

Professor: Dr. Karen Han

Student: Jiaming Li

Ball State University

## YouTube Data Scraping Manual

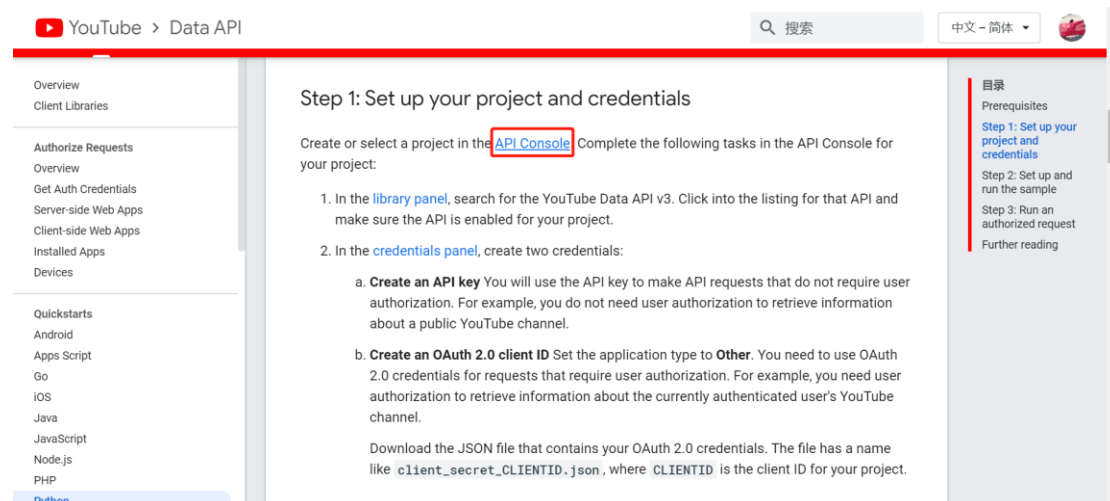
YouTube comment scrape based on Google API

### 1. Preparatory work: (Sorry that my Google is all in Chinese)

<https://developers.google.com/youtube/v3/quickstart/python>

Click on the url just given to register for a Google cloud platform

scrolled down this page when you see this:

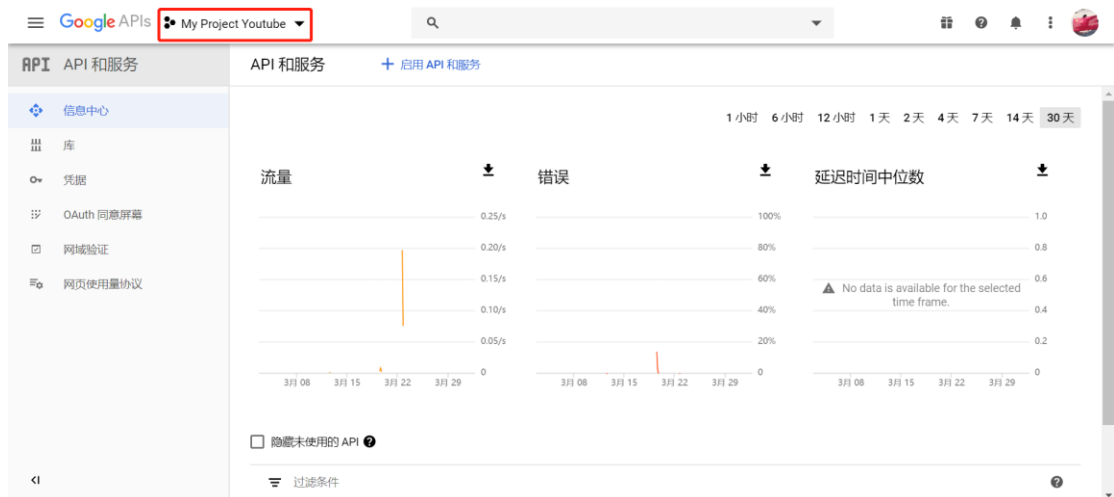


Follow the steps on this page.

