Getting data from APIs

Getting and Cleaning Data


HTTP
httr
R package

HTTP

Hypertext Transfer Protocol

Allows messages to be sent on the Internet!

httr

R package

Enables YOU to work with these data in R!

2. Request goes to the website

Photo by rawpixel on Unsplash

1. Type URL

3. Website is displayed

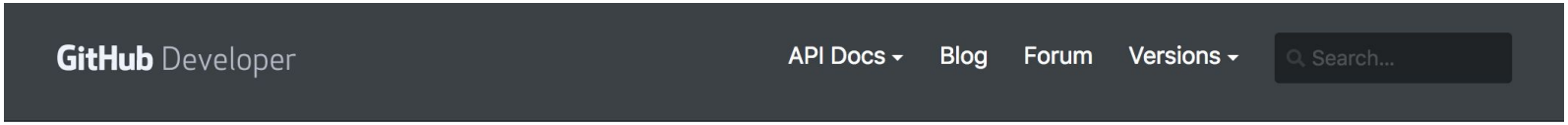
Photo by rawpixel on Unsplash

2. **API endpoint**
interprets request

```
GET(url = URL)
```

1. type **API request**

3. Get **API response**



REST API v3

Reference Guides Libraries

Overview

This describes the resources that make up the official GitHub REST API v3. If you have any problems or requests please contact GitHub support.

i. Current version
ii. Schema
iii. Authentication
iv. Parameters
v. Root endpoint
vi. GraphQL global node IDs
vii. Client errors
viii. HTTP redirects
ix. HTTP verbs
x. Hypermedia
xi. Pagination
xii. Rate limiting
xiii. User agent required
xiv. Conditional requests

https://developer.github.com/v3

```r
## load package
library(httr)
library(dplyr)

## Save GitHub username as variable
username <- 'janeeverydaydoe'

## Save base endpoint as variable
url_git <- 'https://api.github.com/'

## Construct API request
api_response <- GET(url = paste0(url_git,
'users/', username, '/repos'))
```

API endpoint

API request

```
## See variables in response
names(api_response)
```

```
> names(api_response)
 [1] "url"         "status_code" "headers"     "all_headers" "cookies"     "content"
 [7] "date"        "times"       "request"     "handle"
```

What API request from `httr` returns

```r
## Check Status Code of request
api_response$status_code

## Extract content from API response
repo_content <- content(api_response)
```

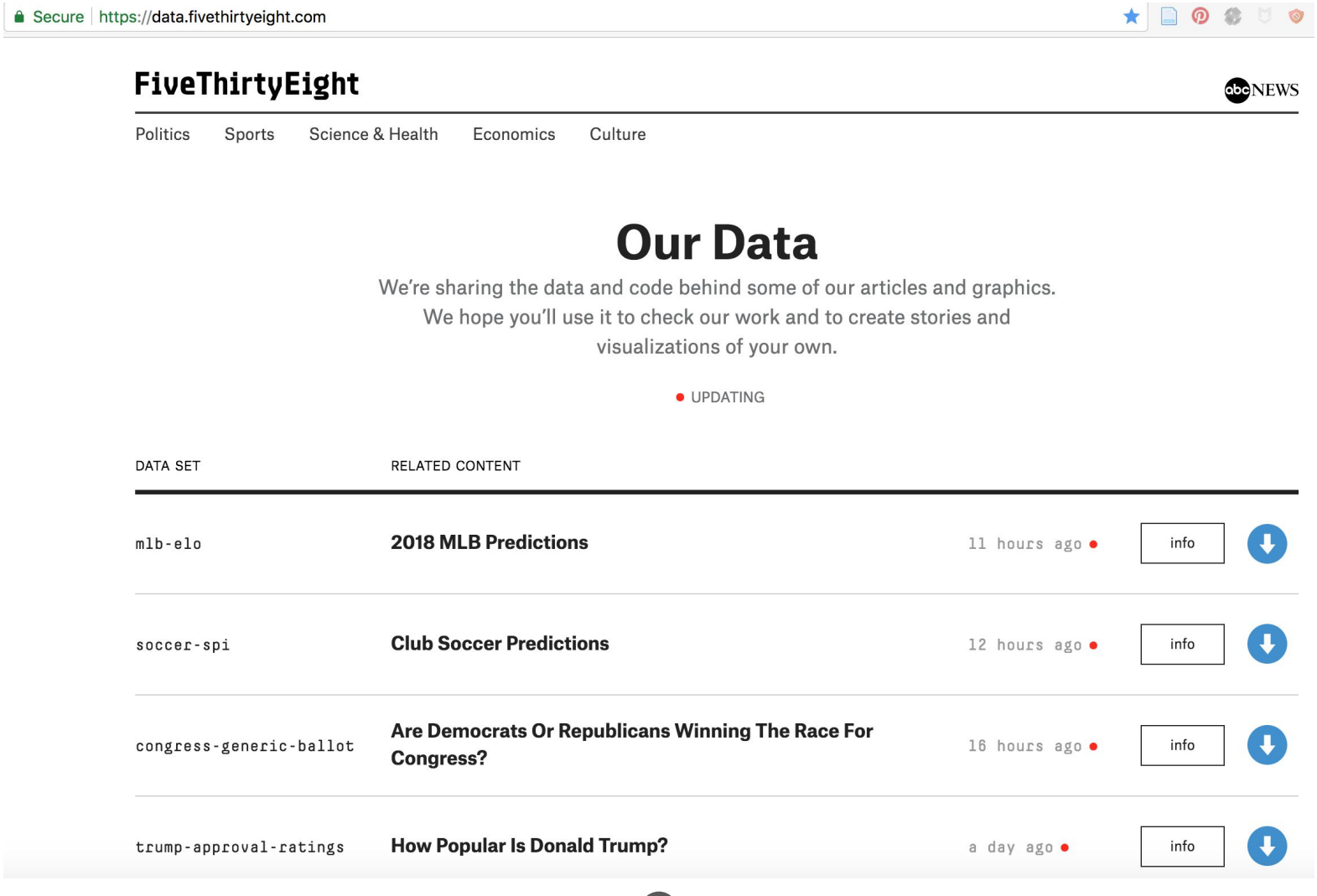
```
> api_response$status_code
[1] 200

> repo_content <- content(api_response)
```

'200' means request was successful!

