

DATA 607 Assignment 9

RJM

2019-10-26

Slection of the libraries:

```
library(httr)
library(jsonlite)
library(knitr)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##   filter, lag

## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
```

```
library(tidyverse)
```

```
## -- Attaching packages ----- tidyverse 1.2.1 --

## v ggplot2 3.2.1    v readr    1.3.1
## v tibble  2.1.3    v purrr   0.3.2
## v tidyr   1.0.0    v stringr 1.4.0
## v ggplot2 3.2.1    v forcats 0.4.0

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x purrr::flatten() masks jsonlite::flatten()
## x dplyr::lag()     masks stats::lag()
```

The code to extract the articles based on the mention of HIV in them from the dates of Jan 01, 2016 to Oct 31, 2019.

```
apikey <- '1pHIViWxgD31sA9yA5uvfIkyi8wv7Zn0'

path <- "http://api.nytimes.com/svc/search/v2/articlesearch.json"
# install.packages("httr")
```

```
library(httr)
r <- GET(path, query=list(q="HIV",
                          "api-key"=apikey,
                          "begin_date"=20160101,
                          "end_date"=20191031))
response <- content(r, as = "text", encoding = "UTF-8")
```

Creation of a dataframe using from JSON:

```
df <- fromJSON(response, flatten = TRUE) %>%
  data.frame()

head(df)
```

```
##      status
## 1      OK
## 2      OK
## 3      OK
## 4      OK
## 5      OK
## 6      OK
##
##                                     copyright
## 1 Copyright (c) 2019 The New York Times Company. All Rights Reserved.
## 2 Copyright (c) 2019 The New York Times Company. All Rights Reserved.
## 3 Copyright (c) 2019 The New York Times Company. All Rights Reserved.
## 4 Copyright (c) 2019 The New York Times Company. All Rights Reserved.
## 5 Copyright (c) 2019 The New York Times Company. All Rights Reserved.
## 6 Copyright (c) 2019 The New York Times Company. All Rights Reserved.
##
##                                     response.docs.web_ur
## 1      https://www.nytimes.com/aponline/2019/10/26/us/ap-us-hiv-medication-lockers-miami.htm
## 2      https://www.nytimes.com/aponline/2019/10/24/world/africa/ap-af-uganda-gay-rights.htm
## 3 https://www.nytimes.com/reuters/2019/10/22/business/22reuters-health-globalfund-climate-change.htm
## 4      https://www.nytimes.com/2019/08/08/movies/vision-portraits-review.htm
## 5      https://www.nytimes.com/2019/09/05/world/africa/pope-africa-photos.htm
## 6      https://www.nytimes.com/2018/10/25/movies/1985-film-review-cory-michael-smith.htm
##
## 1
## 2      Ugandan police have detained 16 men on suspicion of homosexuality and human traffickin
## 3 Climate change is making it harder to eradicate deadly epidemics, with rising temperatures helping
## 4                                     A documentary present
## 5                                     The pontiff is on a three-country tour
## 6                                     The film is a m
##
## 1
## 2                                     KAMPALA, Uganda - Ugandan police
## 3                                     GENEVA - Climate change is making it
## 4 Filmmakers have always faced an obstacle when it comes to depicting blindness on screen; some topi
## 5                                     MAPUTO, Mozambique - Pope Francis concluded his trip to L
## 6      The AIDS crisis, which will remain a global crisis until a definitive cure for t
##
```

```
