# Assignment 9

Heleine Fouda

2023-11-05

Assignment: 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 Data Frame.

## Getting started

```r
# Getting the API key
api_key <- "l0cER50thCByvsfBr3SqQcYsOufP4Iuw"
# Getting the URL
url <-("https://api.nytimes.com/svc/topstories/v2/world.json?api-key=")
# Building the full URL
full_url <- paste("https://api.nytimes.com/svc/topstories/v2/world.json?api-key=l0cER50thCByvsfBr3SqQcY
topstories <- GET(full_url)
# Check status code
topstories$status_code
```

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

```r
# Data set summary
summary(topstories)
```

```
##               Length Class      Mode
## url                1 -none-     character
## status_code        1 -none-     numeric
## headers           22 insensitive list
## all_headers        1 -none-     list
## cookies            7 data.frame  list
## content        71094 -none-     raw
## date               1 POSIXct    numeric
## times              6 -none-     numeric
## request            7 request    list
## handle             1 curl_handle externalptr
```

```r
# Data set structure
str(topstories)
```

```
## List of 10
##  $ url        : chr "https://api.nytimes.com/svc/topstories/v2/world.json?api-key=l0cER50thCByvsfBr3S
##  $ status_code: int 200
##  $ headers    :List of 22
##   ..$ date                    : chr "Sun, 05 Nov 2023 08:45:53 GMT"
##   ..$ content-type            : chr "application/json"
##   ..$ transfer-encoding       : chr "chunked"
##   ..$ connection              : chr "keep-alive"
##   ..$ cache-control           : chr "max-age=600, public"
```

```
##   ..$ x-envoy-upstream-service-time: chr "79"
##   ..$ server                      : chr "envoy"
##   ..$ content-encoding            : chr "gzip"
##   ..$ x-envoy-decorator-operation : chr "du-topstories-api.publishing.nyti.nyt.net:443/*"
##   ..$ accept-ranges               : chr "bytes"
##   ..$ via                         : chr "1.1 varnish"
##   ..$ age                         : chr "575"
##   ..$ x-served-by                 : chr "cache-iad-kcgs7200043-IAD"
##   ..$ x-cache                     : chr "HIT"
##   ..$ x-cache-hits                : chr "1"
##   ..$ x-timer                     : chr "S1699173953.182992,VS0,VE1"
##   ..$ vary                        : chr "Accept-Encoding"
##   ..$ access-control-allow-origin : chr "*"
##   ..$ access-control-allow-headers : chr "Accept, Content-Type, X-Forwarded-For, X-Prototype-Version
##   ..$ access-control-expose-headers: chr "Content-Length, X-JSON"
##   ..$ access-control-allow-methods : chr "GET, OPTIONS"
##   ..$ strict-transport-security    : chr "max-age=63072000; preload; includeSubdomains"
##   ..- attr(*, "class")= chr [1:2] "insensitive" "list"
## $ all_headers:List of 1
##   ..$ :List of 3
##   .. ..$ status : int 200
##   .. ..$ version: chr "HTTP/1.1"
##   .. ..$ headers:List of 22
##   .. .. ..$ date                        : chr "Sun, 05 Nov 2023 08:45:53 GMT"
##   .. .. ..