

Assignment Web APIs

Samuel I Kigamba

October 27, 2019

The New York Times web site provides a rich set of APIs, as described here: <https://developer.nytimes.com/apis> You will need to start by signing up for an API key.

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.

Load all the relevant libraries (install packages if none exists)

```
library(dplyr)
library(stringr)
library(httr)
library(jsonlite)
library(tidyr)
```

Define the HTTP request and in this case from the NYT website : <https://developer.nytimes.com/docs/articlesearch-product/1/routes/articlesearch.json/get> Follow the instructions in the link provided above to set an API key for use to access the required data

```
HTTP_request <- "https://api.nytimes.com/svc/search/v2/articlesearch.json"
```

Define the search Query in this case to search for business articles

```
search_query <- "?q=business"
```

Apply filters to limit the results to business and finance. The following values are allowed: day_of_week, document_type, ingredients, news_desk, pub_month, pub_year, section_name, source, subsection_name, type_of_material

```
filters <- "&fq=document_type(article)&fq=news_desk(business)&fq=type_of_material(business)&fq=keyword(")
```

Define the start (01/01/2015) and end dates of today (10/27/2019)

```
start_date <- "&begin_date=20150101"
end_date <- "&end_date=20191027"
```

Sort by relevance (other allowable parameters include: oldest and newest)

```
sort_by <- "&sort=relevance"
```

Hide API key Include your API key generated from the website into a separate line of code and set the include parameter in the R code block to FALSE

```
hidden_API_key <- api_key
```

```
business_articles <- fromJSON(paste0(HTTP_request, search_query, filters, start_date, end_date, sort_by),
results <- business_articles$response$docs
```

Follow the link below to see all the possible parameters <https://developer.nytimes.com/docs/articlesearch-product/1/routes/articlesearch.json/get>

```
business_df <- data.frame('Date' = format(as.Date(results$pub_date), "%B %d, %Y"),
                          'news_desk' = results$news_desk,
                          #'Headline' = results$headline$main,
                          'source' = results$source,
                          'URL' = results$web_url
                          )
```

```
business_df
```

##	Date	news_desk	source
## 1	May 06, 2015	Business	The New York Times
## 2	May 01, 2017	Letters	The New York Times
## 3	July 14, 2016	Business	The New York Times
## 4	March 02, 2016	Business	The New York Times
## 5	July 19, 2017	Letters	The New York Times
## 6	November 22, 2017	Business	The New York Times
## 7	September 19, 2015	Business Day	The New York Times
## 8	September 15, 2015	Business	The New York Times
## 9	June 25, 2018	Letters	The New York Times
## 10	March 29, 2019	Washington	The New York Times
## 1			https://www.nytimes.com/2015/05/07/business/to-cuba
## 2			https://www.nytimes.com/2017/05/01/opinion/what-business-school
## 3			https://www.nytimes.com/2016/07/15/business/for-trump-business-leaders-are-more-eli
## 4			https://www.nytimes.com/2016/03/02/business/government-meets-goal-set-in-1994-for-womens-busin
## 5			https://www.nytimes.com/2017/07/19/opinion/helping-local
## 6			https://www.nytimes.com/2017/11/22/business/net-neutrality-sma
## 7			https://www.nytimes.com/interactive/2015/09/19/business/smallbusiness/entrepreneurship-lessons-far
## 8			https://www.nytimes.com/2015/09/15/business/stanford-business-school-dean-to-resign-o
## 9			https://www.nytimes.com/2018/06/25/opinion/small-busin
## 10			https://www.nytimes.com/2019/03/29/us/politics/linda-mcmahon-small-b