

# hw12.R

mateuszzezula

Sun Dec 2 18:21:48 2018

```
setwd("/users/mateuszzezula/Box/1. First Semester/Programming for Data Science/Homework/")
```

```
library(pacman)
p_load(httr, magrittr, dplyr)
```

```
# KEYS -- ultimately decided abstract the API keys.
```

```
KEYS <- function() {
  API_key.news <- "c6c10679081f4b43ab4352aea936e71b"
  API_key.google <- "AIzaSyC__Z4AlCtIarOddPbhBnHcy_weBGIulw4"
  google.translate.base <- "https://translation.googleapis.com/language/translate/v2?key="
  google.speech.base <- "https://texttospeech.googleapis.com/v1/text:synthesize?key="

  url.google.translate <- paste0(google.translate.base, API_key.google)
  url.google.speech <- paste0(google.speech.base, API_key.google)

  return(list(API_key.news = API_key.news,
              google.translate = url.google.translate,
              google.speech = url.google.speech))
}
KEYS()$google.speech
```

```
## [1] "https://texttospeech.googleapis.com/v1/text:synthesize?key=AIzaSyC__Z4AlCtIarOddPbhBnHcy_weBGIulw4"
```

```
# en.translator -- Takes character vectors and returns English translation as df.
```

```
en.translator <- function(url, text) {
  source <- "fr"
  target <- "en"
  df_output <- setNames(data.frame(matrix(ncol = 1, nrow = length(text))), "Translated_Titles")
  for (i in 1:length(text)) {
    body <- paste(" ",
                  "'q':", text[i], " ",
                  "'source':", source, " ",
                  "'target':", target, " ",
                  "'format':'text'",
                  " ", sep = " ")
    z <- httr::POST(url$google.translate, body = body)
    z <- jsonlite::fromJSON(rawToChar(z$content))
    df_output[i, "Translated_Titles"] <- z$data$translations$translatedText
  }
  return(df_output)
}
sample.text <- "Merci beaucoup!"
en.translator(KEYS(), sample.text)
```

```
## Translated_Titles
## 1 Thank you so much!
```

```
# create.news.url -- Serves as the foundation for df.article().
```

```
create.news.url <- function(url, date) {
  base <- "https://newsapi.org/v2/everything?sources=le-monde&pageSize=100&apiKey="
  start.date <- as.Date(date, format = "%Y-%m-%d")
  end.date <- start.date + 1
  url.start <- paste0("&from=", start.date)
  url.end <- paste0("&to=", end.date)
  url <- paste0(base, url$API_key.news, url.start, url.end, "&sortBy=popularity")
  return(url)
}
create.news.url(KEYS(), "2018-11-15")
```

```
## [1] "https://newsapi.org/v2/everything?sources=le-monde&pageSize=100&apiKey=c6c10679081f4b43ab4352aea936e71b&from=2018-11-15&to=2018-11-16&sortBy=popularity"
```

```
# df.article -- returns df of top 100 article on specified date. (Note: A function that returns a
# data.frame between any two dates, though instructive, is beyond the scope of this homework.)
df.article <- function(url, start.date) {
  url.input <- create.news.url(url, start.date)
  z <- httr::GET(url.input)
  z <- jsonlite::fromJSON(rawToChar(z$content))
  z <- z$articles
  return(z)
}
head(df.article(KEYS(), "2018-11-15"))
```

```
## source.id source.name author
## 1 le-monde Le Monde <NA>
## 2 le-monde Le Monde <NA>
## 3 le-monde Le Monde <NA>
## 4 le-monde Le Monde Le Monde
## 5 le-monde Le Monde Laure Gasparotto
## 6 le-monde Le Monde Adrien Pécout

## title
## 1 Schiappa favorable à une mesure anti-fessée bientôt examinée à l'Assemblée
## 2 Plus de 1 500 migrants de la « caravane » sont arrivés à la frontière avec les Etats-Unis
## 3 Californie : alors que le bilan s'alourdit, plus de 600 personnes portées disparues
## 4 Un cratère de la taille de Paris découvert sous la glace du Groenland
## 5 Bellevoye, le whisky tricolore qui compte triple
## 6 Les joueuses du XV de France se préparent à devenir semi-professionnelles

## description
## 1 La secrétaire d'Etat à l'égalité entre les femmes et les hommes s'est exprimée à ti
tre « personnel » sur cette proposition de loi, soutenant « le droit à une éducation sans violence ».
## 2 Quelque 3 000 autres arriveront dans la
soirée ou dans la nuit. Donald Trump a promis de les empêcher de rentrer sur le territoire américain.
## 3 Le « Camp Fire », qui sévit dans le nord de l'Etat de
puis le 8 novembre, a déjà fait 63 morts. La liste des portés disparus a, elle, été revue à la hausse
## 4 Un énorme
cratère a été découvert sous la glace du Groenland, ont annoncé des chercheurs danois le 14 novembre.
## 5 Au cœur des Landes, Les Bienheureux, soit Alexandre Sirech et son compère Jean Moueix, peaufinent leur produ
ction 100 % hexagonale. Sa particularité ? Une triple distillation issue d'une recherche approfondie.
## 6 La fédération propose, pour la saiso
n en cours, des contrats à mi-temps, inférieurs à 1 600 euros net par mois, hors primes de résultats.

## url
## 1 https://www.lemonde.fr/societe/article/2018/11/15/schiappa-favorable-a-une-mesure-anti-fessee-bientot-
examinee-a-l-assemblee_5384032_3224.html
