

Demographic estimates from the Facebook Marketing API

Connor Gilroy

2017-10-15

```
library(httr)
library(jsonlite)
library(stringr)
library(tidyverse)
library(yaml)
```

Setup: config and url

We need two pieces of information from the config file: an access token, and an ads account. Make sure the ads account number is prefixed with “act_”.

```
fb_cfg <- yaml.load_file("facebook_config.yml")
```

The url for accessing ‘reach’ estimates is based off of the ads account, so we construct it by pasting strings onto the base url

```
fb_url <- "https://graph.facebook.com/v2.10"  
fb_ads_url <- str_c(fb_url, fb_cfg$ad_account_id,  
                    "reachestimate", sep = "/")
```

Example 1: US population on Facebook

```
targeting_spec <- '{"geo_locations": {"countries": ["US"]}}'

fb_query <- list(
  access_token = fb_cfg$access_token,
  currency = "USD",
  optimize_for = "NONE",
  targeting_spec = targeting_spec
)

r <- GET(fb_ads_url, query = fb_query)
```

Example 2: Young men and women in Washington State

In the state of Washington, how many men and women between the ages of 20 and 30 are on Facebook?

To answer this, we'll use two targeting specs from JSON files.

```
ts1 <- read_file("targeting_specs/targeting_spec_01.json")  
ts2 <- read_file("targeting_specs/targeting_spec_02.json")
```

```
cat(ts1)
```

```
## {  
##   "geo_locations": {  
##     "regions": [{  
##       "key": "3890",  
##       "name": "Washington"  
##     }]  
##   },  
##   "genders": [1],  
##   "age_min": 20,  
##   "age_max": 30  
## }
```

```
make_fb_ads_request <- function(ts) {  
  # Avoid rate limiting!  
  # Don't make too many requests in a short period of time.  
  Sys.sleep(5)  
  
  fb_query <- list(  
    access_token = fb_cfg$access_token,  
    currency = "USD",  
    optimize_for = "NONE",  
    targeting_spec = minify(ts)  
  )  
  
  GET(fb_ads_url, query = fb_query)  
}
```

```
r1 <- make_fb_ads_request(ts1)
r2 <- make_fb_ads_request(ts2)
```


We need to combine the request information from the targeting spec with the estimated number of users from the response.

As with the requests, we'll do this using a function.

```
process_fb_response <- function(ts, r) {  
  # targeting spec  
  ts_df <-  
    fromJSON(ts) %>%  
    as_data_frame() %>%  
    unnest(geo_locations)  
  
  # response  
  r_df <-  
    content(r, as = "parsed")$data %>%  
    as_data_frame() %>%  
    select(users)  
  
  bind_cols(ts_df, r_df)  
}
```

```
bind_rows(  
  process_fb_response(ts1, r1),  
  process_fb_response(ts2, r2)  
)
```

genders	age_min	age_max	key	name	users
1	20	30	3890	Washington	820000
2	20	30	3890	Washington	760000

Exercises

Exercise 1: Pick 3 other countries and report the number of Facebook users in each. Compare these numbers to the actual populations from the World Bank or some other source.

Use two-digit country codes for countries:

https://en.wikipedia.org/wiki/ISO_3166-1_alpha-2

Exercise 2: Pick another US state and age range, and compare the numbers of men and women. You can look up the key that Facebook uses for each state in the file provided.

Challenge exercise: Get the Facebook user population for each US state. Compare these estimates to population estimates from the Census ACS. Note that you'll need to make a separate call to the API for each state, which will take several minutes.