Assignment7

Jose Rodriguez

2022-10-30

Overview

In this project we explore how to connect to an API. We specifically connect to the New York Times API to request the latest list of best-selling books. Ideally, we want to receive our data from the API in JSON format. Once stored locally in our machine, it'll be converted to a flat-file. In this case a dataframe. We want the dataframe to be "tidy", and therefore will execute some functions to unnest lists found in the json data.

In this project we use the following library packages:

- jsonlite
- httr
- tidyverse

Connecting to API

In this section we set up our base url to make our initial request to the API. We make use of paste0() to seamlessly concatenate our private New York Times Developer API key. With the use of GET() we send our request and containerize it. Once we have our data locally, we check the status of the response by using status_code(). A 200 code means we successfully connected. Additionally, we use warn_for_status() to alert us of a failed connection; this function only executes if the connection is unsuccessful.

```
base_url <- "https://api.nytimes.com/svc/books/v3/lists/current/culture.json?api-key="
r <- GET(paste0(base_url, api_key)) #get name of best sellers list
status_code(r)
```

```
## [1] 200
```

```
warn_for_status(r)
```

Retrieving API Data

Next we use content() to retrieve the information stored in our variable. Second, we use fromJSON() to convert from a JSON object to an R object. Once an R object we can navigate the hierarchy until we reach our list of books [newdata$results$books]. We combined the JSON to R conversion with as.data.frame() to get our data into a "flat" state.

However, our dataframe has 2 complex columns. The columns have dataframes as values.

```
data <- content(r, as="text")
newdata <- fromJSON(data, flatten=TRUE)
books_df <- as.data.frame(newdata$results$books)

knitr::kable(head(books_df[, 1:26]), "simple")
```

rank	rank_last_week	weeks_on_list	asterisk	dagger	primary_isbn10	primary_isbn13	publisher
1	0	0	0	0	1455539740	9781455539741	Grand Central/Melcher
2	0	0	0	0	0399588175	9780399588174	Spiegel & Grau
3	0	0	0	0	1455565385	9781455565382	Grand Central
4	0	0	0	0	0062571370	9780062571373	HarperDesign
5	0	0	0	0	1250058902	9781250058904	St. Martin's
6	0	0	0	0	0062316095	9780062316097	Harper

Tidying Data

The final task is to tidy our dataframe. In this case we chose to pivot wider, however, pivot longer can also be done. For sake of this project we pivot wider to reduce the number of observations. We pipe the unnest_wider function for each complex column. The result is a fully flat file.

```
final_df <- books_df |>
  unnest_wider(isbns) |>
  unnest_wider(buy_links)

knitr::kable(head(final_df[, 1:28]), "simple")
```

rank	rank_last_week	weeks_on_list	asterisk	dagger	primary_isbn10	primary_isbn13	publisher
1	0	0	0	0	1455539740	9781455539741	Grand Central/Melcher
2	0	0	0	0	0399588175	9780399588174	Spiegel & Grau
3	0	0	0	0	1455565385	9781455565382	Grand Central
4	0	0	0	0	0062571370	9780062571373	HarperDesign
5	0	0	0	0	1250058902	9781250058904	St. Martin's
6	0	0	0	0	0062316095	9780062316097	Harper