$ content-type                : chr "application/json"
##   .. .. ..$ transfer-encoding           : chr "chunked"
##   .. .. ..$ connection                  : chr "keep-alive"
##   .. .. ..$ cache-control               : chr "max-age=600, public"
##   .. .. ..$ x-envoy-upstream-service-time: chr "79"
##   .. .. ..$ server                      : chr "envoy"
##   .. .. ..$ content-encoding            : chr "gzip"
##   .. .. ..$ x-envoy-decorator-operation : chr "du-topstories-api.publishing.nyti.nyt.net:443/*"
##   .. .. ..$ accept-ranges               : chr "bytes"
##   .. .. ..$ via                         : chr "1.1 varnish"
##   .. .. ..$ age                         : chr "575"
##   .. .. ..$ x-served-by                 : chr "cache-iad-kcgs7200043-IAD"
##   .. .. ..$ x-cache                     : chr "HIT"
##   .. .. ..$ x-cache-hits                : chr "1"
##   .. .. ..$ x-timer                     : chr "S1699173953.182992,VS0,VE1"
##   .. .. ..$ vary                        : chr "Accept-Encoding"
##   .. .. ..$ access-control-allow-origin : chr "*"
##   .. .. ..$ access-control-allow-headers : chr "Accept, Content-Type, X-Forwarded-For, X-Prototype-Ve
##   .. .. ..$ access-control-expose-headers: chr "Content-Length, X-JSON"
##   .. .. ..$ access-control-allow-methods : chr "GET, OPTIONS"
##   .. .. ..$ strict-transport-security    : chr "max-age=63072000; preload; includeSubdomains"
##   .. .. ..- attr(*, "class")= chr [1:2] "insensitive" "list"
## $ cookies    :'data.frame': 0 obs. of  7 variables:
##   ..$ domain    : logi(0)
##   ..$ flag      : logi(0)
##   ..$ path      : logi(0)
##   ..$ secure    : logi(0)
##   ..$ expiration: 'POSIXct' num(0)
##   ..$ name      : logi(0)
##   ..$ value     : logi(0)
```

```
## $ content     : raw [1:71094] 7b 22 73 74 ...
## $ date        : POSIXct[1:1], format: "2023-11-05 08:45:53"
## $ times       : Named num [1:6] 0 0.00247 0.05629 0.41399 0.71082 ...
##  ..- attr(*, "names")= chr [1:6] "redirect" "namelookup" "connect" "pretransfer" ...
## $ request     :List of 7
##  ..$ method    : chr "GET"
##  ..$ url       : chr "https://api.nytimes.com/svc/topstories/v2/world.json?api-key=l0cER50thCByvsfB:
##  ..$ headers   : Named chr "application/json, text/xml, application/xml, */*"
##  .. ..- attr(*, "names")= chr "Accept"
##  ..$ fields    : NULL
##  ..$ options   :List of 2
##  .. ..$ useragent: chr "libcurl/7.64.1 r-curl/5.0.0 httr/1.4.6"
##  .. ..$ httpget  : logi TRUE
##  ..$ auth_token: NULL
##  ..$ output    : list()
##  .. ..- attr(*, "class")= chr [1:2] "write_memory" "write_function"
##  ..- attr(*, "class")= chr "request"
## $ handle      :Class 'curl_handle' <externalptr>
## - attr(*, "class")= chr "response"
```

```r
# Data set names
names(topstories)
```

```
## [1] "url"         "status_code" "headers"     "all_headers" "cookies"
## [6] "content"     "date"        "times"       "request"     "handle"
```

```r
top_stories <- content(topstories, as = "text")
```

```
## No encoding supplied: defaulting to UTF-8.
```