## 1
## 2           Ugandan police have detained 16 men on suspicion of homosexuality and human traffickin
## 3 Climate change is making it harder to eradicate deadly epidemics, with rising temperatures helping
## 4                                     A documentary present
## 5                                     The pontiff is on a three-country tour
## 6                                     The film is a m
##   response.docs.source
## 1             AP
## 2             AP
## 3             Reuters
## 4   The New York Times
## 5   The New York Times
## 6   The New York Times
##
## 1
## 2
## 3
## 4 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
## 5
## 6
##
## 1
## 2
## 3
## 4
## 5 subject, organizations, persons, glocations, glocations, glocations, subject, glocations, organiza
## 6
##       response.docs.pub_date response.docs.document_type
## 1 2019-10-26T12:17:04+0000          article
## 2 2019-10-24T13:44:07+0000          article
## 3 2019-10-22T15:24:20+0000          article
## 4 2019-08-08T11:00:05+0000          article
## 5 2019-09-05T09:23:48+0000          article
## 6 2018-10-25T11:00:01+0000          article
##       response.docs.news_desk response.docs.section_name
## 1              None            U.S.
## 2              None            World
## 3              None      Business Day
## 4        Weekend            Movies
## 5        Foreign            World
## 6        Weekend            Movies
##       response.docs.type_of_material
## 1              News
## 2              News
## 3              News
## 4        Review
## 5              News
## 6        Review
##                               response.docs._id
## 1 nyt://article/01396498-cda4-5db2-9baa-4f037fc38deb
## 2 nyt://article/726e81a6-ab8f-5811-a939-c64439fac456
## 3 nyt://article/7500f04f-cd0e-543e-91bc-73713a411102
## 4 nyt://article/8b456319-0028-528b-be6e-bdae25f74a0
## 5 nyt://article/48e01bc3-2861-5a56-bbba-a3659a1fdb16
```

```

## 6 nyt://article/2bb89e2d-04ec-5b8a-ba5b-d1eca5a06c96
## response.docs.word_count
## 1 670
## 2 408
## 3 404
## 4 413
## 5 1024
## 6 295
## response.docs.uri
## 1 nyt://article/01396498-cda4-5db2-9baa-4f037fc38deb
## 2 nyt://article/726e81a6-ab8f-5811-a939-c64439fac456
## 3 nyt://article/7500f04f-cd0e-543e-91bc-73713a411102
## 4 nyt://article/8b456319-0028-528b-be6e-bdaeb25f74a0
## 5 nyt://article/48e01bc3-2861-5a56-bbba-a3659a1fdb16
## 6 nyt://article/2bb89e2d-04ec-5b8a-ba5b-d1eca5a06c96
## response.docs.subsection_name response.docs.print_page
## 1 <NA> <NA>
## 2 Africa <NA>
## 3 <NA> <NA>
## 4 <NA> 8
## 5 Africa <NA>
## 6 <NA> 9
## response.docs.headline.main
## 1 Medication Lockers Help Miami's Homeless Living With HIV
## 2 Ugandan Police Detain 16 Men Over Suspected Homosexuality
## 3 Climate Change Hampers Progress on Fighting Epidemics: Global Fund
## 4 'Vision Portraits' Review: Blindness as a Way of Seeing
## 5 Pope Francis, in Africa, Urges Mozambique to Put Past Tensions Aside
## 6 Review: In '1985,' a Young Man Hides a Plague From His Family
## response.docs.headline.kicker response.docs.headline.content_kicker
## 1 <NA> NA
## 2 <NA> NA
## 3 <NA> NA
## 4 <NA> NA
## 5 <NA> NA
## 6 <NA> NA
## response.docs.headline.print_headline
## 1 Medication Lockers Help Miami's Homeless Living With HIV
## 2 Ugandan Police Detain 16 Men Over Suspected Homosexuality
## 3 Climate Change Hampers Progress on Fighting Epidemics: Global Fund
## 4 Vision Portraits
## 5 <NA>
## 6 1985
## response.docs.headline.name response.docs.headline.seo
## 1 NA NA
## 2 NA NA
## 3 NA NA
## 4 NA NA
## 5 NA NA
## 6 NA NA
## response.docs.headline.sub response.docs.byline.original
## 1 NA By The Associated Press
## 2 NA By The Associated Press
## 3 NA By Reuters