## 2 https://www.lemonde.fr/ameriques/article/2018/11/15/mexique-plus-de-1-500-migrants-de-la-caravane-sont-
arrive-a-la-frontiere_5384106_3222.html
## 3 https://www.lemonde.fr/planete/article/2018/11/16/californie-alors-que-le-bilan-s-alourdit-plus-de-600-perso
nes-portees-disparues_5384164_3244.html
## 4 https://www.lemonde.fr/sciences/video/2018/11/16/un-cratere-de-la-taille-de-paris-decouvert-sous-la
-glace-du-groenland_5384699_1650684.html
## 5 https://www.lemonde.fr/vins/article/2018/11/16/bellevoys-le-whisky-tricolore
e-qui-compte-triple_5384493_3527806.html
## 6 https://www.lemonde.fr/rugby/article/2018/11/15/rugby-les-joueuses-du-xv-de-france-se-preparent-a-devenir-se
mi-professionnelles_5383797_1616937.html

## urlToImage
## 1 https://img.lemde.fr/2018/11/15/37/0/4683/2339/1440/720/60/0/a9f1c
4b_WFWerUH94RwoJ9HoysgHLZuY.jpg
## 2 https://img.lemde.fr/2018/11/15/92/0/5220/2610/1440/72
0/60/0/0ca4868_5744101-01-06.jpg
## 3 https://img.lemde.fr/2018/11/16/0/0/3000/1500/1440/720/60/0/1337d8b_ab7e7dc8d5454fb5b1b16243bf877010-ab7e7dc
8d5454fb5b1b16243bf877010-0.jpg
## 4 https://img.lemde.fr/2018/11/16/14/0/1920/960/644/322/60/0/d8de7
f3_TlyxsyBD6gYUr70N7RPl0qv5.jpg
## 5 https://img.lemde.fr/2018/11/14/1015/0/1813/903/1440/720/60/0/5f81ef6_2018111745.
0.2220228852bellevoys01_web.jpg
## 6 https://img.lemde.fr/2018/11/09/37/0/2888/1442/1440/72
0/60/0/3162756_5707424-01-06.jpg
## publishedAt
## 1 2018-11-15T16:18:29Z
## 2 2018-11-15T18:34:08Z
## 3 2018-11-16T03:28:50Z
## 4 2018-11-16T18:31:34Z
## 5 2018-11-16T13:02:23Z
## 6 2018-11-15T08:08:04Z

## content
## 1 Photographie d'illustration d'un enfant frappé par sa mère, le 1er mars 2015, à Paris. LOIC VENANCE / AFP La s
ecrétaire d'Etat à l'égalité entre les femmes et les hommes, Marlène Schiappa, s'est dite jeudi 15 novembre favorabl
e à titre « personnel » à une proposition de loi, soutenant « le droit à une éducation sans violence ».
## 2 Des migrants arrivant à Tijuana, à la frontière des Etats-Unis, le 15 novembre 2018. GUILLERMO ARIAS / AFP P
lus de 1 500 migrants de la « caravane » sont arrivés à Tijuana, à la frontière avec les Etats-Unis. Dans la mati
née, vingt-deux autobus transportant d... [+1838 chars]
```

```
## 3 A la recherche de restes humains dans les décombres de la ville de Paradise, ravagée par les flammes, le 15
novembre. Terry Chea / AP Le bilan humain du « Camp Fire », incendie qui ravage depuis plusieurs jours, le nord d
e la Californie, ne cesse de salourdir... [+1508 chars]
## 4 Large de 31 kilomètres, il est plus grand que Washington ou Paris. Cest surtout le premier cratère de ce typ
e jamais découvert au Groenland, et sous une couche de glace en général. Il figure parmi les 25 plus grands sur T
erre, selon létude publiée dans la rev... [+1055 chars]
## 5 Cest un petit village coquet, mais sans cachet particulier, à une petite cinquantaine de kilomètres de Borde
aux, dans les Landes de Gascogne. Un bourg tranquille dun peu plus de 700 habitants, habitué depuis des lustres a
u calme. Mais, ces dernières années, L... [+6339 chars]
## 6 Article réservé aux abonnés Lors du match France-Nouvelle-Zélande, le 9 novembre au stade Mayol de Toulon. C
HRISTOPHE SIMON / AFP Elles nen sont pas encore au stade où elles pourraient en vivre comme dun « métier » à temp
s plein, mais les choses bougent. Cet ... [+2662 chars]
```

```
# get.title -- returns translated titles that contain specified keyword.
get.title <- function(url, start.date, keyword) {
  df <- df.article(url, start.date)
  title <- df$title
  en.title <- en.translator(url, title)
  output <- dplyr::filter(en.title, grepl(keyword, Translated_Titles))
  return(output)
}
get.title(KEYS(), "2018-11-15", "California")
```

```
##
## 1 California: As the death toll rises, over 600 people are missing
## 2 California fires: electricity companies charged
## 3 California: the human toll is growing, fires are still raging
## 4 Fires in California: "I was wondering if I was going to find the road on the other side of the flames"
```

```
newspaper.text <- get.title(KEYS(), "2018-11-15", "California")[1,"Translated_Titles"]
newspaper.text
```

```
## [1] "California: As the death toll rises, over 600 people are missing"
```

```
read.english <- function(url, text) {
  input <- paste("'input':{'text':", text, "'", sep="")
  voice <- "'voice':{'languageCode':'en-GB','name':'en-GB-Standard-C','ssmlGender':'FEMALE'}"
  output <- "'audioConfig':{'audioEncoding':'MP3'}"
  s.body <- paste("{", input, ",", voice, ",", output, "}", sep="")

  z <- httr::POST(url$google.speech, body = s.body)
  z <- jsonlite::fromJSON(rawToChar(z$content))

  Sys.setenv(output_string = z$audioContent)
  system("echo $output_string > sound_out.txt")
  system("base64 --decode sound_out.txt > sound_out.mp3")
  system("afplay sound_out.mp3")
}
```