

Homework12

Larine Hamied

12/2/2018

```
library(httr)
library(lubridate)
library(base64enc)

newsAPI <- "apiKey=02f3b9bf72be490ca99636ca0921db7e"
googleCloudAPI <- "AIzaSyAy_zmaU5tKjkq-ApMUY6d7aDSq8FHGA4E"
```

I made my API Keys global variables above.

For this exercise, I will use the Arabic financial newspaper, ارقام, which can be transliterated as “Argaam” or “Arqaam” (and translates to “numbers” in English). I received this API from the website newsapi.org

Here is a **get_headlines_date_range** function, where the user must include the name of the source, a start date and end date (in any way (i.e. with a forward slash, hyphen, or no punctuation at all) so long as it is in the order year, month, day) that provides a range, and the API key, “newsAPI.”

```
get_headlines_date_range <- function(source, date1, date2, API_key){
  base <- "https://newsapi.org/v2/everything?"
  sources <- paste("sources=", source, sep="")
  date1 <- paste("from=", ymd(as.character(date1)), sep="")
  date2 <- paste("from=", ymd(as.character(date2)), sep="")
  url <- paste(base, sources, "&", date1, "&", date2, "&", API_key, sep="")
  x <- httr::GET(url=url)
  x <- jsonlite::fromJSON(rawToChar(x$content))
  articles <- x$articles
  return(unlist(articles$title))
}
```

Here is the function with the newspaper Argaam/Arqaam:

```
get_headlines_date_range("argaam","2018/11/30","20181202", newsAPI)
```

```
## [1] "مبيعات مذكرات ميشال أوباما تحقق أرقاماً قياسية خلال 15 يوماً - أرقام"
## [2] "روسيا تُعلن خطة إقامة مستوطنة على القمر بحلول عام 2040 - أرقام"
## [3] "مجلس الوحدة الاقتصادية العربية يتبنى \"رؤية عربية مشتركة للاقتصاد الرقمي\" - أرقام"
## [4] "زلزال قوي يضرب منطقة شمال ولاية \"ألاسكا\" الأمريكية وتحذيرات من \"تسونامي\" - أرقام"
## [5] "مخلوق بحري غامض في جسم كائن فضائي يخيف نيوزيلندا (صور) - أرقام"
## [6] "مسؤولون كوريون جنوبيون في الشمال لدرس إعادة ربط السكك الحديدية بين البلدين - أرقام"
## [7] "حالة الطقس المتوقعة اليوم الجمعة - أرقام"
## [8] "ولي العهد يلتقي دولة رئيس الوزراء الهندي - أرقام"
## [9] "شاهد.. \"المكافحة\" تكشف عن أحد أساليب تهريب المخدرات إلى المملكة - أرقام"
```

Here is the function with the newspaper Argaam/Arqaam (notice they are spelled differently - interestingly enough, the url would only work with Argaam above and Arqaam below). The keyword I used is “الوقود” or **fuel**, giving me headlines about fuel prices generally in the Middle East, and particularly news on the protests going on in France.

```
get_headlines_keyword <- function(source, keyword, API_key) {
  base <- "https://newsapi.org/v2/everything?"
  sources <- paste("sources=", source, sep="")
  keyword <- paste("q=", keyword, sep="")
  url <- paste(base, sources, "&", keyword, "&", API_key, sep="")
  x <- httr::GET(url=url)
  x <- jsonlite::fromJSON(rawToChar(x$content))
  articles <- x$articles
  return(unlist(articles$title))
}
```

Here is a **get_headlines_keyword** function, where the user must include the name of the source, a keyword, and the API key, “newsAPI.”

```
get_headlines_keyword("arqaam","الوقود", newsAPI)
```

```
## [1] "ربع مليون فرنسي يغلقون الشوارع احتجاجاً على ارتفاع أسعار الوقود"
## [2] "\"احتجاجات فرنسا: وفاة متظاهرة مع بدء حركة\"السترات الصفراء\"
## [3] "فرنسا: الشرطة تستخدم الغاز والمياه لتفريق احتجاجات\"السترات الصفراء\" في باريس"
## [4] "فرنسا: مئات الجرحى مع استمرار الاحتجاجات ضد رفع أسعار الوقود"
## [5] "كيف تخطط الهند لاستغلال شواطئها في إنتاج وقود نووي؟"
## [6] "حركة\"السترات الصفراء\".. اعتقال عشرات المتظاهرين في باريس"
## [7] "السودان: أزمة الوقود تطيح قيادات من وزارة النفط والبنك المركزي - العربي الجديد"
## [8] "فرنسا.. مصرع متظاهرة وإصابة العشرات باحتجاجات على فرض ضريبة الوقود - الغد الأردني"
## [9] "لبنان يمنع الوقود عن الطيران الإيراني والسوري - الفجر"
## [10] "السترات الصفراء\" تقود غضب الفرنسيين ضد زيادة أسعار الوقود - الاقتصادية\"
## [11] "مئات الجرحى باحتجاجات فرنسا وشعبية ماكرون في انهيار - الجزيرة"
## [12] "مقتل متظاهرة في حادث خلال احتجاجات بفرنسا - روسيا اليوم"
## [13] "أميركا قد توقف تزويد الوقود لطائرات التحالف بقيادة السعودية - العربي الجديد"
## [14] "احتجاجات الفرنسيين على زيادة ضرائب الوقود تهوي بشعبية ماكرون - هسبريس"
## [15] "واشنطن قد توقف تزويد الوقود في حرب اليمن - الحرة"
## [16] "باريس تشتعل مجدداً باحتجاجات رفع أسعار الوقود - المصري اليوم"
## [17] "باريس تستعد لموجة احتجاجات جديدة بسبب ارتفاع أسعار الوقود - اليوم السابع"
## [18] "احتجاجات ضد أسعار الوقود والضرائب في فرنسا - صحيفة الوثام الالكترونية"
## [19] "باريس.. تأهب أمني في مواجهة احتجاجات الوقود - العربية نت"
## [20] "شعبية ماكرون تتراجع مع استمرار الاحتجاج على زيادة ضرائب الوقود - إرم نيوز"
```

