to read in the JSON data, and transform it into an R DataFrame.

Setup

API Key Request

We created an account with the NYTimes and requested an API key. The key was saved in a config.yml file which will be read into the project using the config package. To reproduce this work, you would need to follow the instructions to create a key here: https://developer.nytimes.com/get-started. Then save the key into your own config.yml file.

Install libraries

```
library(tidyverse)
library(config)
library(httr)
library(jsonlite)
```

Load API Key

```
cfg <- config::get()

base_url <- 'https://api.nytimes.com/svc/books/v3/lists/'
api_key <- paste('?api-key=',cfg$nytimes$api_key,sep = '')
```

List Names Service

The lists/names service returns a list of all the NYT Best Sellers Lists. Some lists are published weekly and others monthly. The response includes when each list was first published and last published.

```r
list_names <- 'names.json'

list_html <- httr::GET(paste(base_url,list_names,api_key,sep = ''))
list_json <- fromJSON(content(list_html,'text'))

if (list_json$status == 'OK') {
  lists <- as_tibble(list_json$results)

  lists
}
```

```
## # A tibble: 59 x 6
##    list_name    display_name  list_name_encoded oldest_publishe~ newest_publishe~
##    <chr>        <chr>         <chr>             <chr>            <chr>
##  1 Combined P~  Combined Pri~ combined-print-a~ 2011-02-13       2021-10-31
##  2 Combined P~  Combined Pri~ combined-print-a~ 2011-02-13       2021-10-31
##  3 Hardcover ~  Hardcover Fi~ hardcover-fiction 2008-06-08       2021-10-31
##  4 Hardcover ~  Hardcover No~ hardcover-nonfic~ 2008-06-08       2021-10-31
##  5 Trade Fict~  Paperback Tr~ trade-fiction-pa~ 2008-06-08       2021-10-31
##  6 Mass Marke~  Paperback Ma~ mass-market-pape~ 2008-06-08       2017-01-29
##  7 Paperback ~  Paperback No~ paperback-nonfic~ 2008-06-08       2021-10-31
##  8 E-Book Fic~  E-Book Ficti~ e-book-fiction    2011-02-13       2017-01-29
##  9 E-Book Non~  E-Book Nonfi~ e-book-nonfiction 2011-02-13       2017-01-29
## 10 Hardcover ~  Hardcover Ad~ hardcover-advice  2008-06-08       2013-04-21
## # ... with 49 more rows, and 1 more variable: updated <chr>
```

List Data Service

The lists/{date}/{name} service returns the books on the best sellers list for the specified date and list name. To read from the most recent list, use current in place of the date. The name of the list can be found in the list_name_encoded attribute.

Current Hardcover Nonfiction List

The json that is returned contains details about the list as well as the list of books that are on it.

```r
current_nf_html <- httr::GET(paste(base_url,'current/hardcover-nonfiction.json',api_key,sep = ''))
current_nf_json <- fromJSON(content(current_nf_html,'text'))

if (current_nf_json$status == 'OK') {
  current_nf <- as_tibble(current_nf_json$results$books)
  current_nf %>%
    select(rank,rank_last_week, weeks_on_list, title, author)
}
```

```
## # A tibble: 15 x 5
##     rank rank_last_week weeks_on_list title                                author
##    <int>          <int>         <int> <chr>                                <chr>
##  1     1              1             2 THE STORYTELLER                      Dave ~
##  2     2              0             1 TO RESCUE THE REPUBLIC               Bret ~
##  3     3              0             1 THE BOYS                             Ron H~
##  4     4              2             4 PERIL                                Bob W~
##  5     5              0             1 MIDNIGHT IN WASHINGTON               Adam ~
##  6     6              0             1 THE BEATLES: GET BACK                the B~
##  7     7              3             2 TASTE                                Stanl~
##  8     8              5             4 VANDERBILT                           Ander~
##  9     9              6            14 AMERICAN MARXISM                     Mark ~
## 10    10              0             1 IT'S BETTER TO BE FEARED             Seth ~
## 11    11              0             1 E.R. NURSES                          James~
## 12    12              9             2 THERE IS NOTHING FOR YOU HERE        Fiona~
## 13    13              0             1 WHERE THE DEER AND THE ANTELOPE PLAY Nick ~
## 14    14              0             1 RIGGED                               Molli~
## 15    15              7             2 THE DYING CITIZEN                    Victo~
```

There are a number of details about each book on the list as seen in the list of attributes.

```
glimpse(current_nf)
```

```
## Rows: 15
## Columns: 26
## $ rank               <int> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15
## $ rank_last_week     <int> 1, 0, 0, 2, 0, 0, 3, 5, 6, 0, 0, 9, 0, 0, 7
## $ weeks_on_list      <int> 2, 1, 1, 4, 1, 1, 2, 4, 14, 1, 1, 2, 1, 1, 2
## $ asterisk           <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0
## $ dagger             <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0
## $ primary_isbn10     <chr> "0063076098", "0063039540", "006306524X", "198218~
## $ primary_isbn13     <chr> "9780063076099", "9780063039544", "9780063065246"~
## $ publisher          <chr> "<U+200E>Dey Street", "Custom House", "Morrow", "Simon & ~
## $ description        <chr> "A memoir by the musician known for his work with~
## $ price              <chr> "0.00", "0.00", "0.00", "0.00", "0.00", "0.00", "~
## $ title              <chr> "THE STORYTELLER", "TO RESCUE THE REPUBLIC", "THE~
## $ author             <chr> "Dave Grohl", "Bret Baier with Catherine Whitney"~
## $ contributor        <chr> "by Dave Grohl", "by Bret Baier with Catherine Wh~
## $ contributor_note   <chr> "", "", "", "", "", "", "", "", "", "", "", "", "~
## $ book_image         <chr> "https://storage.googleapis.com/du-prd/books/imag~
## $ book_image_width   <int> 331, 329, 329, 331, 329, 417, 331, 329, 324, 329,~
## $ book_image_height  <int> 500, 500, 500, 500, 500, 500, 500, 500, 500, 500,~
## $ amazon_product_url <chr> "https://www.amazon.com/dp/0063076098?tag=NYTBSRE~
## $ age_group          <chr> "", "", "", "", "", "", "", "", "", "", "", "", "~
## $ book_review_link   <chr> "", "", "", "", "", "", "", "", "", "", "", "", "~
## $ first_chapter_link <chr> "", "", "", "", "", "", "", "", "", "", "", "", "~
## $ sunday_review_link <chr> "", "", "", "", "", "", "", "", "", "", "", "", "~
## $ article_chapter_link <chr> "", "", "", "", "", "", "", "", "", "", "", "", "~
## $ isbns              <list> [<data.frame[2 x 2]>], [<data.frame[2 x 2]>], [<d~
## $ buy_links          <list> [<data.frame[6 x 2]>], [<data.frame[6 x 2]>], [<d~
## $ book_uri           <chr> "nyt://book/93914c6c-1313-51de-aa16-047a120720a9"~
```

Fiction List from Last Year

The API also allows you to go back in time to view older lists. Here is the hardcover fiction list from last year.

```
old_f_html <- httr::GET(paste(base_url,'lists/2020-10-25/hardcover-fiction.json',api_key,sep = ''))
old_f_json <- fromJSON(content(old_f_html,'text'))

if (old_f_json$status == 'OK') {
  old_f <- as_tibble(old_f_json$results$books)
  old_f %>%
    select(rank,rank_last_week, weeks_on_list, title, author)
}
```

Summary

The NY Times API allows us to programmatically access data in a standardized format for further analysis.