Leveraging the jsonlite and the tidyr libraries

```r
# Assuming 'top_stories' is a valid JSON character string
top_stories_df <- fromJSON(top_stories)

# Flatten the 'results' element if it's a nested list
if ("results" %in% names(top_stories_df)) {
  top_stories_df <- top_stories_df$results
}
kable(top_stories_df)
```

| section | subsection | title |
|---|---|---|
| world | europe | French Nuns and Climate Activists Square Off Over Plans for a Megachurch |
| | | |
| world | europe | 'I Am Dreaming It Will Stop': A Deadlocked War Tests Ukrainian Morale |
| world | middleeast | In the Aftermath of the Hamas Attack: Portraits of Survivors |
| world | americas | The New Enemies of Argentina's Far Right: Swifties and the BTS Army |
| world | europe | Zelensky Rebuke of Top General Signals Rift in Ukrainian Leadership |
| world | asia | Earthquake in Western Nepal Kills More Than 150 |
| world | asia | Militant Attack on Air Base Is Latest Episode to Unnerve Pakistan |
| world | asia | Chapter 3: Bargaining for Time |
| world | australia | Why China and Australia Are Reconciling. Sort of. |
| world | asia | 'Just Like Medicine': A New Push for Divorce in a Nation Where It's Illegal |
| world | africa | As a Teen, She Loved Video Games. Now She's Using A.I. to Try to Quash Malaria. |
| world | africa | U.S. Faces Tricky Questions With African Trade Group |
| world | europe | Flooding in Tuscany Leaves Six Dead |
| world | europe | Hamburg Airport Closed After Armed Man Drives Onto Tarmac, Prompting Standoff |
| world | middleeast | Blinken Meets Arab Ministers in Bid to Calm Outrage Over Gaza Airstrikes |
| us | politics | U.S. Officials Outline Steps to Israel to Reduce Civilian Casualties |
| briefing | | Rothko, in Pain and Glory |
| world | americas | Remembering Riopelle, the First Canadian Modern Artist Recognized by the World |
| world | middleeast | Rashida Tlaib Posts Video Accusing Biden of Supporting 'Genocide' |
| world | middleeast | Israel Used 2,000-Pound Bombs in Strike on Jabaliya, Analysis Shows |
| world | middleeast | Hamas put wounded fighters on departure lists, delaying efforts to evacuate foreigners, a U.S. offici |
| world | middleeast | Palestinian TV Reporter and 11 Family Members Killed in Gaza |
| world | middleeast | Blinken Urges 'Pauses' in Fighting to Deliver Aid for Gaza Civilians |
| world | middleeast | Many are killed near a Gaza hospital, its chief says, in a strike the Israelis say was aimed at Hama |
| world | europe | David Kirke, Who Made the First Modern Bungee Jump, Dies at 78 |
| world | asia | Japan and Philippines, Wary of China, Look to Expand Military Ties |
| world | americas | Is Fluminense the Team of the Future? |
| world | asia | New Delhi Chokes as Annual Curse of Pollution Returns With a Vengeance |
| us | politics | Biden Hosts Leaders of South American Nations for an Economic Summit |
| world | europe | Russia Sends Biggest Drone Barrage in Weeks, Ukrainian Officials Say |
| world | europe | Headwind Cycling Race Called Off Over Too Much Wind |
| world | middleeast | In High-Stakes Speech, Hezbollah's Leader Stops Short of Call to Expand Hamas War |
| world | middleeast | As Gaza War Enters New Phase, Israel Faces Pressure Over Civilian Deaths |

```r
#Count of Section
top_stories_df <- data.frame(
  "section" = character(34),
  "subsection" = character(34),
  "title" = character(34),
  "abstract" = character(34),
  "url" = character(34),
  "uri" = character(34),
  "byline" = character(34),
  "item_type" = character(34),
  "updated" = character(34),
  "created_date" = character(34),
  "published_date" = character(34),
  "material_type_facet" = character(34),
  "kicker"= character(34),
  "des_facet" = character(34),
  "org_facet" = character(34),
```

```r
  "per_facet" = character(34),
  "geo_facet" = character(34),
  "multimedia" = character(34),
  "short_url" = character(34)
)


# Check the structure of the data frame
str(top_stories_df)
```

```
## 'data.frame':    34 obs. of  19 variables:
##  $ section            : chr  "" "" "" "" ...
##  $ subsection         : chr  "" "" "" "" ...
##  $ title              : chr  "" "" "" "" ...
##  $ abstract           : chr  "" "" "" "" ...
##  $ url                : chr  "" "" "" "" ...
##  $ uri                : chr  "" "" "" "" ...
##  $ byline             : chr  "" "" "" "" ...
##  $ item_type          : chr  "" "" "" "" ...
##  $ updated            : chr  "" "" "" "" ...
##  $ created_date       : chr  "" "" "" "" ...
##  $ published_date     : chr  "" "" "" "" ...
##  $ material_type_facet: chr  "" "" "" "" ...
##  $ kicker             : chr  "" "" "" "" ...
##  $ des_facet          : chr  "" "" "" "" ...
##  $ org_facet          : chr  "" "" "" "" ...
##  $ per_facet          : chr  "" "" "" "" ...
##  $ geo_facet          : chr  "" "" "" "" ...
##  $ multimedia         : chr  "" "" "" "" ...
##  $ short_url          : chr  "" "" "" "" ...
```

```r
  colnames <- c(
  "section",
  "subsection",
  "title",
  "abstract",
  "url",
  "uri",
  "byline",
  "item_type",
  "updated",
  "created_date",
  "published_date",
  "material_type_facet",
  "kicker",
  "des_facet",
  "org_facet",
  "per_facet",
  "geo_facet",
  "multimedia",
  "short_url"
)
colnames
```

```
##  [1] "section"             "subsection"          "title"
##  [4] "abstract"            "url"                 "uri"
```

```
##  [7] "byline"              "item_type"           "updated"
## [10] "created_date"        "published_date"      "material_type_facet"
## [13] "kicker"              "des_facet"           "org_facet"
## [16] "per_facet"           "geo_facet"           "multimedia"
## [19] "short_url"
```

```r
byline <- top_stories_df %>%
  group_by(byline) %>%
  summarise(num=n()) %>%
  arrange(desc(num))
byline
```

```
## # A tibble: 1 x 2
##   byline   num
##   <chr>  <int>
## 1 ""        34
```