

Case Studies Using Social Media Marketing APIs

Ridhi Kashyap

Department of Sociology
Leverhulme Centre for Demographic Science
Nuffield College
University of Oxford

SICSS-Duke
June 21, 2022

Accessing Digital Trace Data

- ▶ Screen scraping
- ▶ Querying data from application programming interfaces (APIs)
- ▶ Archives or data sharing platforms (e.g. Google Trends).
- ▶ Data sharing agreements with data owners.

Screen Scraping

- ▶ The process of extracting data from web pages in an automated way
- ▶ Lots of data are available on the web – but not all of it is appropriate to use for research.

Considerations when Screen Scraping

- ▶ Read the website's Terms of Service: are you allowed to do this?
- ▶ Larger websites like Facebook, Instagram, NY Times do not allow these practices – but some provide structured access to data through APIs.
- ▶ Even if terms of service does not prohibit this – are the data sensitive? Could the use of the data harm in some way? Is their use ethical?

A Simple Example: Wikipedia Page

World Bank Group (2020) [edit]

Data of the World Bank Group for 2020.^{[3][4][5]} The values in the World Bank Group tables are rounded. All calculations were done on raw data, therefore, due to the nuances of rounding places illusory inconsistencies of indicators arose, with a size of 0.01 year. Since this entry is prone to vandalism and political bias it is recommended to check the original source.

	Countries	all	male	female	gender gap	Δ 2019 all	Δ 2019 male	Δ 2019 female	Δ 2019 gen. gap
		♦	♦	♦	♦	♦	♦	♦	♦
1	Japan	84.62	81.64	87.74	6.10	0.26	0.23	0.29	0.06
2	Singapore	83.74	81.50	86.10	4.60	0.15	0.10	0.20	0.10
3	South Korea	83.43	80.50	86.50	6.00	0.20	0.20	0.20	0.00
4	Norway	83.21	81.60	84.90	3.30	0.25	0.30	0.20	-0.10
5	Australia	83.20	81.20	85.30	4.10	0.30	0.30	0.30	0.00
6	Switzerland	83.10	81.10	85.20	4.10	-0.80	-1.00	-0.60	0.40
7	Iceland	83.07	81.70	84.50	2.80	-0.10	0.00	-0.20	-0.20
8	Israel	82.70	80.70	84.80	4.10	-0.10	-0.30	0.10	0.40
9	Malta	82.65	80.80	84.60	3.80	-0.20	-0.40	0.00	0.40
10	Sweden	82.41	80.70	84.20	3.50	-0.70	-0.80	-0.60	0.20
11	Italy	82.34	80.10	84.70	4.60	-1.15	-1.30	-1.00	0.30

Figure:

https://en.wikipedia.org/wiki/List_of_countries_by_life_expectancy

Wikipedia Page HTML

```
1 <!DOCTYPE html>
2 <html class="client-nojs" lang="en" dir="ltr">
3 <head>
4 <meta charset="UTF-8"/>
5 <title>List of countries by life expectancy - Wikipedia</title>
6 <script>document.documentElement.className="client-js";RLCONF={"wgBreakFrames":false,"wgSeparatorTransformTable":[],"wgDigitTransformTable":[]}
7 "Life expectancy","wgPageContentLanguage":"en","wgPageContentModel":"wikitext","wgRelevantPageName":"List_of_countries_by_life_expectancy","wgRev...
8 RLSTATE="ext.globalCsJs.user.styles":"ready","site.styles":"ready","user.styles":"ready","ext.globalCsJs.user":"ready","user":"ready","user.optio...
9 "ext.gadget.switcher","ext.centralauth.centralautologin","ext.popups","ext.uls.compactlinks","ext.uls.interface","ext.growthExperiments.SuggestedEdit...
10 <script>(RlQ>window.RlQ||[]).push(function(){mw.loader.implement("user.options#21251",function($,jQuery,require,module){mw.user.tokens.set({"patrolT...
11 <link rel="stylesheet" href="/w/load.php?lang=en&modules=ext.cite.styles&7Cext.uls.interlanguage&7Cext.visualEditor.desktopArticleTarget.noscri...
12 <script async="" src="/w/load.php?lang=en&modules=startup&only=scripts&raw=1&map=skin=vector"></script>
13 <meta name="ResourceLoaderDynamicItems" content="">
14 <link rel="stylesheet" href="/w/load.php?lang=en&modules=site.styles&only=styles&skin=vector"/>
15 <link rel="stylesheet" href="/w/load.php?lang=en&modules=site.styles&skin=vector"/>
16 <meta name="generator" content="MediaWiki 1.39.0-wmf.16"/>
17 <meta name="referrer" content="origin"/>
18 <meta name="referrer" content="origin-when-crossorigin"/>
19 <meta name="referrer" content="origin-when-cross-origin"/>
20 <meta name="format-detection" content="telephone=no"/>
21 <meta property="og:image" content="https://upload.wikimedia.org/wikipedia/commons/thumb/5/52/Comparison_gender_life_expectancy_WHO.svg/1200px-Compar...
22 <meta property="og:image:height" content="954"/>
23 <meta property="og:image" content="https://upload.wikimedia.org/wikipedia/commons/thumb/5/52/Comparison_gender_life_expectancy_WHO.svg/800px-Compar...
24 <meta property="og:image:width" content="800"/>
25 <meta property="og:image:height" content="636"/>
26 <meta property="og:image" content="https://upload.wikimedia.org/wikipedia/commons/thumb/5/52/Comparison_gender_life_expectancy_WHO.svg/640px-Compar...
27 <meta property="og:image:width" content="640"/>
28 <meta property="og:image:height" content="509"/>
29 <meta property="og:title" content="List of countries by life expectancy - Wikipedia"/>
30 <meta property="og:type" content="website"/>
31 <link rel="preconnect" href="//upload.wikimedia.org"/>
32 <link rel="alternate" media="only screen and (max-width: 720px)" href="//en.m.wikipedia.org/wiki/List_of_countries_by_life_expectancy"/>
33 <link rel="alternate" type="application/x-wiki" title="Edit this page" href="/w/index.php?title=List_of_countries_by_life_expectancy&action=edit...
34 <link rel="apple-touch-icon" href="/static/apple-touch/wikipedia.png"/>
35 <link rel="shortcut icon" href="/static/favicon/wikipedia.ico"/>
36 <link rel="search" type="application/opensearchdescription+xml" href="/w/opensearch_desc.php" title="Wikipedia (en)"/>
37 <link rel="EditURI" type="application/rsd+xml" href="https://en.wikipedia.org/w/api.php?action=rsd"/>
38 <link rel="license" href="https://creativecommons.org/licenses/by-sa/3.0/"/>
39 <link rel="canonical" href="https://en.wikipedia.org/wiki>List_of_countries_by_life_expectancy"/>
40 <link rel="dns-prefetch" href="//meta.wikimedia.org" />
41 <link rel="dns-prefetch" href="//login.wikimedia.org" />
42 </head>
43 <body class="mediawiki ltr sitedir-ltr mw-hide-empty-elt ns-0 ns-subject mw-editable page-List_of_countries_by_life_expectancy rootpage-List_of_coun...
44 <div id="mw-head-base" class="noprint"></div>
45 <div id="content" class="mw-body" role="main">
46 <a id="top"></a>
47 <div id="mw-navigation"><!-- CentralNotice --></div>
```