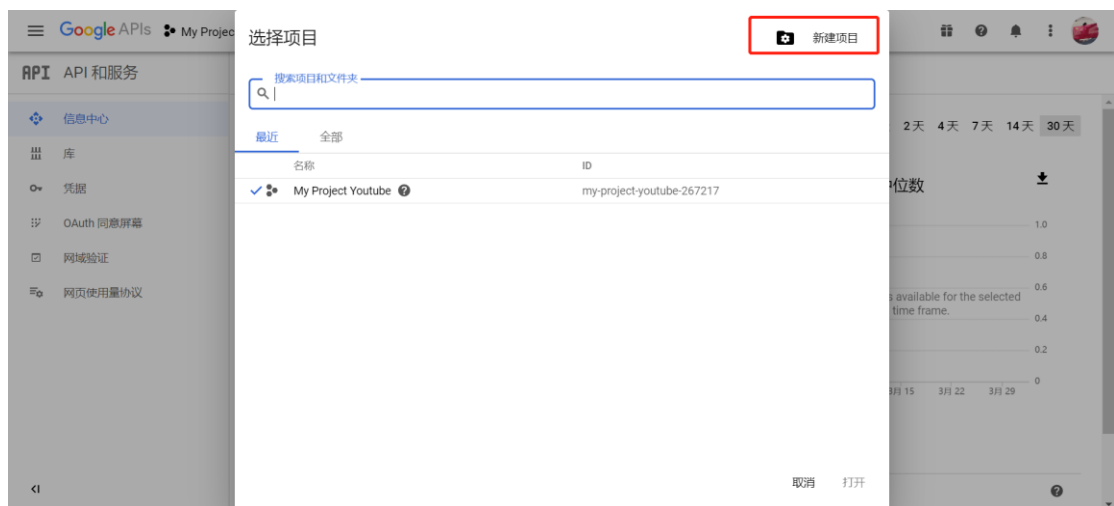
(1) Open API Console link that I circled below.

(2) Login to Google account and access to Google API, as shown in the following

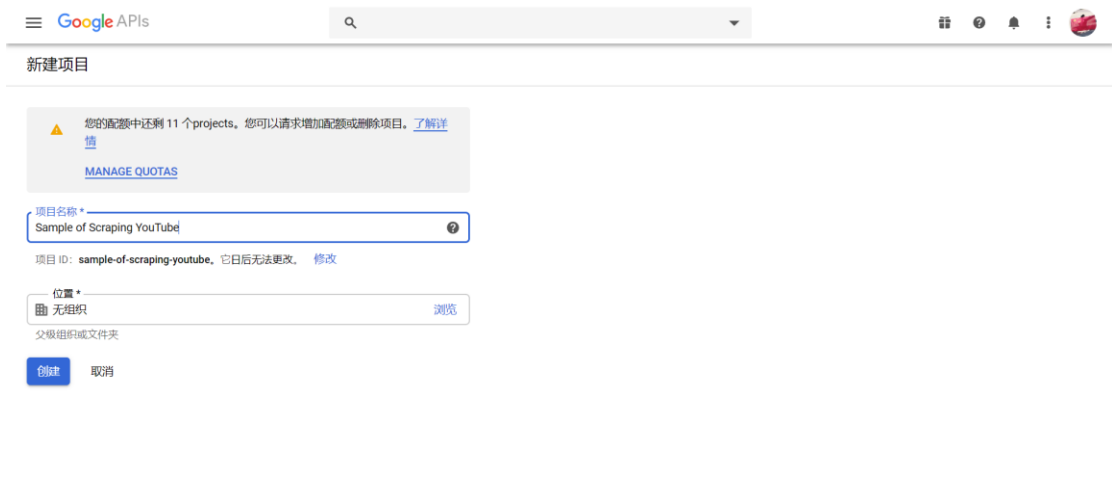
figure :( because I have an existing project, the top bar shows one of my projects)



