

# Web API Assignment

Matthew Tillmawitz

2024-10-27

## Fetching the NYT Best Sellers List

This assignment lends itself better to an actual R Script file but for ease of publishing as well as viewing the results it will be done in R Markdown. Note that this code assumes you have already registered an application and authorized it to call the /lists API as well as stored the API key in your .Renvirom. The API call is performed in the first block to allow for data manipulation and experimentation without constantly calling the API.

```
response <- GET(
  "https://api.nytimes.com/svc/books/v3/lists/full-overview.json",
  query = list(
    "api-key" = Sys.getenv("NYT_API_KEY"),
    "date" = "current"
  )
)
```

The lists API returns the best seller lists in a format that is a bit tricky to handle due to the books in each list being interpreted as a nested dataframe as demonstrated below. We are able to use the simplifyDataFrame argument in the fromJSON function to coerce the data nicely into a dataframe and allow us to make use of the pluck function to pull out the fields we want.

```
data <- fromJSON(rawToChar(response$content), simplifyDataFrame = TRUE) |>
  pluck("results", "lists") |>
  as_tibble()

data |>
  select(list_name, books)
```

```
## # A tibble: 18 x 2
##   list_name                books
##   <chr>                   <list>
## 1 Combined Print and E-Book Fiction <df [15 x 26]>
## 2 Combined Print and E-Book Nonfiction <df [15 x 26]>
## 3 Hardcover Fiction          <df [15 x 26]>
## 4 Hardcover Nonfiction       <df [15 x 26]>
## 5 Trade Fiction Paperback     <df [15 x 26]>
## 6 Paperback Nonfiction        <df [15 x 26]>
## 7 Advice How-To and Miscellaneous <df [10 x 26]>
## 8 Childrens Middle Grade Hardcover <df [10 x 26]>
## 9 Picture Books              <df [10 x 26]>
## 10 Series Books               <df [10 x 26]>
```

```
## 11 Young Adult Hardcover          <df [10 x 26]>
## 12 Audio Fiction                  <df [15 x 26]>
## 13 Audio Nonfiction                <df [15 x 26]>
## 14 Business Books                 <df [10 x 26]>
## 15 Graphic Books and Manga         <df [15 x 26]>
## 16 Mass Market Monthly             <df [15 x 26]>
## 17 Middle Grade Paperback Monthly  <df [10 x 26]>
## 18 Young Adult Paperback Monthly   <df [10 x 26]>
```

To handle the column of dataframes we can use the `unnest` function which preserves the context of the parent rows and populates it for the expanded tibbles. We do a little reordering and drop any columns for which none of the books have data and are left with a nicely formatted dataframe we can use for analysis.

```
bestsellers <- data|>
  unnest(books) |>
  # Reorder columns for better readability
  select(
    list_id,
    list_name,
    rank,
    title,
    author,
    rank_last_week,
    weeks_on_list,
    publisher,
    description,
    everything()
  ) |>
  select(where(~ !all(is.na(.) | . == "")))

bestsellers
```

```
## # A tibble: 230 x 29
##   list_id list_name    rank title author rank_last_week weeks_on_list publisher
##   <int> <chr>      <int> <chr> <chr>      <int>      <int> <chr>
## 1     704 Combined P~    1 THE ~ Micha~    0          1 Little, ~
## 2     704 Combined P~    2 THE ~ Freid~    1          3 Poisoned~
## 3     704 Combined P~    3 FOUR~ Rebec~    5         65 Red Tower
## 4     704 Combined P~    4 COUN~ Nicho~    2          4 Random H~
## 5     704 Combined P~    5 THE ~ Freid~    3         67 Grand Ce~
## 6     704 Combined P~    6 THE ~ Krist~    6         37 St. Mart~
## 7     704 Combined P~    7 IRON~ Rebec~    7         42 Red Tower
## 8     704 Combined P~    8 A CO~ Sarah~    8         28 Bloomsbu~
## 9     704 Combined P~    9 IT E~ Colle~   11        143 Atria
## 10    704 Combined P~   10 INTE~ Sally~   10          4 Farrar, ~
## # i 220 more rows
## # i 21 more variables: description <chr>, list_name_encoded <chr>,
## #   display_name <chr>, updated <chr>, age_group <chr>,
## #   amazon_product_url <chr>, book_image <chr>, book_image_width <int>,
## #   book_image_height <int>, book_review_link <chr>, book_uri <chr>,
## #   btrn <chr>, contributor <chr>, contributor_note <chr>, created_date <chr>,
## #   price <chr>, primary_isbn10 <chr>, primary_isbn13 <chr>, ...
```