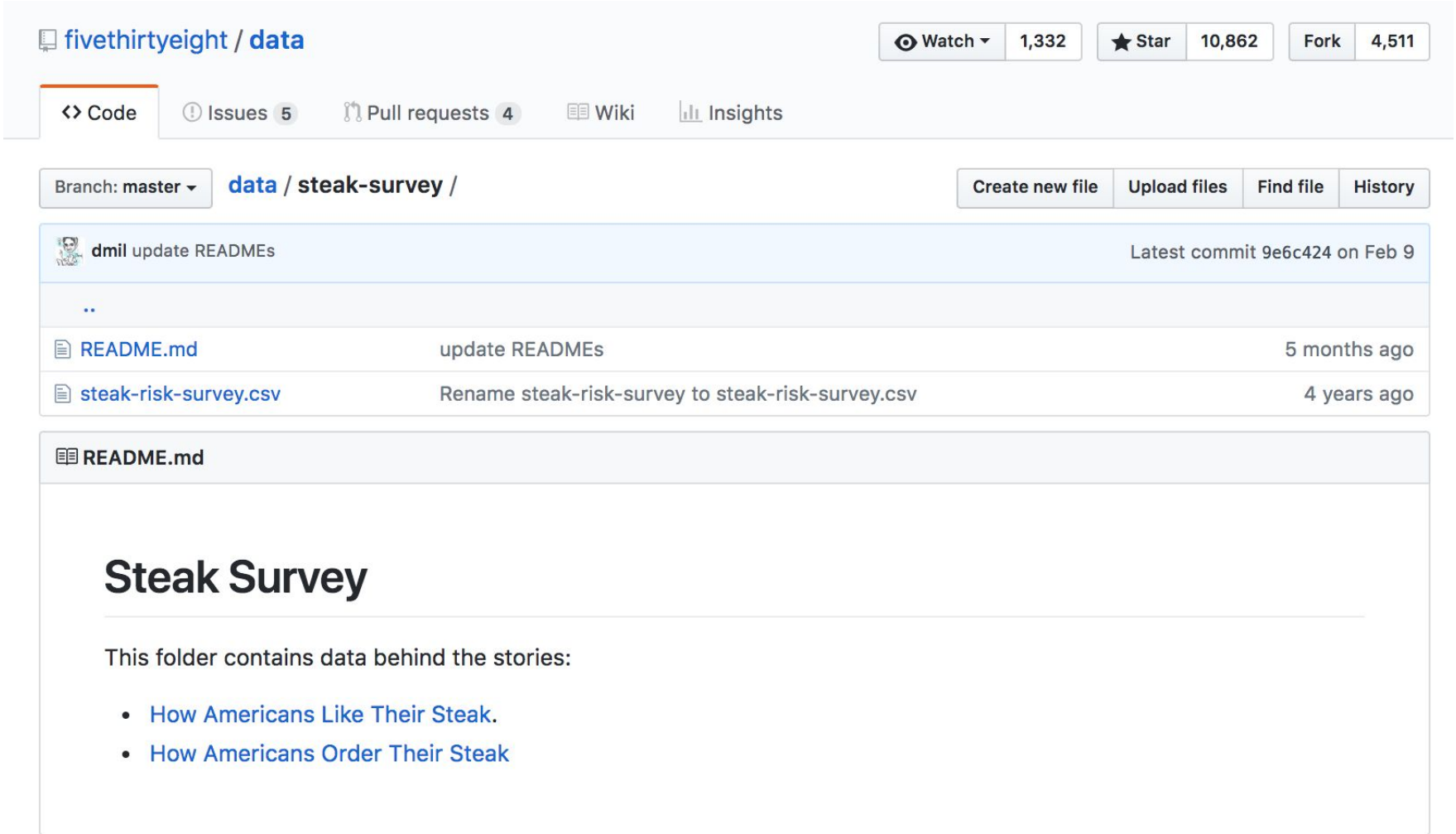
`content()` extracts contents from API request

```
## function to get name and URL for each repo
lapply(repo_content, function(x) {
  df <- data_frame(repo = x$name,
                   address = x$html_url)}) %>%
  bind_rows()
```

```
> lapply(repo_content, function(x) {
+     df <- data_frame(repo = x$name,
+                      address = x$html_url)}) %>%
+     bind_rows()
# A tibble: 6 x 2
  repo                              address
  <chr>                             <chr>
1 first_project                     https://github.com/JaneEverydayDoe/first_project
2 hello-world                       https://github.com/JaneEverydayDoe/hello-world
3 janeeverydaydoe.github.com        https://github.com/JaneEverydayDoe/janeeverydaydoe.g…
4 my_first_project                  https://github.com/JaneEverydayDoe/my_first_project
5 newproject                        https://github.com/JaneEverydayDoe/newproject
6 Temporary_add_to_version_control https://github.com/JaneEverydayDoe/Temporary_add_to_…
```



https://data.fivethirtyeight.com



https://github.com/fivethirtyeight/data/tree/master/steak-survey

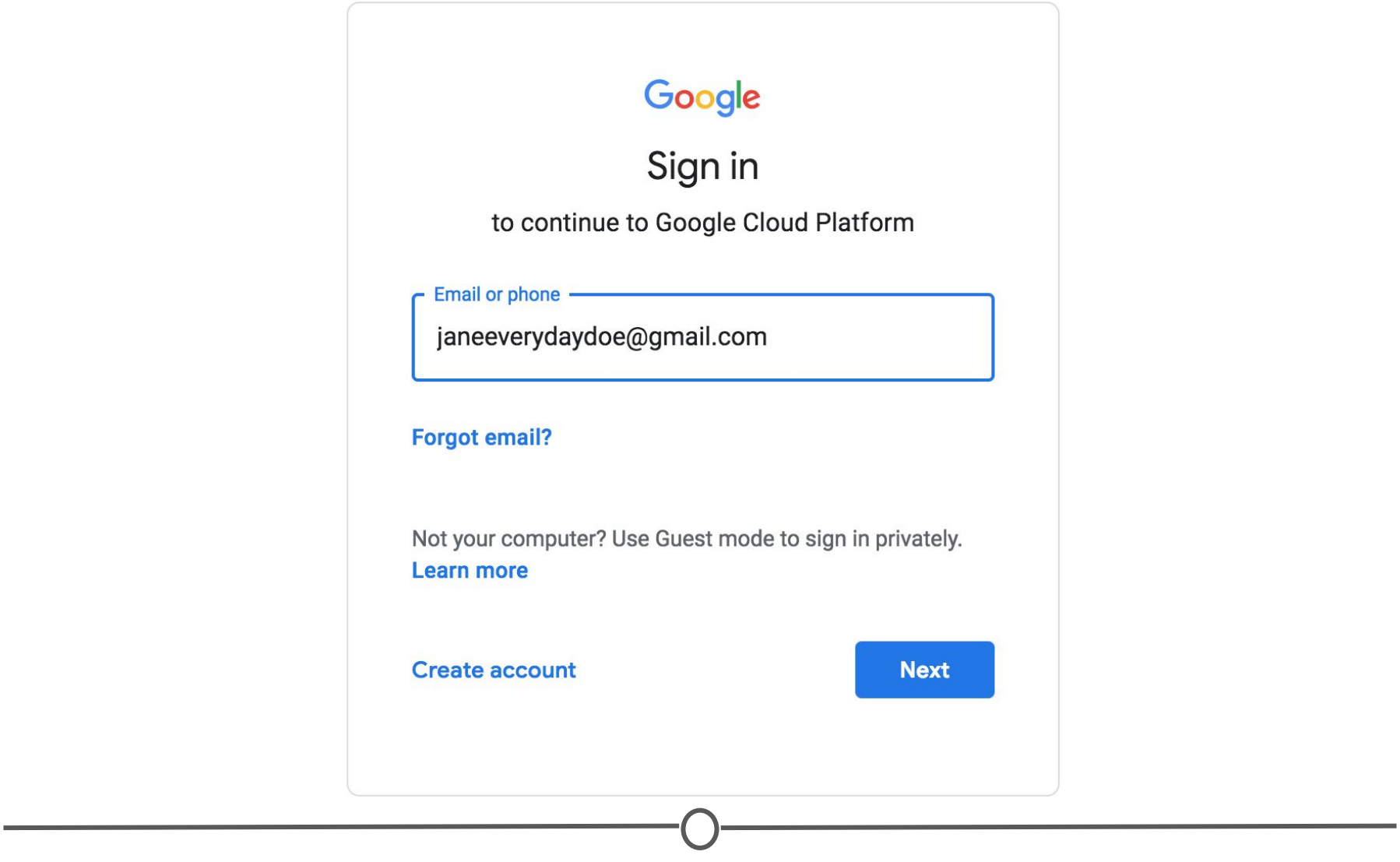
```
> ## Make API request
> api_response <- GET(url = "https://raw.githubusercontent.com/fivethirtyeight/data/master/steak-survey/steak-risk-survey.csv")
>
> ## Extract content from API response
> df_steak <- content(api_response, type="text/csv")
No encoding supplied: defaulting to UTF-8.
Parsed with column specification:
cols(
  RespondentID = col_double(),
  `Consider the following hypothetical situations: <br>In Lottery A, you have a 50% chance of success, with a payout of $100. <br>In Lottery B, you have a 90% chance of success, with a pay out of $20. <br><br>Assuming you have $10 to bet, would you play Lottery A or Lottery B?` = col_character(),
  `Do you ever smoke cigarettes?` = col_character(),
  `Do you ever drink alcohol?` = col_character(),
  `Do you ever gamble?` = col_character(),
  `Have you ever been skydiving?` = col_character(),
  `Do you ever drive above the speed limit?` = col_character(),
  `Have you ever cheated on your significant other?` = col_character(),
  `Do you eat steak?` = col_character(),
  `How do you like your steak prepared?` = col_character(),
  Gender = col_character(),
  Age = col_character(),
  `Household Income` = col_character(),
  Education = col_character(),
  `Location (Census Region)` = col_character()
)
```