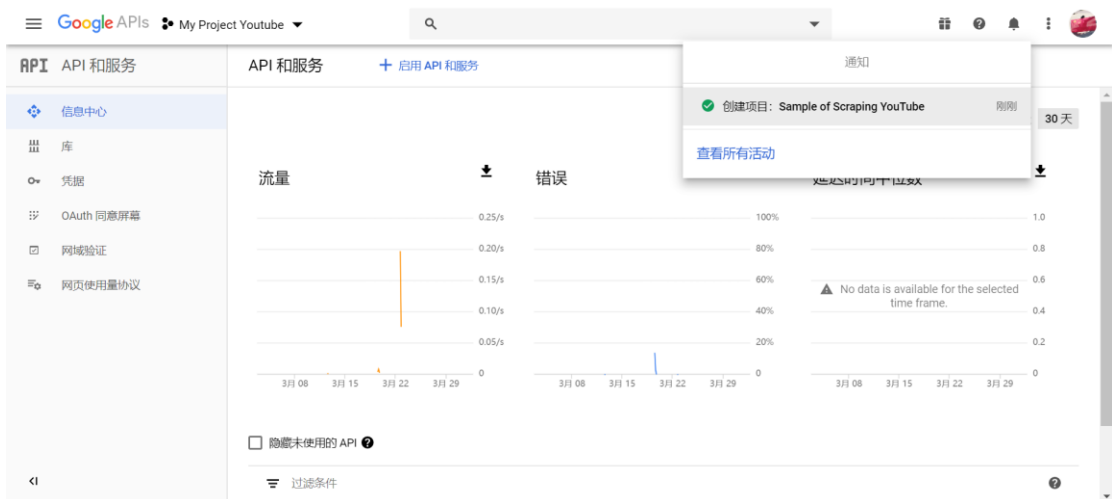
(3) Click “My Project YouTube” (drop-down arrow here) and click " create a new project" in the pop-up box:



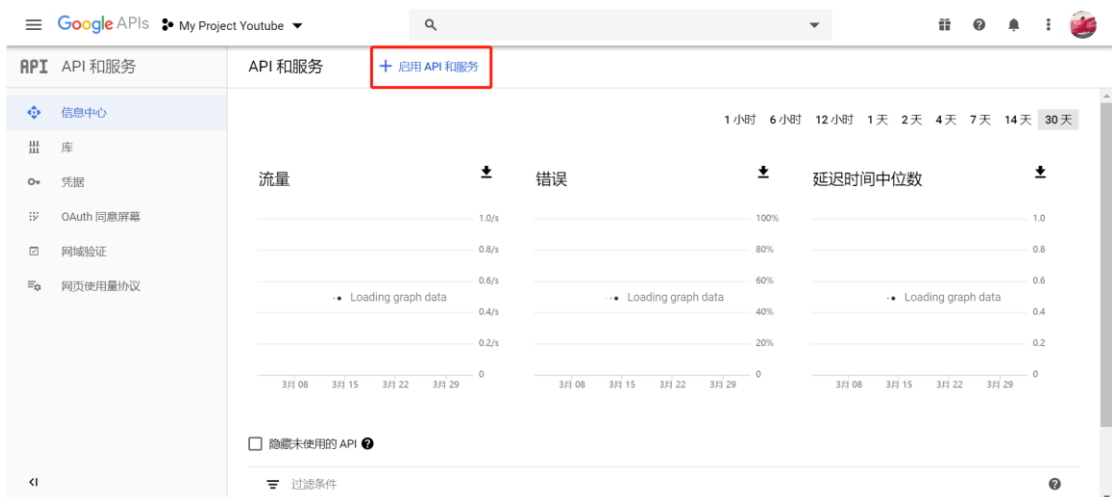
(4) Then name your new project. (Once you create the project, the project name is unchangeable!)