A Simple Example: Wikipedia Page

```
library(rvest)
le_wiki <- read_html("https://en.wikipedia.org/wiki/List_of_countries_by_life_expectancy")
le_table <- html_node(le_wiki, xpath = '//*[@id="mw-content-text"]/div/table[1]')
le_rankings <- html_table(le_table1)
head(le_rankings)
```

...

A tibble: 6 × 9

Countries <chr>	all <dbl>	male <dbl>	female <dbl>	gendergap <dbl>	Δ 2019all <dbl>	Δ 2019male <dbl>	Δ 2019female <dbl>	▶
	NA	NA	NA	NA	NA	NA	NA	NA
Japan	84.62	81.64	87.74	6.1	0.26	0.23	0.29	
Singapore	83.74	81.50	86.10	4.6	0.15	0.10	0.20	
South Korea	83.43	80.50	86.50	6.0	0.20	0.20	0.20	
Norway	83.21	81.60	84.90	3.3	0.25	0.30	0.20	
Australia	83.20	81.20	85.30	4.1	0.30	0.30	0.30	

6 rows | 1–8 of 9 columns

A Simple Example: Wikipedia Page

The screenshot shows a Chrome browser window with the URL <https://en.wikipedia.org/>. The page content discusses life expectancy and its variability across countries. The 'View' menu is open, and the 'Developer' option is highlighted with a tooltip showing sub-options: View Source, Developer Tools, Inspect elements, JavaScript Console, and Allow JavaScript from Apple Events.

	Countries	all	male	female	gender gap	Δ 2019 all	Δ 2019 male	Δ 2019 female	Δ 2019 gen. gap
	#	#	#	#	#	#	#	#	#
1	Japan	84.62	81.64	87.74	6.10	0.26	0.23	0.29	0.06
2	Singapore	83.74	81.50	86.10	4.60	0.15	0.10	0.20	0.10
3	South Korea	83.43	80.50	86.50	6.00	0.20	0.20	0.20	0.00
4	Norway	83.21	81.60	84.90	3.30	0.25	0.30	0.20	-0.10
5	Australia	83.20	81.20	85.30	4.10	0.30	0.30	0.30	0.00
6	Switzerland	83.10	81.10	85.20	4.10	-0.80	-1.00	-0.60	0.40
7	Iceland	83.07	81.70	84.50	2.80	-0.10	0.00	-0.20	-0.20
8	Israel	82.70	80.70	84.80	4.10	-0.10	-0.30	0.10	0.40
9	Malta	82.65	80.80	84.60	3.80	-0.20	-0.40	0.00	0.40

A Simple Example: Wikipedia Page

- the average number of years that a person can expect to live in "full health".^{[1][2]}

Comparing life expectancies across countries can be problematic. For example, due to poor reporting in some countries and various local standards in collecting statistics. This is especially true for Healthy life expectancy, the definition of which criteria may change over time, even within a country.

World Bank Group (2020) [edit]

Data of the World Bank Group for 2020.^{[3][4][5]} The values in the World Bank Group tables are rounded. All

table.wikitable.sortable.static-row-numbers.plainrowheader 701 x 7213.16 ue to the nuances of rounding, in some places illusory
s.srn-white-background.jcu... 0.01 year. Since this entry is prone to vandalism and political
rce.

	Countries	all	male	female	gender gap	Δ 2019 all	Δ 2019 male	Δ 2019 female	Δ 2019 gen. gap
		♦	♦	♦	♦	♦	♦	♦	♦
1	Japan	84.62	81.64	87.74	6.10	0.26	0.23	0.29	0.06
2	Singapore	83.74	81.50	86.10	4.60	0.15	0.10	0.20	0.10
3	South Korea	83.43	80.50	86.50	6.00	0.20	0.20	0.20	0.00
4	Norway	83.21	81.60	84.90	3.30	0.25	0.30	0.20	-0.10
5	Australia	83.20	81.20	85.30	4.10	0.30	0.30	0.30	0.00
6	Switzerland	83.10	81.10	85.20	4.10	-0.80	-1.00	-0.60	0.40

The screenshot shows a context menu open over a table cell in a browser's developer tools. The menu includes options like 'Copy' and 'Copy Xpath'. The background shows a table with data from the World Bank Group (2020) page.

Elements Console Sources Network > 1 1

Copy

Styles Computed

Filter

element.style { text-align: right; }

.mw-parser-output counter-reset:;

.mw-parser-output background: #e0e0e0;

.wikitable { background-color: #f9f9f9; color: black; margin: 1em 0; border: 1px solid #a2a9b1; border-collapse: collapse; }

table { }

Console What's New

Highlights from the Chrome 102 update

Some Limits

- ▶ This was a simple example. Often web pages are complex with many different elements.
- ▶ This yields messy data.
- ▶ Screen scraping can be frustrating or unfeasible. What then?
 - ▶ For complex web pages, crowd worker platforms (e.g. Mechanical Turk) could be an option.
 - ▶ Some web data can be accessed via APIs.

What is an API?

- ▶ Application programming interfaces or APIs are a software intermediary that allows two applications to talk to each other.
- ▶ Web APIs allow one computer (a client) to ask another computer (a server) for some resource over the internet.
- ▶ APIs provide a structured way to access data that are stored in databases that are continuously updated.