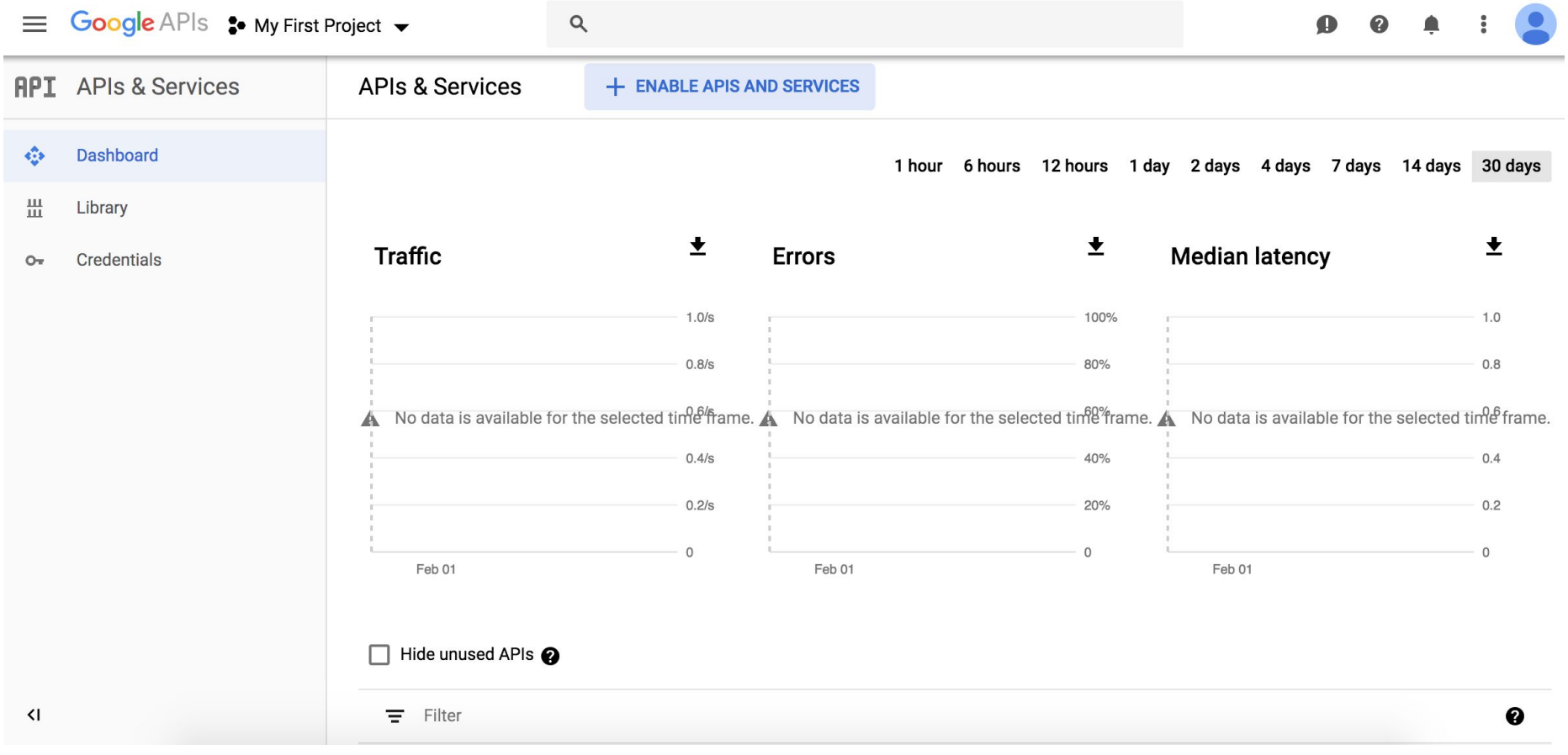
`GET()` makes API request

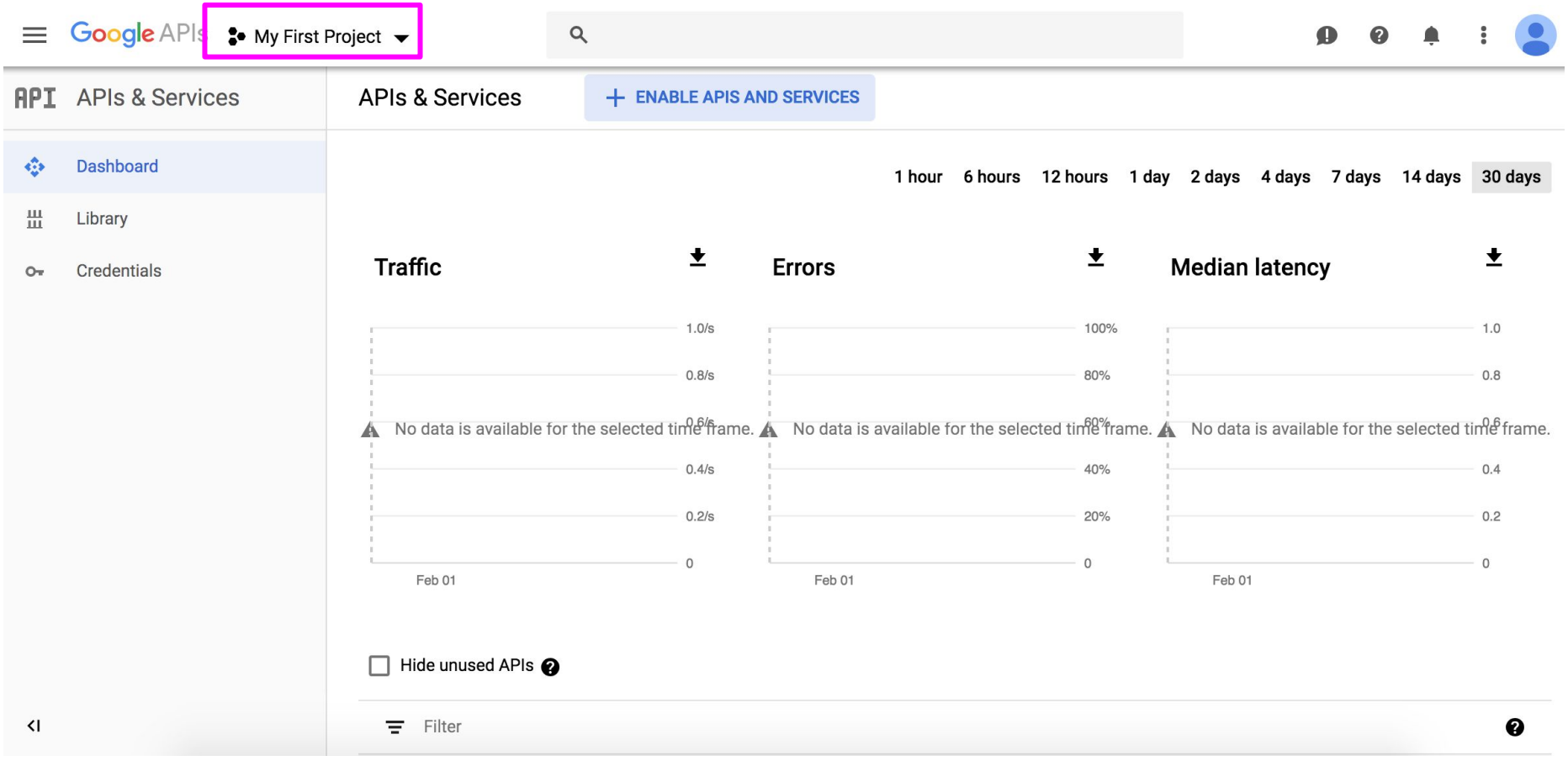
`content()` extracts information

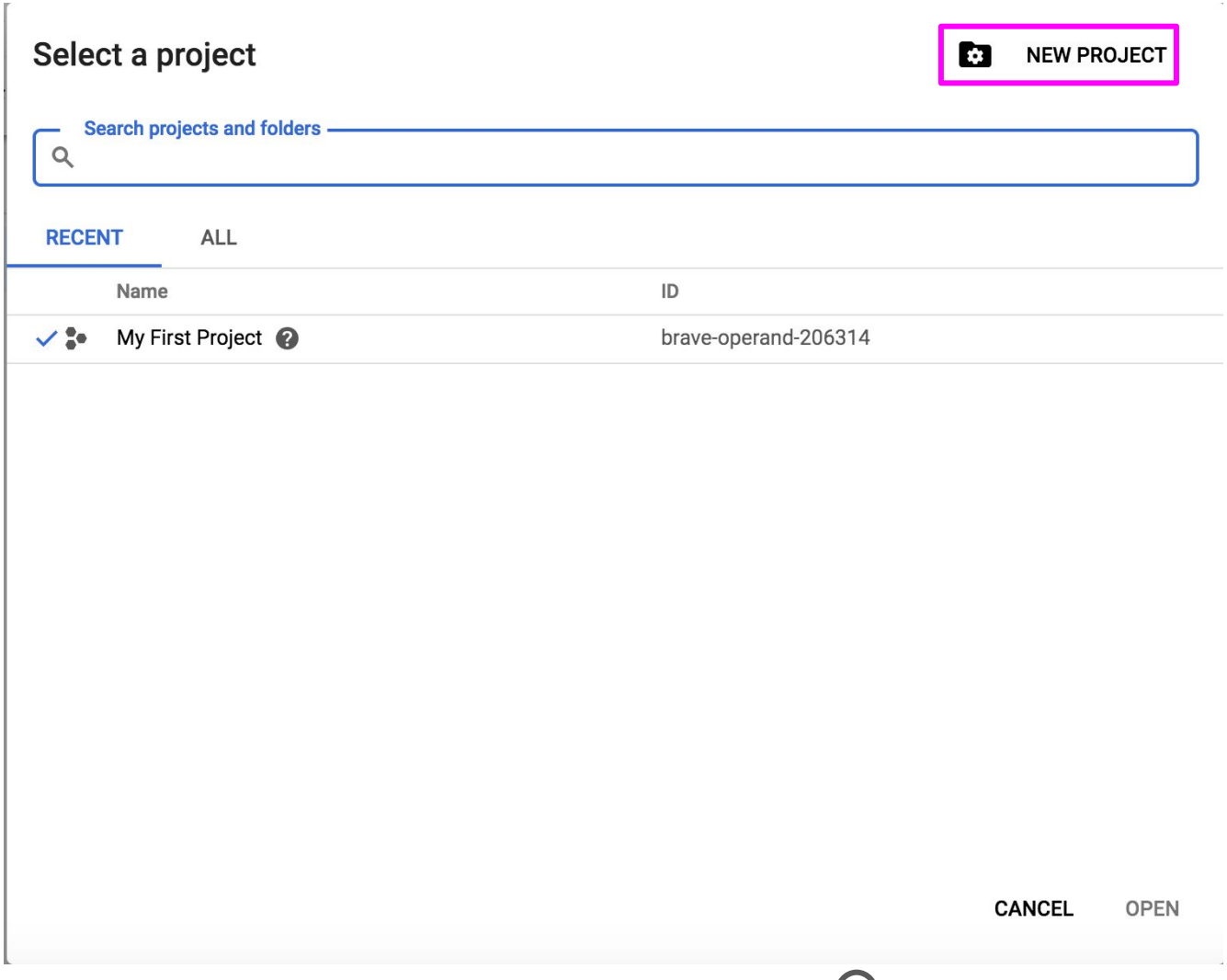

*****
designed by freepik.com

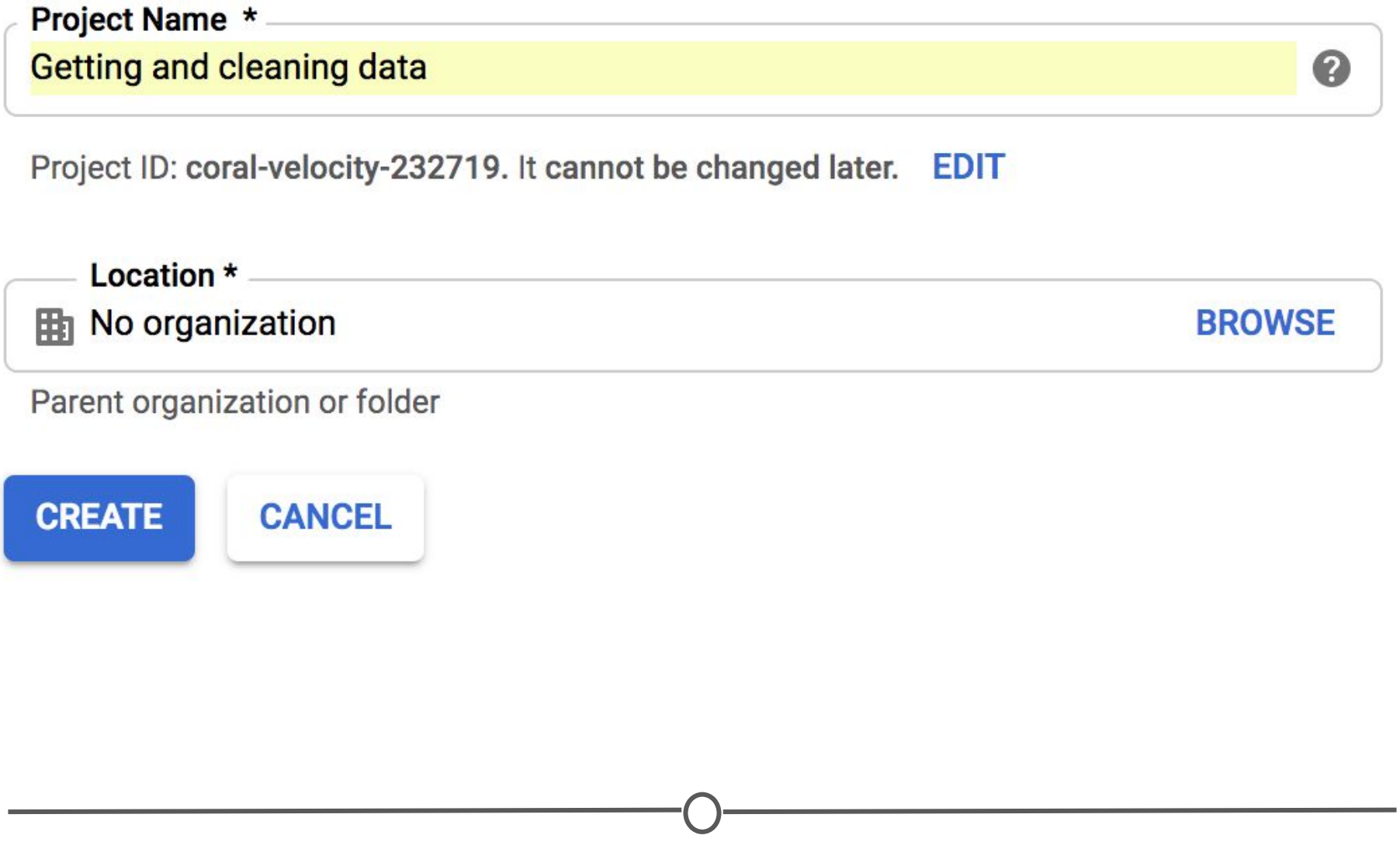
```
myapp = oauth_app("twitter",
                  key = "yourConsumerKeyHere",
                  secret = "yourConsumerSecretHere")
sig = sign_oauth1.0(myapp,
                    token = "yourTokenHere",
                    token_secret = "yourTokenSecretHere")
homeTL =
GET("https://api.twitter.com/1.1/statuses/home_timeline.js
on", sig)
```