```

```
## 4 NA By Ben Kenigsberg
## 5 NA By Jason Horowitz
## 6 NA By Glenn Kenny
## response.docs.byline.person
## 1 NULL
## 2 NULL
## 3 NULL
## 4 Ben, NA, Kenigsberg, NA, NA, reported, , 1
## 5 Jason, NA, Horowitz, NA, NA, reported, , 1
## 6 Glenn, NA, Kenny, NA, NA, reported, , 1
## response.docs.byline.organization response.meta.hits
## 1 The Associated Press 107
## 2 The Associated Press 107
## 3 Reuters 107
## 4 <NA> 107
## 5 <NA> 107
## 6 <NA> 107
## response.meta.offset response.meta.time
## 1 0 17
## 2 0 17
## 3 0 17
## 4 0 17
## 5 0 17
## 6 0 17
```

```
nrow(df)
```

```
## [1] 10
```

Selection and renaming of columns of interest:

```
df2 <- select(df, Date = response.docs.pub_date, Source = response.docs.source,
              Section = response.docs.section_name, "Word Count" = response.docs.word_count, URL =
              response.docs.web_url)
```

```
df2
```

```
## Date Source Section Word Count
## 1 2019-10-26T12:17:04+0000 AP U.S. 670
## 2 2019-10-24T13:44:07+0000 AP World 408
## 3 2019-10-22T15:24:20+0000 Reuters Business Day 404
## 4 2019-08-08T11:00:05+0000 The New York Times Movies 413
## 5 2019-09-05T09:23:48+0000 The New York Times World 1024
## 6 2018-10-25T11:00:01+0000 The New York Times Movies 295
## 7 2018-06-14T20:36:01+0000 The New York Times Health 848
## 8 2019-01-14T21:59:52+0000 <NA> T Brand 1426
## 9 2018-08-20T16:25:19+0000 The New York Times Arts 357
## 10 2018-12-01T10:00:10+0000 The New York Times Health 864
```

```
##
## 1 https://www.nytimes.com/aponline/2019/10/26/us/ap-us-hiv-medication-lockers-miami.htm
```

2 [https://www.nytimes.com/aponline/2019/10/24/world/africa/ap-af-uganda-gay-rights.ht](https://www.nytimes.com/aponline/2019/10/24/world/africa/ap-af-uganda-gay-rights.html)
3 [https://www.nytimes.com/reuters/2019/10/22/business/22reuters-health-globalfund-climate-change.ht](https://www.nytimes.com/reuters/2019/10/22/business/22reuters-health-globalfund-climate-change.html)
4 [https://www.nytimes.com/2019/08/08/movies/vision-portraits-review.ht](https://www.nytimes.com/2019/08/08/movies/vision-portraits-review.html)
5 [https://www.nytimes.com/2019/09/05/world/africa/pope-africa-photos.ht](https://www.nytimes.com/2019/09/05/world/africa/pope-africa-photos.html)
6 [https://www.nytimes.com/2018/10/25/movies/1985-film-review-cory-michael-smith.ht](https://www.nytimes.com/2018/10/25/movies/1985-film-review-cory-michael-smith.html)
7 [https://www.nytimes.com/2018/06/14/health/youth-risk-depression-suicide-opioids.ht](https://www.nytimes.com/2018/06/14/health/youth-risk-depression-suicide-opioids.html)
8 [https://www.nytimes.com/paidpost/university-of-iowa/iowa-neuroscience-institute.ht](https://www.nytimes.com/paidpost/university-of-iowa/iowa-neuroscience-institute.html)
9 [https://www.nytimes.com/2018/08/20/arts/television/three-things-to-watch-this-week.ht](https://www.nytimes.com/2018/08/20/arts/television/three-things-to-watch-this-week.html)
10 [https://www.nytimes.com/2018/12/01/health/aids-day-photography-1980s.ht](https://www.nytimes.com/2018/12/01/health/aids-day-photography-1980s.html)