Application Programming Interface

- ▶ Modern APIs adhere to standards, that make data exchange programmatically accessible, safe and structured
- ▶ Contrast with web scraping.
- ▶ Two important concepts when using APIs
 - ▶ Credentialling
 - ▶ Rate limiting

API Directory

The screenshot shows the ProgrammableWeb API Directory homepage. At the top, there's a navigation bar with links for "LEARN ABOUT APIs", "WHAT IS AN API?", "TUTORIALS", "API CHARTS & RESEARCH", "ADD API & MORE", and social sharing icons. A search bar is also present. Below the navigation is a banner for "MuleSoft Applying DevOps to APIs". The main content area features a heading "Search the Largest API Directory on the Web" with a search input field and a "SEARCH APIs" button. To the right, there's a sidebar for "API UNIVERSITY" with sections for "FEATURED" and "LATEST" news, and a "FOR API PROVIDERS" section with links to "What Are APIs and How Do They Work?", "8 Real World API Strategies and the Keys to Their Success", and "Microservices 101: Understanding and Leveraging Microservices". There's also a "More for API Providers" link. The main content area lists several APIs with their names, descriptions, categories, and submission dates:

API Name	Description	Category	Submitted
Google Maps	[This API is no longer available. Google Maps' services have been split into multiple APIs, including the Static Maps API.]	Mapping	12.05.2005
Twitter	[This API is no longer available. It has been split into multiple APIs, including the Twitter Ads API, Twitter Search Tweets...]	Social	12.08.2006
YouTube	The Data API allows users to integrate their program with YouTube and allows it to perform many of the operations available on the website. It provides the capability to search for videos, retrieve...	Video	02.08.2006
Flickr	The Flickr API can be used to retrieve photos from the Flickr photo sharing service using a variety of feeds - public photos and videos, favorites, friends, group pools, discussions, and more. The...	Photos	09.04.2005
Facebook	[This API is no longer available. Its functions have been split among the following APIs: Facebook Ads, ...]	Social	08.16.2006

On the right side, there's a "Today in APIs" newsletter sign-up form with fields for "Email Address" and a "SUBSCRIBE" button.

Figure: <https://www.programmableweb.com/apis/directory>

httr in R

- ▶ The `httr` package in R is useful to work with APIs
- ▶ You can make a request (to a url) and get a response.
- ▶ Response contains a status, header and body.

Social Media Ad Audience Estimates

- ▶ **Digital census:** How many users of 'x' characteristics (age, gender, location, etc) are on a given platform?

The screenshot shows the Facebook Ads Manager interface. On the left, there's a sidebar with navigation links: Campaign, Ad Account, Ad Set (selected), and Ad. Under Ad Set, Audience is highlighted. The main area shows targeting parameters for an ad set targeting Nigeria. A red box highlights the 'Audience Size' section, which includes a gauge from 'Specific' to 'Broad' (set to 'Broad') and text stating 'Your audience selection is fairly broad.' Below this, it says 'Potential Reach: 15,000,000 people'. Another red box highlights the 'Estimated Daily Results' section, which shows a reach range of '16K - 46K'. At the bottom, there are fields for 'Age' (18 - 65+), 'Gender' (All, Men selected, Women), and 'Detailed Targeting' (All demographics, interests and behaviors). A 'Show More Options' button is at the very bottom.

Figure: Facebook ads manager: <https://www.facebook.com/adsmanager>

Social Media Ad Audience Estimates

Facebook Ads Manager interface showing audience targeting options for a campaign.

Campaign: Ridhi Kashyap (280850...)

Ad Account: Create New

Ad Set: Page, Audience, Placements, Budget & Schedule

Ad: Identity, Format, Media, Text & Links, Languages, Tracking

Locations: Everyone in this location (Nigeria)

Audience Size: Your audience selection is fairly broad. Potential Reach: 9,100,000 people.

Estimated Daily Results: Reach: 17K - 48K

Targeting Options:

- Age: 18 - 65+ (highlighted with a red box)
- Gender: All, Men, Women (Women is selected)
- Detailed Targeting: All demographics, interests and behaviors
- Show More Options

The accuracy of estimates is based on factors like past campaign data, the budget you entered and market data. Numbers are provided to give you an idea of performance for your budget, but are only estimates and don't guarantee results.

Were these estimates helpful?

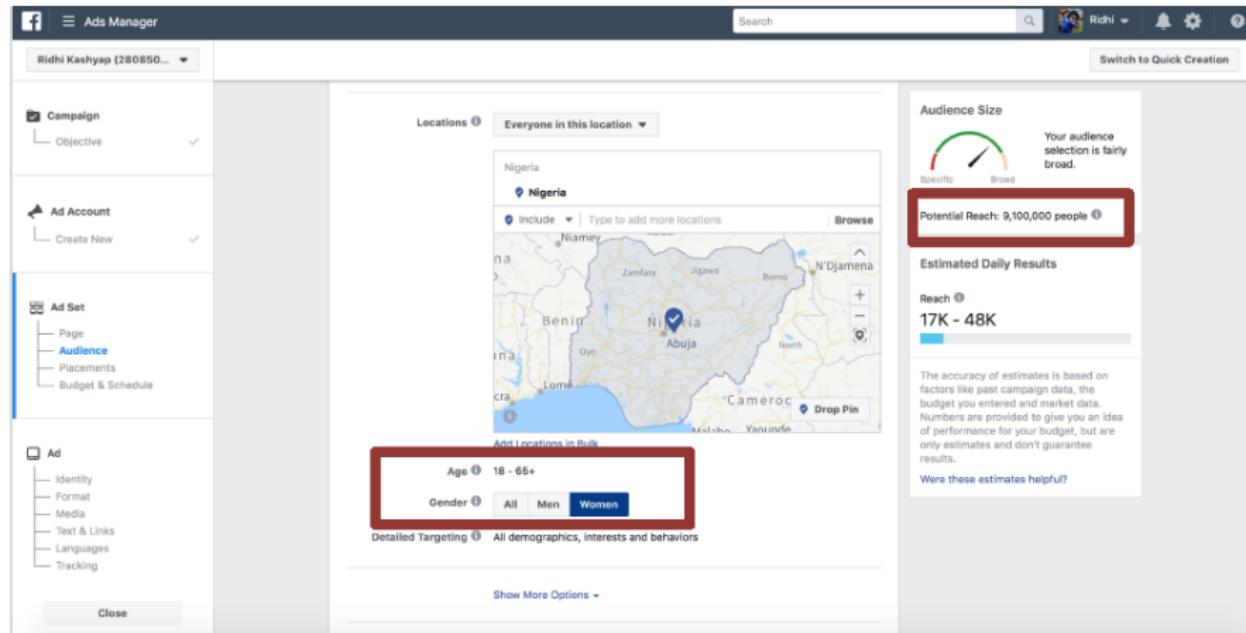


Figure: Facebook ads manager: <https://www.facebook.com/adsmanager>

Social Media Ad Audience Estimates: Updated MAU

New Campaign > New Ad Set > 1 Ad

[Edit](#) [Review](#)