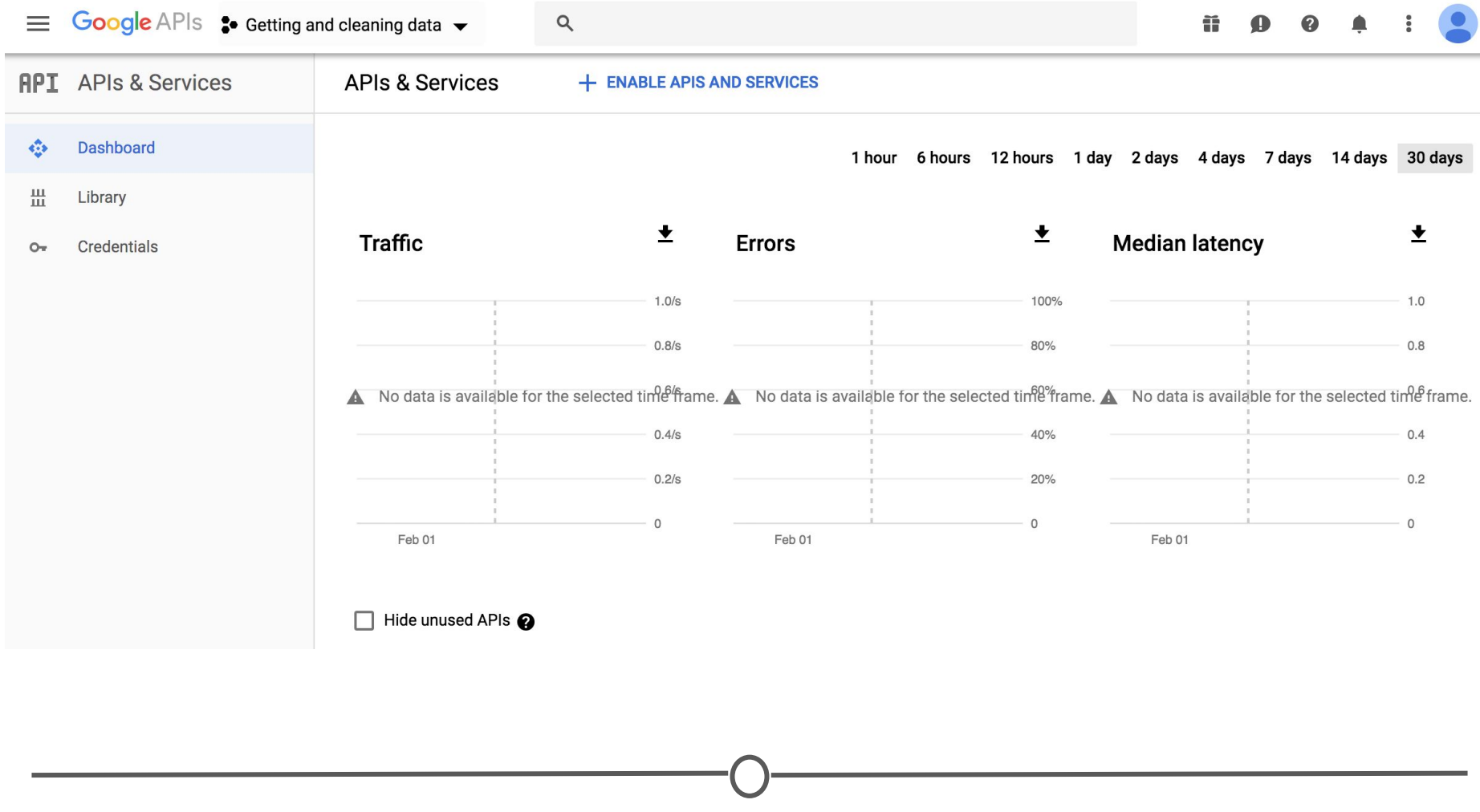

Google
Sign in
to continue to Google Cloud Platform
Email or phone
janeeverydaydoe@gmail.com
Forgot email?
Not your computer? Use Guest mode to sign in privately.
Learn more
Create account
Next

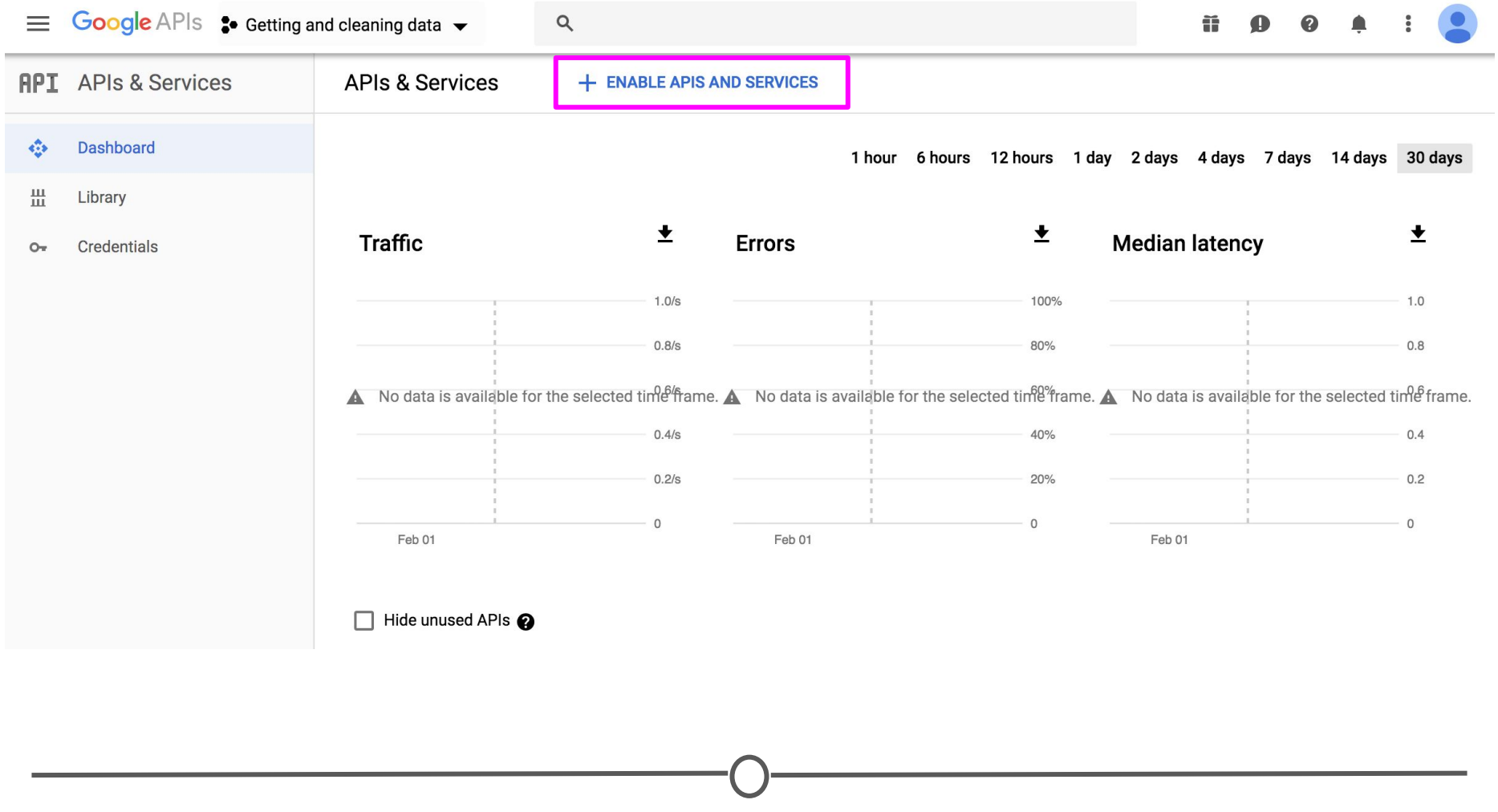

Google APIs
My First Project
APIs & Services
Dashboard
Library
Credentials
APIs & Services
+ ENABLE APIS AND SERVICES
1 hour
6 hours
12 hours
1 day
2 days
4 days
7 days
14 days
30 days
Traffic
1.0/s
0.8/s
0.6/s
0.4/s
0.2/s
0
Feb 01
No data is available for the selected time frame.
Errors
100%
80%
60%
40%
20%
0
Feb 01
No data is available for the selected time frame.
Median latency
1.0
0.8
0.6
0.4
0.2
0
Feb 01
No data is available for the selected time frame.
Hide unused APIs
Filter