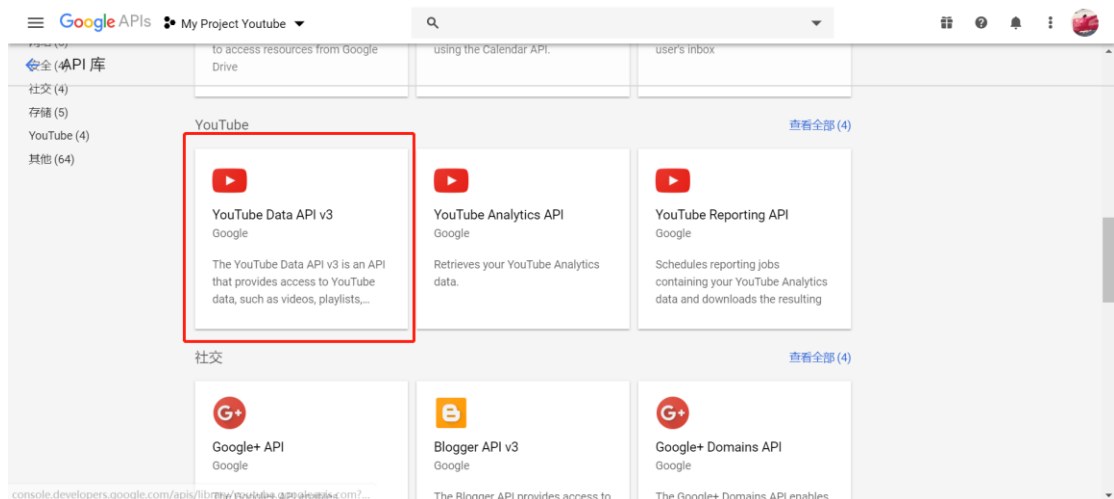


And it will show you that you created successfully.

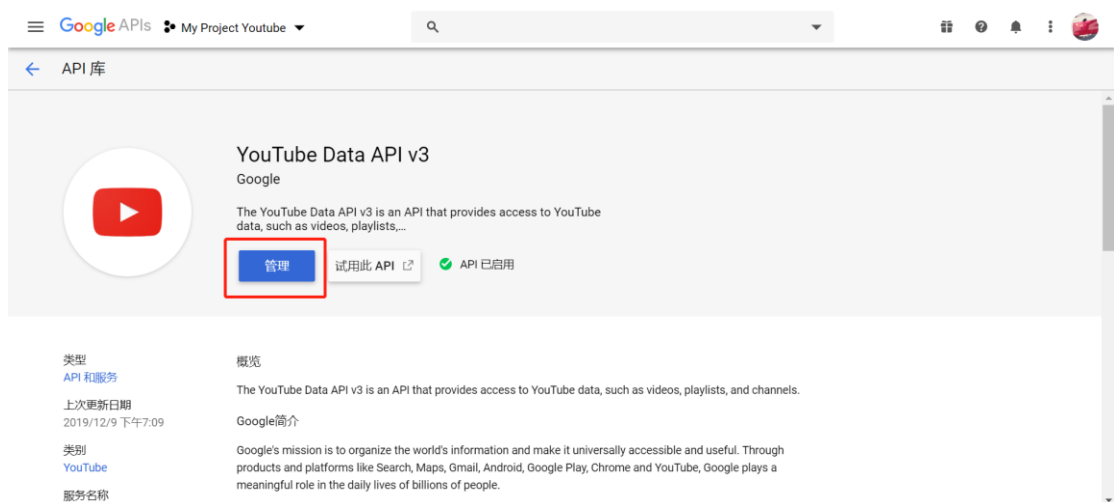


(5) Click “Enable API services” and scrolled down until you see something about YouTube:





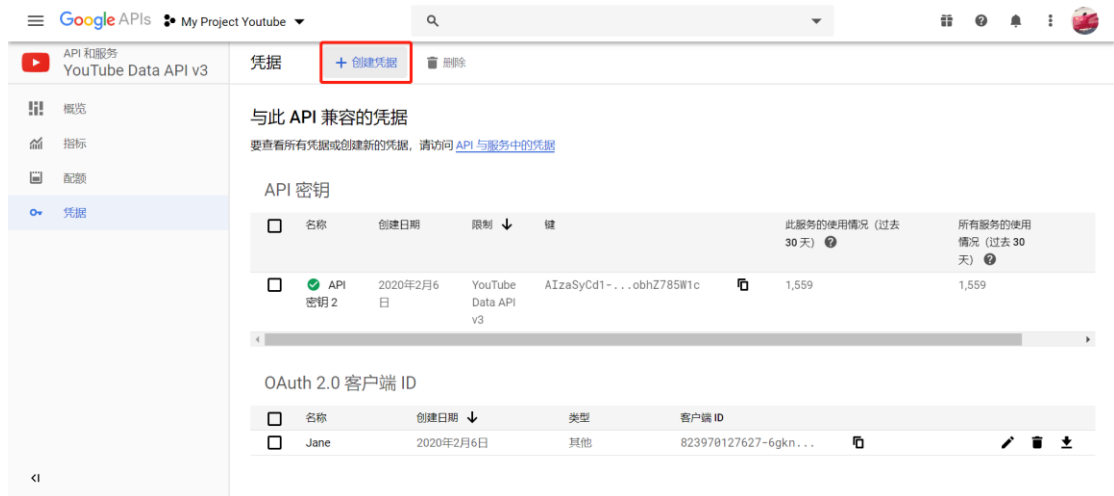
(6) Choose the first one which called “YouTube Data API v3”, click enable. (I’ve already enabled this one before, so it might look different from yours.)