Locations

People living in or recently in this location

Nigeria

Nigeria

Include Search locations Browse

NEGAL
Bamako
GUINEA
BURKINA FASO
BENIN
NIGERIA
Lagos
CENTRAL AFRICAN REP.
S. SUD.
AKRY
LIBERIA
GHANA
CAMEROON
BANGUI
GABON
CONGO
DEI
CC
Drop pin NC

Add locations in bulk

Audience definition

Your audience selection is fairly broad.

Specific Broad

Estimated audience size: 12,200,000 - 14,300,000 i

Estimates may vary significantly over time based on your targeting selections and available data.

Estimated daily results

Estimated daily results aren't available for this campaign since it has a budget optimized across ad sets.

Figure: Facebook ads manager: <https://www.facebook.com/adsmanager>

Social Media Ad Audience Estimates

The screenshot shows the LinkedIn Ads Manager interface for setting audience targeting criteria. A red box highlights the 'Your estimated target audience' section, which displays '4,500,000+ LinkedIn members'. Another red box highlights the 'What gender do you want to target?' section, where 'Male' is selected. The interface includes fields for targeting by location ('Netherlands') and gender, and various targeting options like company name, industry, size, job title, function, seniority, education, fields of study, degrees, and member skills.

Use a matched audience (optional)
Custom targeting options to reach your website visitors, contacts, and target accounts. [Select](#)

Target by the audience below

What location do you want to target? (required)

include [See full list](#)

include Netherlands

Target people who permanently live or work in the selected location(s).
 Deliver ads to people who reside in the selected **location(s)** and are not recent visitors

What gender do you want to target?

All
 Female
 Male

Select specific targeting criteria to zero in on your ideal audience:

Company name	Company industry	Company size	Job title	Job function
Job seniority	Member schools	Fields of study	Degrees	Member skills

Figure: LinkedIn ads manager

New campaign

None of your ads are running - Your campaigns and ad groups are paused or removed. Enable them to begin showing your ads.

[LEARN MORE](#)

1 Create your campaign 2 Confirmation

Edit targeted demographics

DONE

Gender	Age	Parental status	Household income
<input checked="" type="checkbox"/> Female	<input checked="" type="checkbox"/> 18 - 24	<input checked="" type="checkbox"/> Not a parent	<input checked="" type="checkbox"/> Top 10%
<input checked="" type="checkbox"/> Male	<input checked="" type="checkbox"/> 25 - 34	<input checked="" type="checkbox"/> Parent	<input checked="" type="checkbox"/> 11 - 20%
<input checked="" type="checkbox"/> Unknown <small>(?)</small>	<input checked="" type="checkbox"/> 35 - 44	<input checked="" type="checkbox"/> Unknown <small>(?)</small>	<input checked="" type="checkbox"/> 21 - 30%
	<input checked="" type="checkbox"/> 45 - 54		<input checked="" type="checkbox"/> 31 - 40%
	<input checked="" type="checkbox"/> 55 - 64		<input checked="" type="checkbox"/> 41 - 50%
	<input checked="" type="checkbox"/> 65+		<input checked="" type="checkbox"/> Lower 50%
	<input checked="" type="checkbox"/> Unknown <small>(?)</small>		<input checked="" type="checkbox"/> Unknown <small>(?)</small>

Your targeting's reach (?)

Impressions
10B+

What's defining your reach (?) ▼

Your weekly estimates (?)

Enter a bid and budget to see your estimated performance

⚠ Note: Household income targeting is only available in select countries. [Learn more](#)

Underlying Data: Facebook Marketing API

- ▶ You need:
 - ▶ Facebook account
 - ▶ Marketing app with token and an ad account number ("act"): these are the credentials to make calls to the API
 - ▶ Steps for obtaining these credentials are at: https://github.com/ridhi-kashyap/SICSS_Digital_Trace_2022/blob/main/Steps_for_creating_FB_Access_Token.pdf

Facebook Marketing API

- ▶ We can programmatically make a query to the API to retrieve these ad audience estimates.
- ▶ For more information on how targeting specifications work – see <https://developers.facebook.com/docs/marketing-api/buying-api/targeting/>
- ▶ To search for available targeting options – see <https://developers.facebook.com/docs/marketing-api/targeting-search>

Facebook Marketing API

```
#Loading required packages
library(readr)
library(dplyr)
library(jsonlite)
library(httr)

#Specify version of the API
version <- "v13.0"

#Specify your authentication/credentials,
#these should be saved separately in a private file

credential<-read.csv("credentials.csv",header = FALSE)

token <- credential$V1
act <- credential$V2

#We specify the basic URL as a string
Credentials <- paste0('https://graph.facebook.com/',version,'/act_',act,'/delivery_estimate?acces
s_token=',token)
```

- ▶ **version** refers to the version of the API (v 13.0 is current)
- ▶ **act** is the ad account number
- ▶ **token** is the access token
- ▶ Important to remember rate limiting when working with APIs and making multiple calls (Sys.sleep)

Facebook Marketing API: Basic Targeting Spec

```
targeting_spec_simple <- '{"geo_locations":{"countries":["GB"]}}'  
  
query_char <- list(  
  include_headers="false",  
  method="get",  
  optimization_goal="REACH",  
  suppress_http_code=1,  
  targeting_spec = targeting_spec_simple)
```

- ▶ Targeting specifications need to be specified in a JSON array
 - but we are treating it as a string here
- ▶ This is a call for one country's audience estimates.
- ▶ Other arguments for the query: we want the reach estimate, which is obtained via an HTTP get request