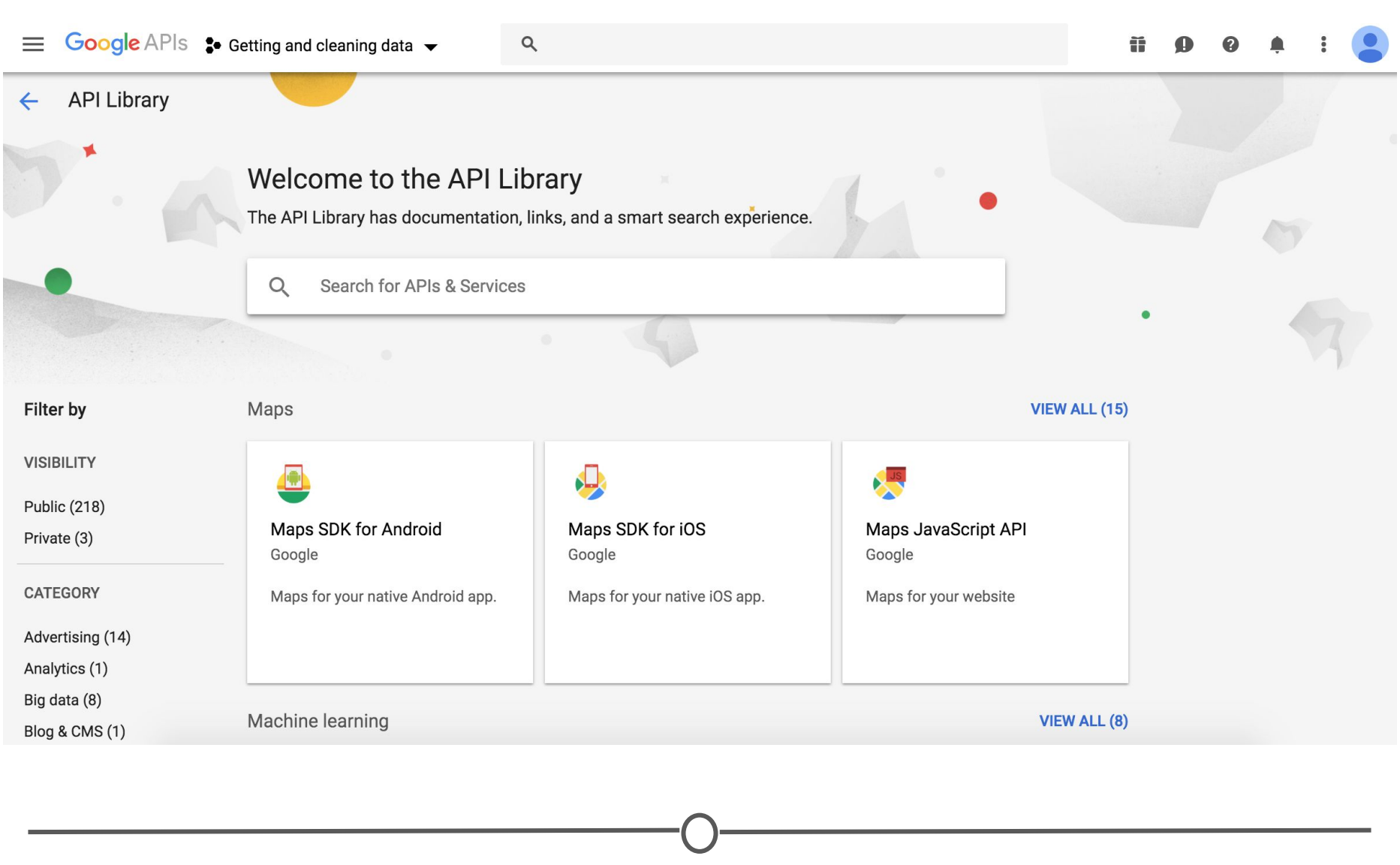

Google APIs
My First Project
APIs & Services
Dashboard
Library
Credentials
APIs & Services
ENABLE APIS AND SERVICES
1 hour
6 hours
12 hours
1 day
2 days
4 days
7 days
14 days
30 days
Traffic
Errors
Median latency
1.0/s
0.8/s
0.6/s
0.4/s
0.2/s
0
100%
80%
60%
40%
20%
0
1.0
0.8
0.6
0.4
0.2
0
Feb 01
Feb 01
Feb 01
No data is available for the selected time frame.
No data is available for the selected time frame.
No data is available for the selected time frame.
Hide unused APIs
Filter

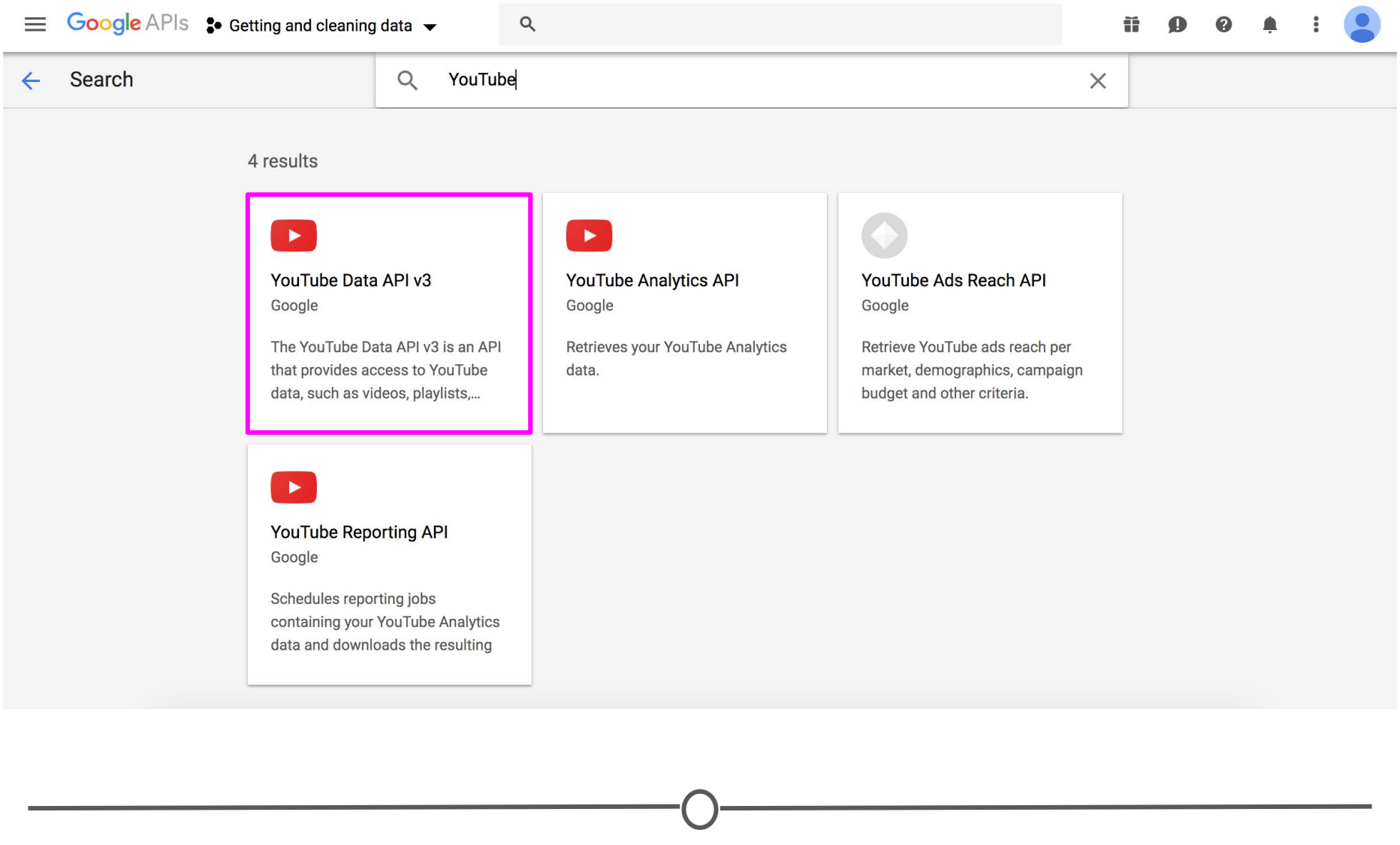

Select a project
NEW PROJECT
Search projects and folders
RECENT
ALL
Name
ID
My First Project
brave-operand-206314
CANCEL
OPEN


Project Name *
Getting and cleaning data
Project ID: coral-velocity-232719. It cannot be changed later. EDIT
Location *
No organization
BROWSE
Parent organization or folder
CREATE
CANCEL


Google APIs
Getting and cleaning data
APIs & Services
Dashboard
Library
Credentials
APIs & Services
ENABLE APIS AND SERVICES
1 hour
6 hours
12 hours
1 day
2 days
4 days
7 days
14 days
30 days
Traffic
Errors
Median latency
1.0/s
0.8/s
0.6/s
0.4/s
0.2/s
0
100%
80%
60%
40%
20%
0
1.0
0.8
0.6
0.4
0.2
0
Feb 01
No data is available for the selected time frame.
Hide unused APIs


Google APIs
Getting and cleaning data
APIs & Services
Dashboard
Library
Credentials
APIs & Services
+ ENABLE APIS AND SERVICES
1 hour
6 hours
12 hours
1 day
2 days
4 days
7 days
14 days
30 days
Traffic
Errors
Median latency
1.0/s
0.8/s
0.6/s
0.4/s
0.2/s
0
100%
80%
60%
40%
20%
0
1.0
0.8
0.6
0.4
0.2
0
Feb 01
No data is available for the selected time frame.
Hide unused APIs


Google APIs
Getting and cleaning data
API Library
Welcome to the API Library
The API Library has documentation, links, and a smart search experience.
Search for APIs & Services
Filter by
VISIBILITY
Public (218)
Private (3)
CATEGORY
Advertising (14)
Analytics (1)
Big data (8)
Blog & CMS (1)
Maps
VIEW ALL (15)
Maps SDK for Android
Google
Maps for your native Android app.
Maps SDK for iOS
Google
Maps for your native iOS app.
Maps JavaScript API
Google
Maps for your website
Machine learning
VIEW ALL (8)


Google APIs
Getting and cleaning data
Search
YouTube
4 results
YouTube Data API v3
Google
The YouTube Data API v3 is an API that provides access to YouTube data, such as videos, playlists,...
YouTube Analytics API
Google
Retrieves your YouTube Analytics data.
YouTube Ads Reach API
Google
Retrieve YouTube ads reach per market, demographics, campaign budget and other criteria.
YouTube Reporting API
Google
Schedules reporting jobs containing your YouTube Analytics data and downloads the resulting

YouTube Data API v3

Google

The YouTube Data API v3 is an API that provides access to YouTube data, such as videos, playlists,...

