

607 Assignment 9

Lu Beyer

Overview

I am looking to access, read, and store data from the official NYTimes API portal into a usable R dataframe.

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr      1.1.4      v readr      2.1.5
## v forcats    1.0.0      v stringr   1.5.1
## v ggplot2    3.4.4      v tibble    3.2.1
## v lubridate  1.9.3      v tidyr     1.3.1
## v purrr      1.0.2
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(jsonlite)
```

```
##
## Attaching package: 'jsonlite'
##
## The following object is masked from 'package:purrr':
##
##     flatten
```

From the NYTimes API portal, I was granted an API key linked to my account.

```
key = 'aF3xU5VRcQVibOrgoG12U7RX9SsUs2Nk'
```

I found an interesting API, a list of historical best selling books.

```
url = 'https://api.nytimes.com/svc/books/v3/lists/best-sellers/history.json?api-key='
```

I used the base url provided by the API portal, adding my key, to create a full url to access the API.

```
a <- paste0(url, key)
```

Using the jsonlite package function fromJSON(), I requested the data from the API, and assigned it to a list.

```
b <- fromJSON(a)
```

I converted the relevant data from the list into a dataframe, which can now be more easily utilized within R.

```
c <- as.data.frame(b$results)
```

Conclusion

I found this to be a useful workflow for quickly and easily pulling official data from their source. While JSON format is typically a bit harder to read, as discussed in previous assignments, their structure allows for easy machine readability, allowing us to work with the data quickly after receiving. This especially allows for a more streamlined process since APIs seem to be fairly standardized.