Facebook Marketing API: Requesting an Estimate

```
query_val1 <- GET(url = Credentials, query = query_char) %>% content(as="text",encoding =  
"UTF-8") %>% fromJSON  
query_val1<-query_val1$data  
query_val1  
#The query provides three counts - 1. estimated daily active users (dau)  
#                                two monthly active user  
#                                2. mau_upper_bound  
#                                3. mau_lower_bound  
  
query_val1$estimate_dau  
query_val1$estimate_mau_lower_bound  
query_val1$estimate_mau_upper_bound  
...
```

- ▶ We use GET from the `httr` package to obtain the estimates.
- ▶ These are returned as a JSON object.
- ▶ We extract them using the command `fromJSON` from `JSONlite`

Facebook Marketing API: Output JSON

```
{  
  "data": [  
    {  
      "daily_outcomes_curve": [  
        {  
          "spend": 0,  
          "reach": 0,  
          "impressions": 0,  
          "actions": 0  
        }  
      ],  
      "estimate_dau": 46394479,  
      "estimate_mau_lower_bound": 45500000,  
      "estimate_mau_upper_bound": 53600000,  
      "estimate_ready": true  
    }  
  ]  
}
```

- ▶ If we put the query into a browser this is what we get
- ▶ This is a JSON object.

Facebook Marketing API: Output JSON

```
query_val1 <- GET(url = Credentials, query = query_char) %>% content(as="text",encoding = "UTF-8") %>% fromJSON  
query_val1<-query_val1$data  
query_val1
```

```
##   daily_outcomes_curve estimate_dau estimate_mau_lower_bound  
## 1          0, 0, 0      46394479           45500000  
##   estimate_mau_upper_bound estimate_ready  
## 1                  53600000            TRUE
```

```
#The query provides three counts - 1. estimated daily active users (dau)  
#                                two monthly active user  
#  
#                                2. mau_upper_bound  
#                                3. mau_lower_bound  
  
query_val1$estimate_dau
```

```
## [1] 46394479
```

```
query_val1$estimate_mau_lower_bound
```

```
## [1] 45500000
```

```
query_val1$estimate_mau_upper_bound
```

```
## [1] 53600000
```

Facebook Marketing API: Including Age as a Targeting Param

```
target_query <- paste0('{"age_min":', age_min,  
                      ', "age_max":', age_max,  
                      ', "genders":[', genders, ']',  
                      ', "geo_locations": {"countries":[', countries, '], "location_types": ["home", "recent  
                    "]}}')
```

- ▶ We can add age through the age (min, max) arguments.
- ▶ For making multiple calls, e.g. multiple age groups and countries, it is helpful to write a function and loop.
- ▶ See example at: https://github.com/ridhi-kashyap/SICSS_Digital_Trace_2022

Facebook Marketing Search API: Requesting Available Countries

```
search_url <- "https://graph.facebook.com/v13.0/search"  
country_search <- GET(url = search_url,  
  query=list(  
    type='adgeolocation',  
    location_types='country',  
    access_token=token,  
    limit=1000)) %>%content(as="text", encoding="UTF-8")%>%fromJSON
```

- ▶ We can use the Facebook Marketing Search API to get obtain tables of targeting specifications (e.g. demographics, behaviours).
- ▶ The URL for this is
<https://graph.facebook.com/v13.0/search>

Applications

- ▶ Predicting gender gap indicators:
 - ▶ Internet and mobile access gender gaps and digital skills gender gaps (using Facebook, Google AdWords)
 - ▶ Professional gender gaps (using LinkedIn)
- ▶ Monitoring population displacement in Ukraine during the war

Nowcasting Digital Gender Gaps

Monthly Report
2022-05

latest Internet GG - Online Share Download

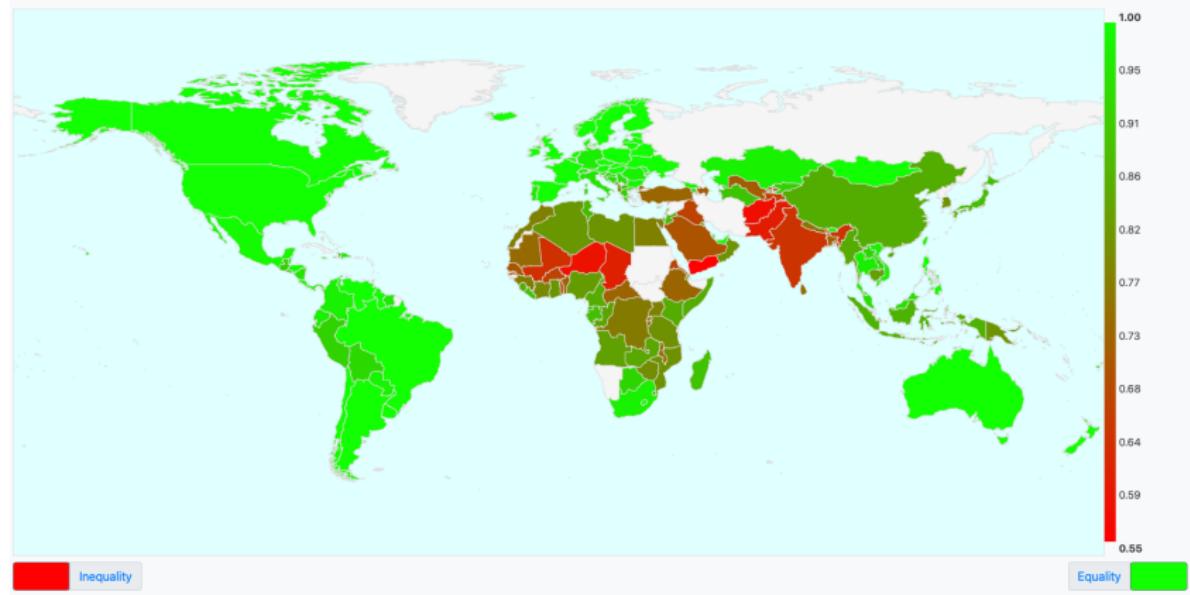
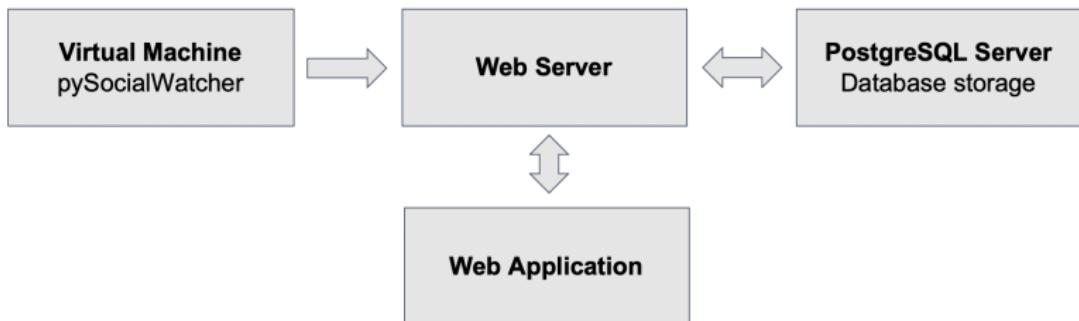