ENABLE TRY THIS API

Type
APIs & services

Last updated
1/9/19, 3:25 PM

Category
YouTube

Service name
youtube.googleapis.com

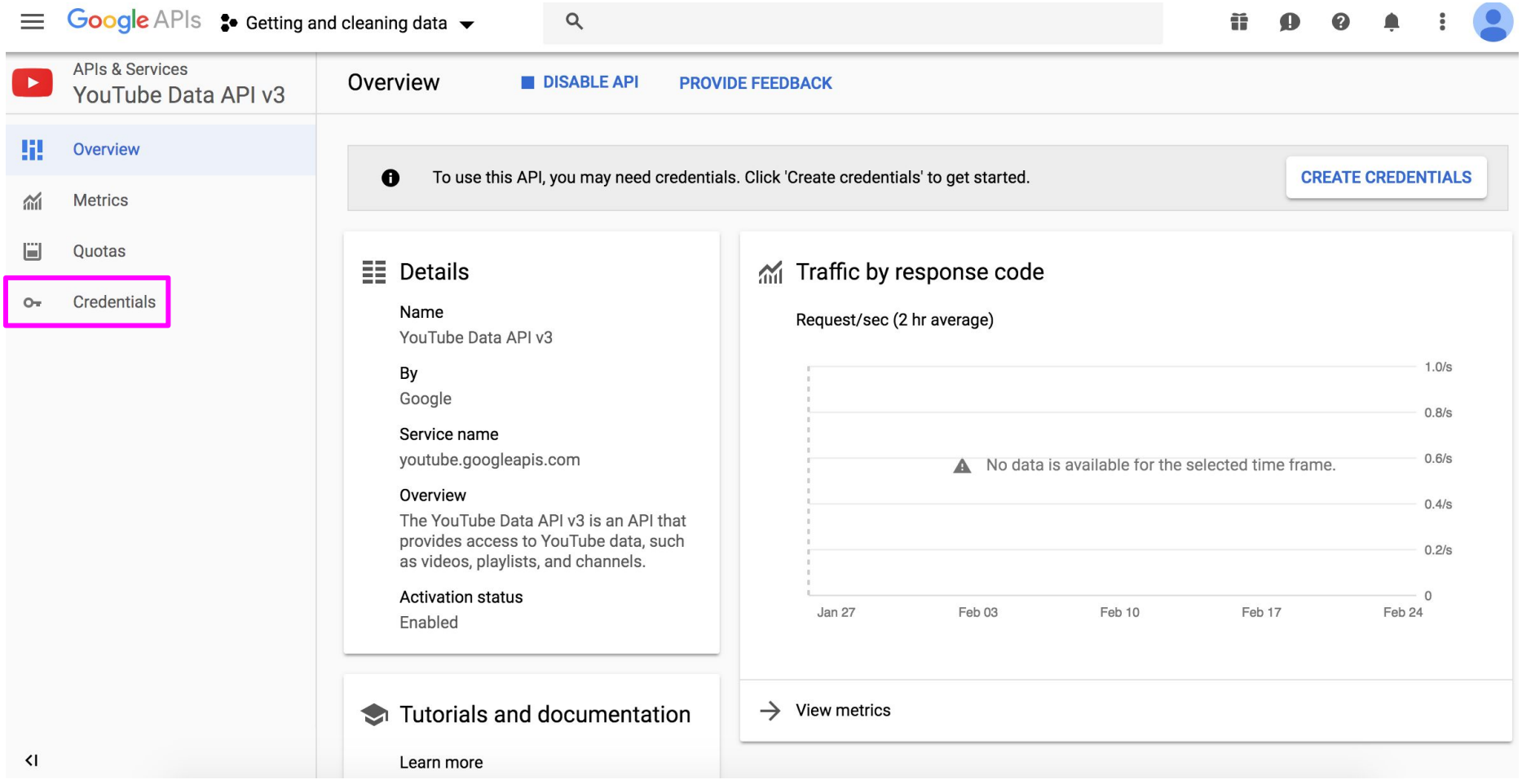
Overview

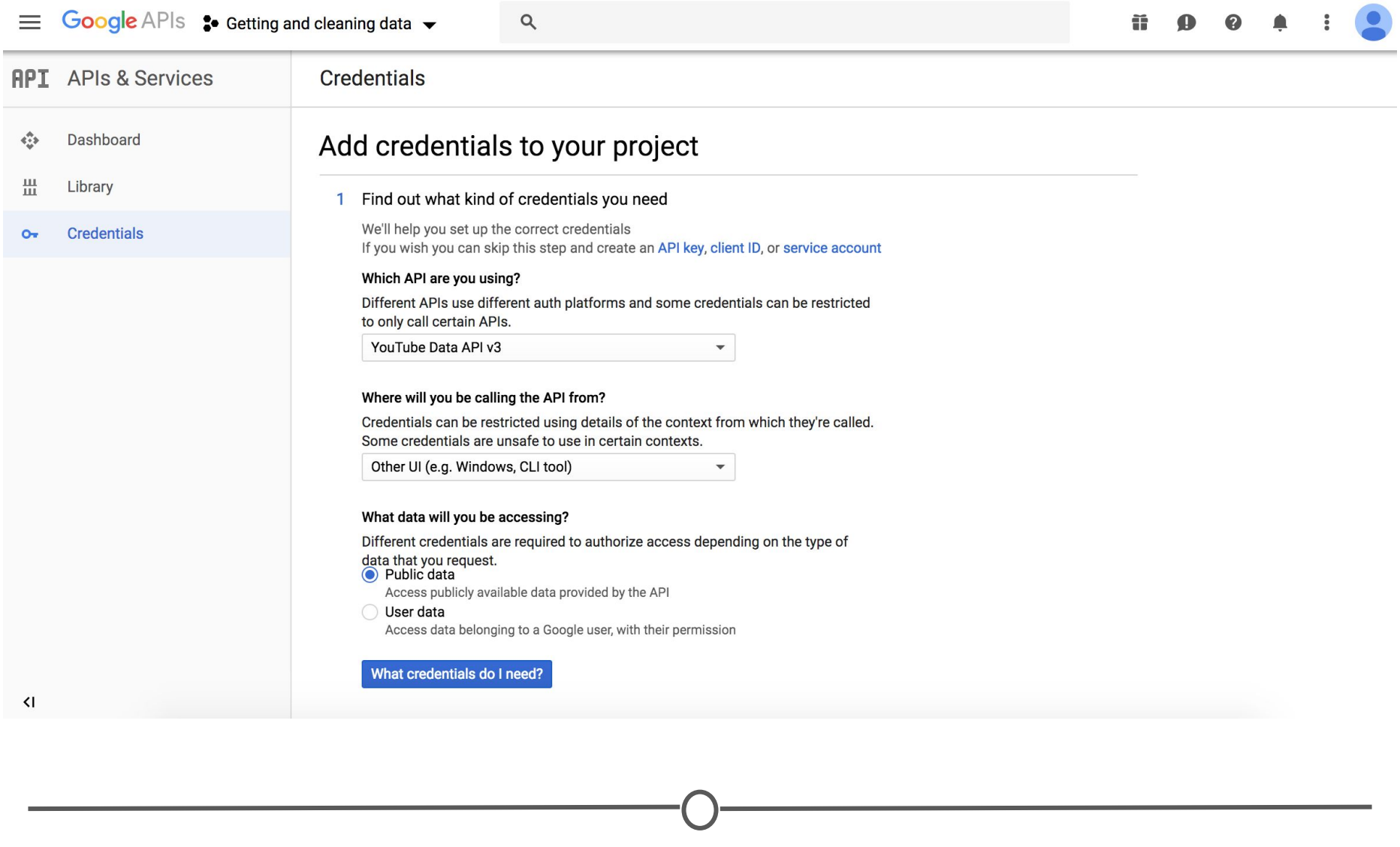
The YouTube Data API v3 is an API that provides access to YouTube data, such as videos, playlists, and channels.