The user can store either of the functions above into an object. Let's do it below with the **get_headlines_keyword** function and store it into an object called **"headlines"**

```
headlines <- get_headlines_keyword("arqaam","الوقود", newsAPI)
```

Here is a **translate_arabic_headlines** function, where the user must pass through an object with whichever headlines they want to use, and an API key for Google Cloud (from above) to use Google Cloud's translate API.

```

translate_arabic_headlines <- function(headlines, API_key){
  source <- "ar"
  target <- "en"
  nh <- length(headlines)
  translated <- rep(NA, nh)
  for (i in 1:nh){
    text <- headlines[i]
    body <- paste("{",
                  "'q':", text, "'",
                  "'source':", source, "'",
                  "'target':", target, "'",
                  "'format':'text'",
                  "}", sep="")

    base <- "https://translation.googleapis.com/language/translate/v2?key="
    url <- paste(base, API_key, sep="")

    x <- httr::POST(url, body = body)
    x <- jsonlite::fromJSON(rawToChar(x$content))
    out <- x$data$translations$translatedText
    translated[i] <- out
  }
  return(translated)
}

```

Translating the headlines from above:

```

translate_arabic_headlines(headlines, googleCloudAPI)

```

[1] "A quarter of a million Frenchmen are closing the streets to protest rising fuel prices"

[2] "France protests: Death of a demonstrator with the start of the movement of \"yellow jackets\""

[3] "France: Police use gas and water to disperse protests of \"yellow jackets\" in Paris"

[4] "France: Hundreds injured as protests continue against fuel price hike"

[5] "How does India plan to exploit its shores to produce nuclear fuel?"

[6] "The movement of the \"yellow jackets\" .. The arrest of dozens of demonstrators in Paris"

[7] "Sudan: Fuel crisis ousts leaders from the Ministry of Oil and the Central Bank - the new Arab"

[8] "France .. Death of a demonstrator and injuring dozens of protests against the imposition of fuel tax - the Jordanian tomorrow"

[9] "Lebanon prevents fuel from Iranian and Syrian flights"

[10] "\"Yellow jackets\" lead the anger of the French against the increase in fuel prices - economic"

[11] "Hundreds injured by protests of France and Macron's popularity in the collapse - the island"

[12] "A demonstrator killed in an accident during protests in France - Russia today"

[13] "America has stopped supplying fuel to coalition aircraft led by Saudi Arabia - the new Arab"

[14] "Protests by the French over the increase in fuel taxes are infuriating the popularity of Macaron - Haspa ris"

[15] "Washington has stopped supplying fuel in the Yemen - free war"

[16] "Paris reacts again with fuel price protests - Egypt Today"

[17] "Paris is preparing for a new wave of protests due to rising fuel prices - the seventh day"

[18] "Protests against fuel prices and taxes in France - electronic harmony newspaper"

[19] "Paris: Security alert in the face of fuel protests"

[20] "Macron's popularity recedes as protests continue to increase fuel taxes - Trem News"

Here is a **text_to_speech** function, where the user must pass through a specific headline, and an API key for Google Cloud (from above) to use the Google Cloud Text to Speech API. This will read out whatever is translated above in US English in a male's voice, and return an mp3 file.

```
text_to_speech <- function(text, API_key) {

  text <- translate_arabic_headlines(text,API_key)

  text <- paste(text, collapse=" ")
  text <- gsub("'", "", text)

  input <- paste("'input':{'text':'", text, "'}", sep="")
  voice <- "'voice':{'languageCode':'en-US','name':'en-US-Standard-B','ssmlGender':'MALE'}"
  output <- "'audioConfig':{'audioEncoding':'MP3'}"

  s_body <- paste("{", input, ",", voice, ",", output, "}", sep="")

  base_url <- "https://texttospeech.googleapis.com/v1/text:synthesize?key="
  url <- paste(base_url, API_key, sep="")

  x <- httr::POST(url, body = s_body)
  x <- jsonlite::fromJSON(rawToChar(x$content))
  Sys.setenv(output_string=x$audioContent)
  system("echo $output_string > sound_out.txt")
  system("base64 --decode sound_out.txt > sound_out.mp3")
  system("afplay sound_out.mp3")
}
```

I arbitrarily picked the 8th headline because I was interested to hear if the Speech would pause accordingly with ellipsis (...).

```
text_to_speech(headlines[8], googleCloudAPI)
```