Figure: Gender gaps in Internet access predicted using Facebook gender gap index available at www.digitalgendergaps.org.

¹Fatehkia, Masoomali, Ridhi Kashyap, and Ingmar Weber. "Using Facebook ad data to track the global digital gender gap." *World Development* 107 (2018): 189-209.

Backend Setup



LinkedIn Gender Gap Index

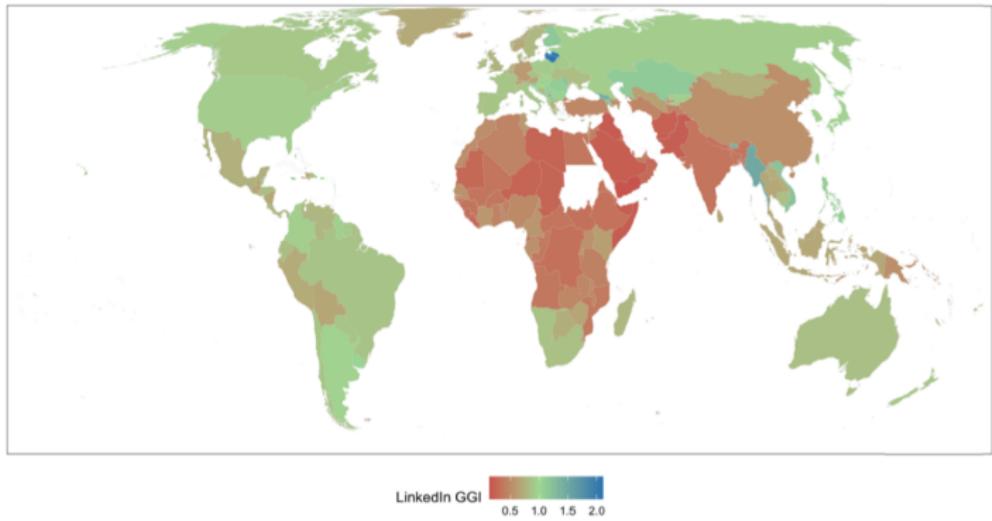


Figure: LinkedIn gender gap index (GGI) (female-to-male ratio of LinkedIn users), across the world

¹Kashyap, Ridhi and Florianne Verkroost "Analysing global professional gender gaps using LinkedIn advertising data." *EPJ Data Science* 10(1): 39

Professional Gender Gaps

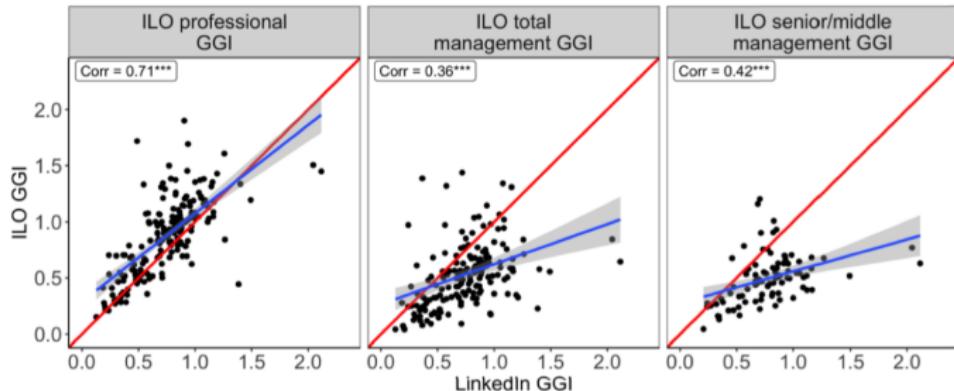
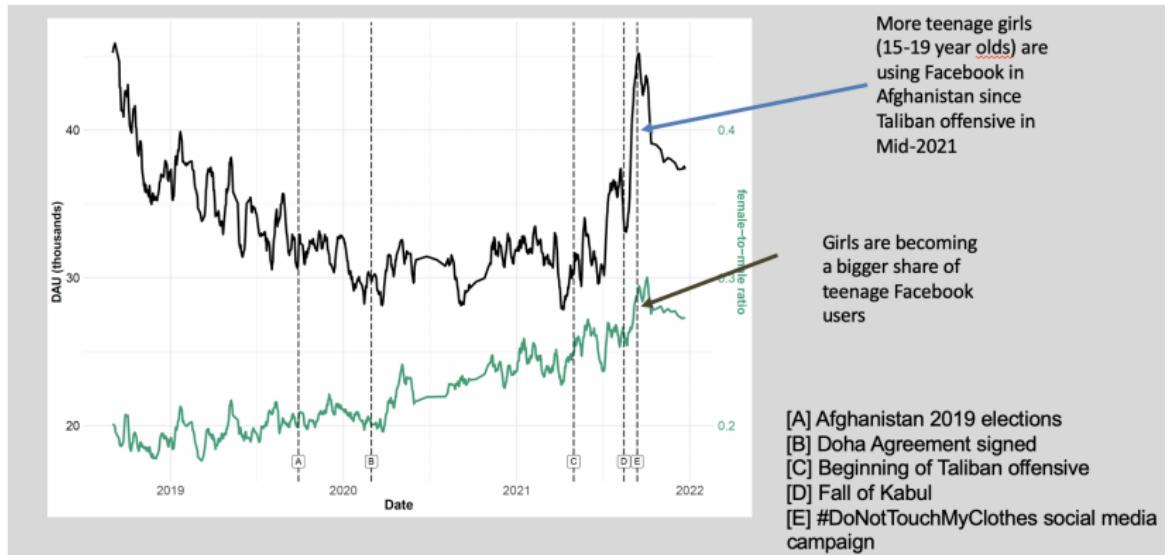