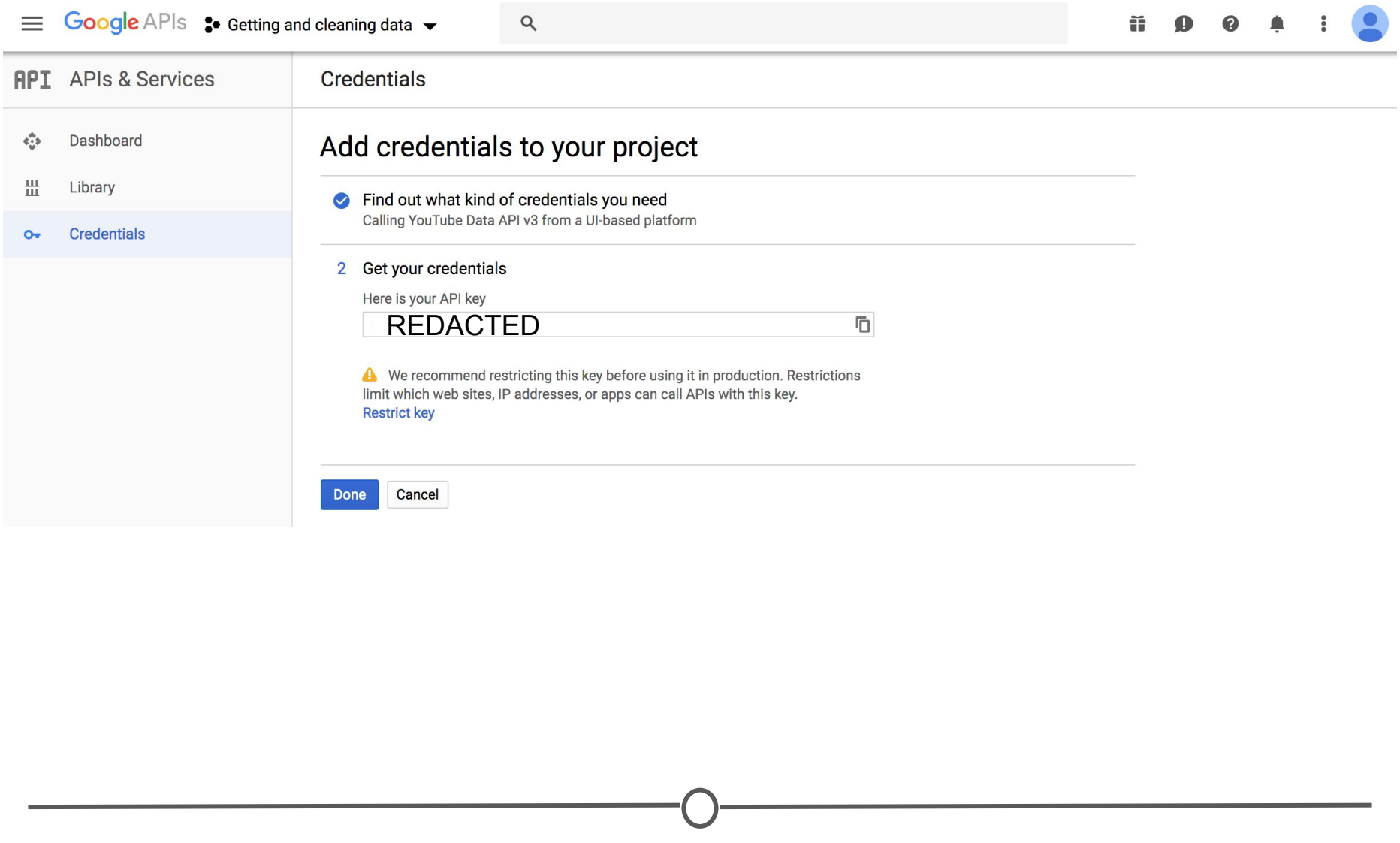
About Google

Google's mission is to organize the world's information and make it universally accessible and useful. Through products and platforms like Search, Maps, Gmail, Android, Google Play, Chrome and YouTube, Google plays a meaningful role in the daily lives of billions of people.

Tutorials and documentation


Google APIs
Getting and cleaning data
APIs & Services
YouTube Data API v3
Overview
Metrics
Quotas
Credentials
Overview
DISABLE API
PROVIDE FEEDBACK
To use this API, you may need credentials. Click 'Create credentials' to get started.
CREATE CREDENTIALS
Details
Name
YouTube Data API v3
By
Google
Service name
youtube.googleapis.com
Overview
The YouTube Data API v3 is an API that provides access to YouTube data, such as videos, playlists, and channels.
Activation status
Enabled
Traffic by response code
Request/sec (2 hr average)
1.0/s
0.8/s
0.6/s
0.4/s
0.2/s
0
No data is available for the selected time frame.
Jan 27
Feb 03
Feb 10
Feb 17
Feb 24
View metrics
Tutorials and documentation
Learn more


Google APIs
Getting and cleaning data
APIs & Services
Credentials
Dashboard
Library
Credentials
Add credentials to your project
1 Find out what kind of credentials you need
We'll help you set up the correct credentials
If you wish you can skip this step and create an API key, client ID, or service account
Which API are you using?
Different APIs use different auth platforms and some credentials can be restricted to only call certain APIs.
YouTube Data API v3
Where will you be calling the API from?
Credentials can be restricted using details of the context from which they're called. Some credentials are unsafe to use in certain contexts.
Other UI (e.g. Windows, CLI tool)
What data will you be accessing?
Different credentials are required to authorize access depending on the type of data that you request.
Public data
Access publicly available data provided by the API
User data
Access data belonging to a Google user, with their permission
What credentials do I need?


Google APIs
Getting and cleaning data
APIs & Services
Dashboard
Library
Credentials
Credentials
Add credentials to your project
Find out what kind of credentials you need
Calling YouTube Data API v3 from a UI-based platform
2 Get your credentials
Here is your API key
REDACTED
We recommend restricting this key before using it in production. Restrictions limit which web sites, IP addresses, or apps can call APIs with this key.
Restrict key
Done
Cancel

```r
# Save your API as a variable.
# NEVER share this with anybody -- including accidentally sharing on GitHub!
API_key <- "REDACTED"

# Base URL for the YouTube API
Base_URL <- "https://www.googleapis.com/youtube/v3/"

# Our query term, formatted as per YouTube API requirements
Query <- "search?q=kittens&maxResults=10&order=date&part=snippet&key="

# Combine each of these to perform the query using GET()
Response <- GET(url = paste0(Base_URL, Query, API_key))

# Check that our request worked!
Response$status_code

# See what we got back!
test <- content(Response)

# Just look at the video titles
for(i in 1:10) {
  print(content(Response)$items[[i]]$snippet$title)
}

```

```
> for(i in 1:10) {
+   print(content(Response)$items[[i]]$snippet$title)
+ }
[1] "So many cute kittens videos compilation 2018"
[1] "Learn How Baby Kittens Grow: 0-8 Weeks!"
[1] "Cute is Not Enough - Cute Cats and Kittens Doing Funny Things 2018 #11"
[1] "猫部屋ライブ映像　Cats & Kittens room 【Miaou みゃう】"
[1] "LIVE: Stanley the Tiny Kitten - TinyKittens.com"
[1] "Cute Cats and Little Kittens Meowing and Talking Compilation"
[1] "Three Little Kittens | +More Nursery Rhymes &amp;amp; Kids Songs - CoCoMelon"
[1] "♥Cute Cats and Kittens Doing Funny Things 2018♥ #3 - Funny Cat compilation"
[1] "Tigers reaction to the kittens"
[1] "10 Tips for Tricky Bottle Baby Kittens"
```

```
25  # Using JSONlite to query the YouTube API
26
27  # Perform query
28  Response2 <- fromJSON(paste0(Base_URL, Query, API_key))
29
30  # Check all results
31  Response2
32
33  # Just look at the video titles
34  Response2$items$snippet$title
35
```

```r
# See how the results of querying the API have changed in the last few minutes
# Use the same query as before
Response_updated <- GET(url = paste0(Base_URL, Query, API_key))

# Lets see how the video titles have changed!

## Before
for(i in 1:10) {
  print(content(Response)$items[[i]]$snippet$title)
}

## Updated
for(i in 1:10) {
  print(content(Response_updated)$items[[i]]$snippet$title)
}

# Best practice: Record the date and time that you accessed the data
Sys.time()

```

BEFORE

```
[1] "So many cute kittens videos compilation 2018"
[1] "Learn How Baby Kittens Grow: 0-8 Weeks!"
[1] "Cute is Not Enough - Cute Cats and Kittens Doing Funny Things 2018 #11"
[1] "猫部屋ライブ映像　Cats & Kittens room 【Miaou みゃう】"
[1] "LIVE: Stanley the Tiny Kitten - TinyKittens.com"
[1] "Cute Cats and Little Kittens Meowing and Talking Compilation"
[1] "Three Little Kittens | +More Nursery Rhymes &amp;amp; Kids Songs - CoCoMelon"
[1] "♥Cute Cats and Kittens Doing Funny Things 2018♥ #3 - Funny Cat compilation"
[1] "Tigers reaction to the kittens"
[1] "10 Tips for Tricky Bottle Baby Kittens"
```

AFTER

```
[1] "Waking With Kittens.. if you haven&amp;#39;t gotten to."
[1] "My kittens"
[1] "Saar &amp;amp; kittens | zaterdag 30 maart"
[1] "Adorable Kittens Live Stream"
[1] "Cutest kittens ever!!"
[1] "Cat Cam - Countdown to Kittens"
[1] "[Vlog] Woke Up Max And Kitty Murka Did Not Lead Kittens  # 93"
[1] "Funny kittens playing."
[1] "Playing with the kittens"
[1] "Cute kittens"
```

Getting data from APIs

Getting and Cleaning Data