

# 607\_Fall2021\_HW8\_NYTimesAPI

MarkSchmalfeld

10/21/2021

```
library(tidytext)
library(httr)
library(RCurl)
library(jsonlite)
```

## Problem Statement

Select one of the New York Times API's, construct an interface in R to read in the JSON data, and transform it into an R Dataframe.

Question: What are the NY Times bestsellers in hard-cover fiction currently, 5 years ago and 10 years ago based on the information in the created data frames.

## Problem Approach

- 1) Select the API from the NY Times list for books and create the API GET setup below.
- 2) Review the NY Times API help section to understand how to limit or target the list for data selection. (there are many more options to do targeted selection of information based on the help and API setup information)
- 3) Run a GET function and confirm it functions with a 200 Ok status code. (for first retrieval or test it out) and then extend to more specific additional data - current, 5 yr ago and 10 yr ago best seller lists for hard cover.
- 4) Convert JSON into a dataframe using a direct JSON URL retrieval and inspect results.

## API Call from NY Times book list:

- 1) Setup an account in NY Times API
- 2) Get an API key and enable for the specific API list / source
- 3) Review the specific API call format for the given type of list or source area (Articles, Headlines, Books, Movies etc)

**Example of the NY Times book list call format:** <https://api.nytimes.com/svc/books/v3/lists/current/hardcover-fiction.json?api-key=NR8j0hBE2jWMktrnSnhdBl83TI3cw8lS>

<https://api.nytimes.com/svc/books/v3/lists/%7BYYYY-MM-DD%7D/%7Blist%7D.json>

#Conclusions

We can find the best sellers as a list or identify it by looking at the list to find the #1 best seller.

A. Currently as of Oct 2021 the #1 Best seller is “State of Terror” by Hillary Clinton & Louise Penny and has one week on the list.

B. 5 yr ago Oct 2016 the #1 Best Seller was “Two by Two” by Nicholas Sparks and at the time had one week on the best seller list.

C. 10 yr ago Oct 2011 the #1 Best Seller was “The best of me” by Nicholas Sparks and at the time had one week on the best seller list.

A lot more could be done with the available API data with more time and work. Evaluation of how long a book stays on list, how it changes position in the rankings and any trends for common authors sitting on the best seller list. Interesting was seeing Nicholas Sparks show up twice but 5 yrs apart on the best seller lists.

```
library(httr)

#NYTIMES_KEY="NR8j0hBE2jWMktrnSnhdBl83TI3cw8lS"

url <- "https://api.nytimes.com/svc/books/v3/lists/current/hardcover-fiction.json?api-key=NR8j0hBE2jWMktrnSnhdBl83TI3cw8lS"

(nybook <- GET(url))

## Response [https://api.nytimes.com/svc/books/v3/lists/current/hardcover-fiction.json?api-key=NR8j0hBE2jWMktrnSnhdBl83TI3cw8lS]
##   Date: 2021-10-25 00:33
##   Status: 200
##   Content-Type: application/json; charset=UTF-8
##   Size: 34.9 kB

urlhistory<- "https://api.nytimes.com/svc/books/v3/lists/2016-10-20/hardcover-fiction.json?api-key=NR8j0hBE2jWMktrnSnhdBl83TI3cw8lS"

(nybookhistory5<- GET(urlhistory))

## Response [https://api.nytimes.com/svc/books/v3/lists/2016-10-20/hardcover-fiction.json?api-key=NR8j0hBE2jWMktrnSnhdBl83TI3cw8lS]
##   Date: 2021-10-25 00:33
##   Status: 200
##   Content-Type: application/json; charset=UTF-8
##   Size: 48.6 kB

urlhistoryten<- "https://api.nytimes.com/svc/books/v3/lists/2011-10-25/hardcover-fiction.json?api-key=NR8j0hBE2jWMktrnSnhdBl83TI3cw8lS"

(nybookhistory10<- GET(urlhistoryten))

## Response [https://api.nytimes.com/svc/books/v3/lists/2011-10-25/hardcover-fiction.json?api-key=NR8j0hBE2jWMktrnSnhdBl83TI3cw8lS]
##   Date: 2021-10-25 00:33
##   Status: 200
##   Content-Type: application/json; charset=UTF-8
##   Size: 49.3 kB
```

```
dfbookscurrent<-fromJSON(url)
#current best sellers hardcover fiction

dfbooks5yrago <-fromJSON(urlhistory)
# best sellers hardcover fiction 5 yr ago

dfbooks10yrago<-fromJSON(urlhistoryten)
# best seller hardcover fiction 10 yr ago
```

## Conclusions

### We can find the best sellers as a list or identify it by looking at the list to find the #1 best seller.

#1 Best Seller	Author	Weeks on List
Current Oct 2021 State of Terror Hillary Clinton & Louise Penny	1	5 yr ago Oct 2016 Two by Two Nicholas Sparks
1 10 yr ago Oct 2011 The best of me	Nicholas Sparks	1

A lot more could be done with the available API data with more time and work. Evaluation of how long a book stays on list, how it changes position in the rankings and any trends for common authors sitting on the best seller list. Interesting was seeing Nicholas Sparks show up twice but 5 yrs apart on the best seller lists.

## R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this: so embed plots, for example:

Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.