Figure: LinkedIn Gender Gap Index (GGI) and its correlations with different International Labour Organization (ILO) Gender Gaps

¹Kashyap, Ridhi and Florianne Verkroost "Analysing global professional gender gaps using LinkedIn advertising data." *EPJ Data Science* 10(1): 39

Further applications

- ▶ Tracking changes during crises, e.g. Afghanistan



Ukraine

February 24, 2022



Objectives

- ▶ Can we use FB marketing API data to monitor population displacement within Ukraine during the war?
- ▶ Produce daily sub-national population estimates inside Ukraine disaggregated by age and sex.
- ▶ Help fill critical data gaps to inform humanitarian response efforts.

Executive Summary
Introduction
Methods
Results
Discussion
Acknowledgements
License
References
Supplementary Material

Ukraine Crisis: Monitoring population displacement through social media activity

2022-06-06

Douglas R. Leasure^{1,3,6,*}, Ridhi Kashyap^{1,3}, Francesco Rampazzo^{2,1,3}, Benjamin Elbers^{3,1}, Claire Dooley^{4,6}, Ingmar Weber⁵, Masoomali Fatehkia², Maksym Bondarenko⁶, Mark Verhagen¹, Arun Frey¹, Jianyi Yan¹, Evelina T. Akimova¹, Robert Trigwell¹, Brian McDonald⁷, Mohamed Bakr⁷, Alessandro Sorichetta⁶, Andrew J. Tatem⁶, Melinda C. Mills^{1,3}

¹ Leverhulme Centre for Demographic Science, Department of Sociology, University of Oxford

² Said Business School, University of Oxford

³ Nuffield College, University of Oxford

⁴ Department of Population Health, London School of Hygiene and Tropical Medicine

⁵ Qatar Computing Research Institute, Hamad bin Khalifa University

⁶ WorldPop, University of Southampton

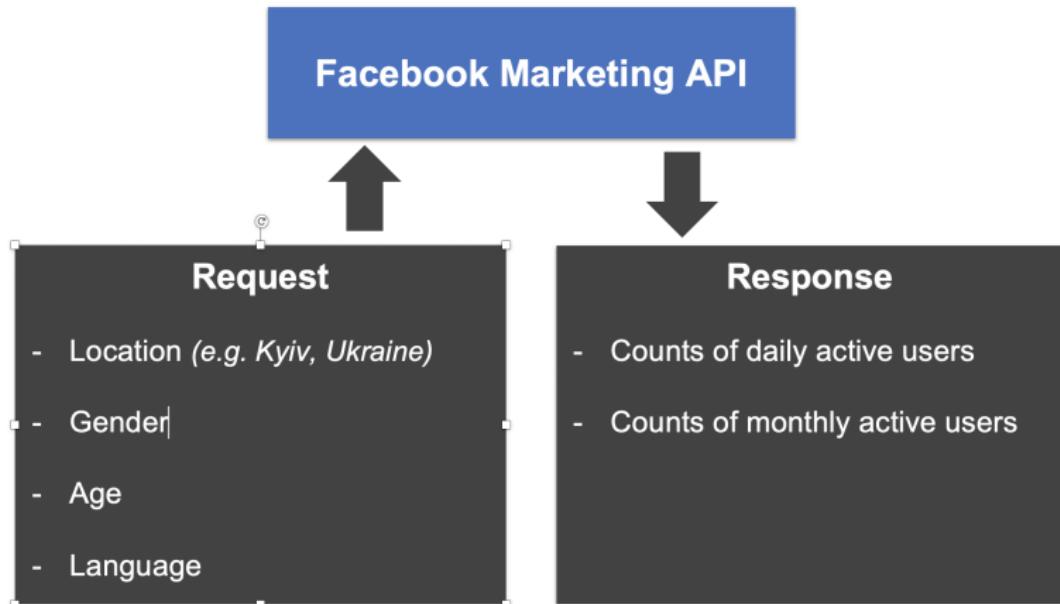
⁷ International Organization for Migration, United Nations

* douglas.leasure@sociology.ox.ac.uk

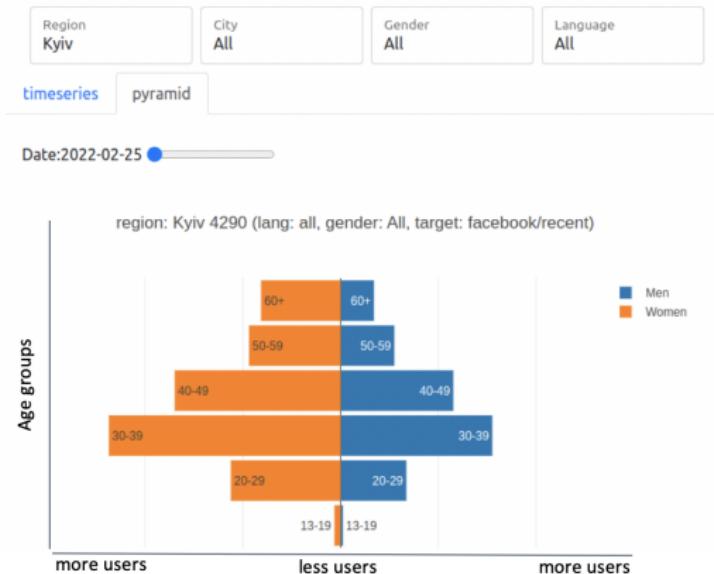


Note: This analysis is a rapid-response effort that has not yet undergone peer-review. All results are provisional and should be interpreted with caution. Version updates will be provided as potential issues are identified, methods are improved, or new data become available.

Figure: <https://doi.org/10.31235/osf.io/6j9wq>



Ukraine



Ukraine



Ukraine

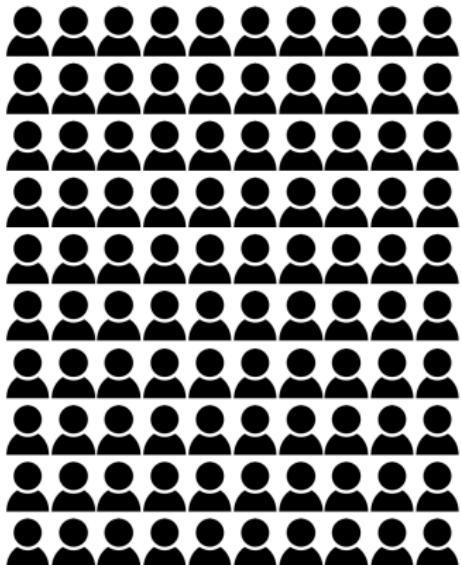


From FB estimates to Population Counts

Population Estimation

*30-34 year old women
in Kyiv before the conflict*

Baseline population = 100



From FB estimates to Population Counts

Population Estimation

*30-34 year old women
in Kyiv before the conflict*

Baseline population = 100

Baseline Facebook users = 25



From FB estimates to Population Counts

Population Estimation

*30-34 year old women
in Kyiv before the conflict*

Baseline population = 100

Baseline Facebook users = 25

Baseline Facebook penetration = 25%



From FB estimates to Population Counts

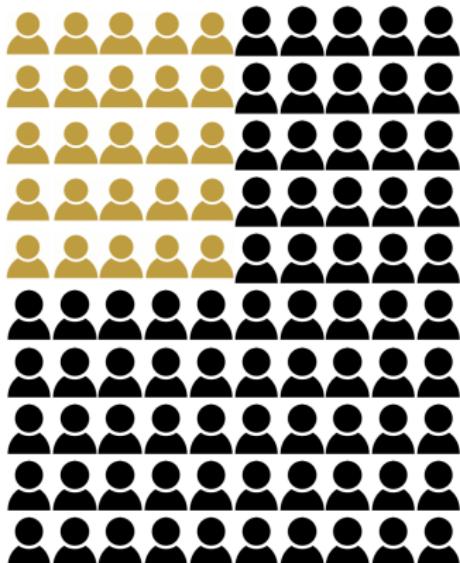
Population Estimation

*30-34 year old women
in Kyiv before the conflict*

Baseline population = 100

Baseline Facebook users = 25

Baseline Facebook penetration = 25%



From FB estimates to Population Counts

Population Estimation

*How many 30-34 year old women
are in Kyiv today?*



Facebook users = 15

POPULATION = USERS / PENETRATION

From FB estimates to Population Counts

Population Estimation

*How many 30-34 year old women
are in Kyiv today?*



Facebook users = 15

Baseline Facebook penetration = 25%

POPULATION = USERS / PENETRATION

From FB estimates to Population Counts

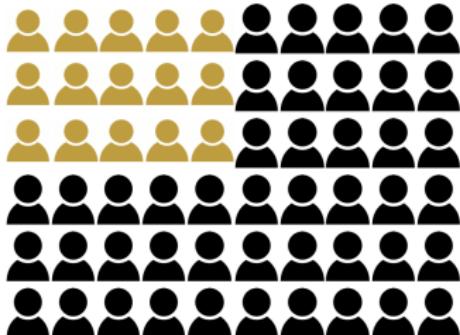
Population Estimation

How many 30-34 year old women are in Kyiv today?

Facebook users = 15

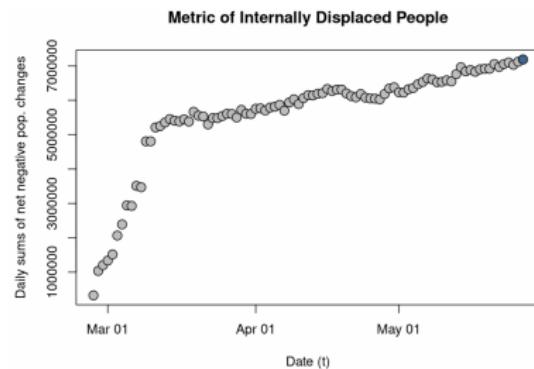
Baseline Facebook penetration = 25%

Current population = $15 / 0.25 = 60$



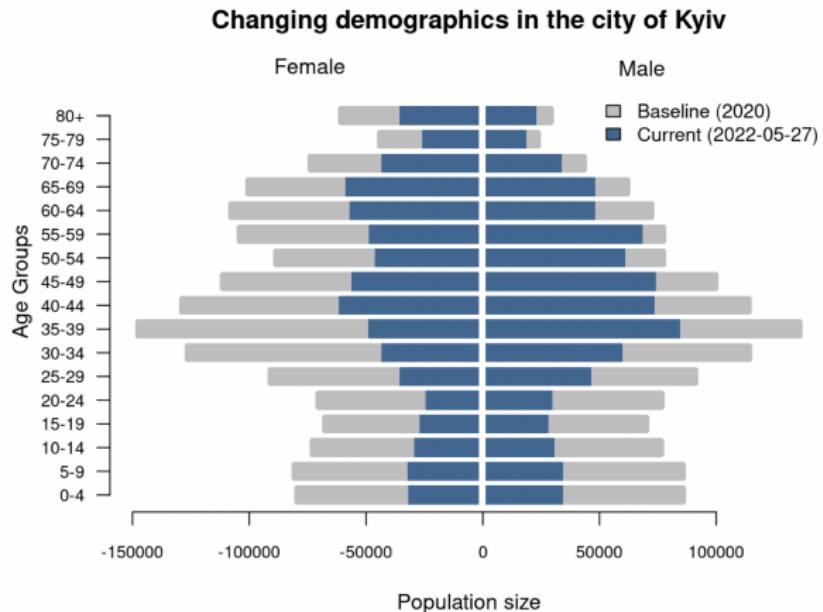
POPULATION = USERS / PENETRATION

Population Displacement Estimates



- ▶ As of May 27th: 7.2M fewer people estimated across all Oblasts where populations declined compared to baseline

Population Displacement Estimates



Summary: Social Media APIs

- ▶ Social media APIs serve as digital censuses that provide aggregated data on user populations.
- ▶ These data can provide valuable resolution for rapidly-evolving situations, or for monitoring social indicators.
- ▶ Yet, a significant shortcoming: no historic data available.
- ▶ Targeting categories change over time; their data generating process is opaque
- ▶ Algorithmic changes, version changes