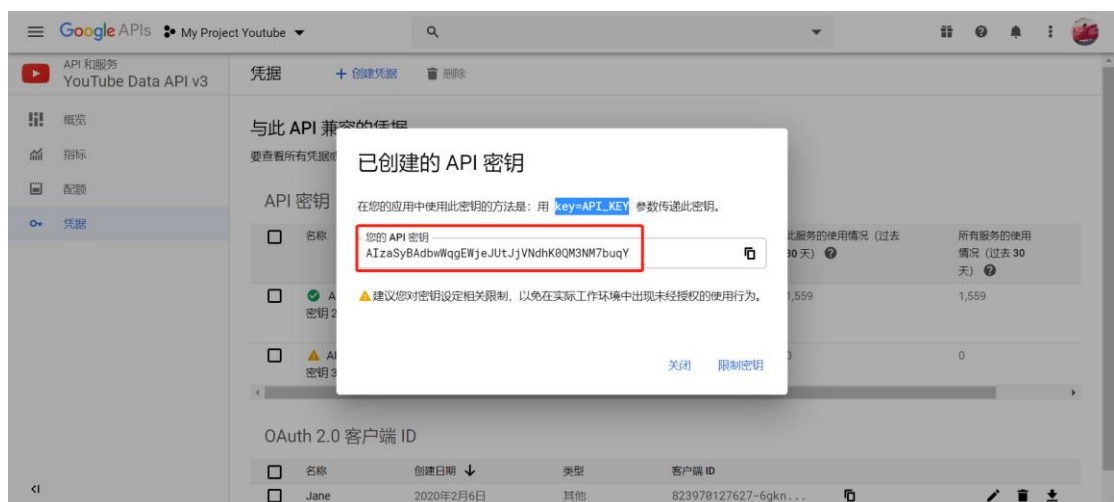
(7) Click “Credentials” on the left side.



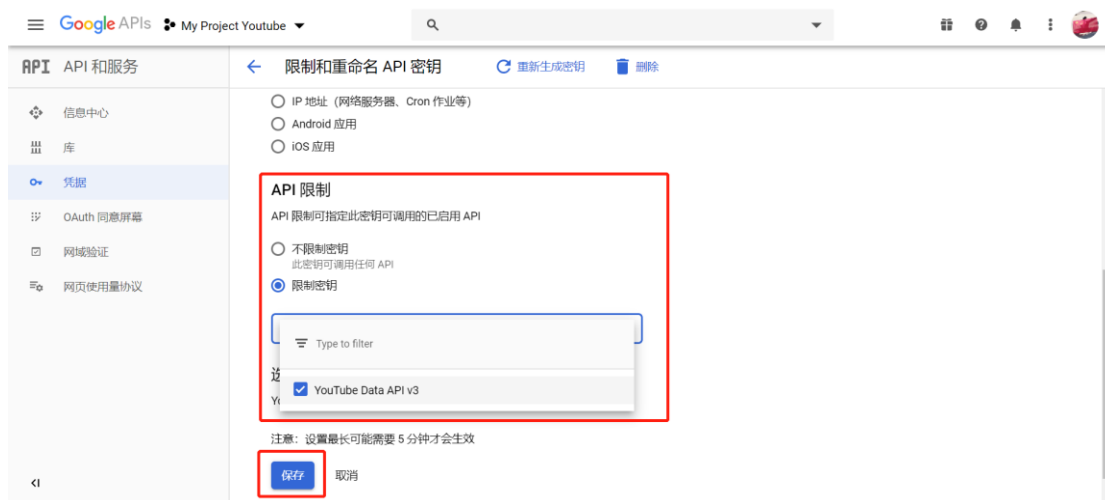
(8) Click “create a credential” and select “API Key”



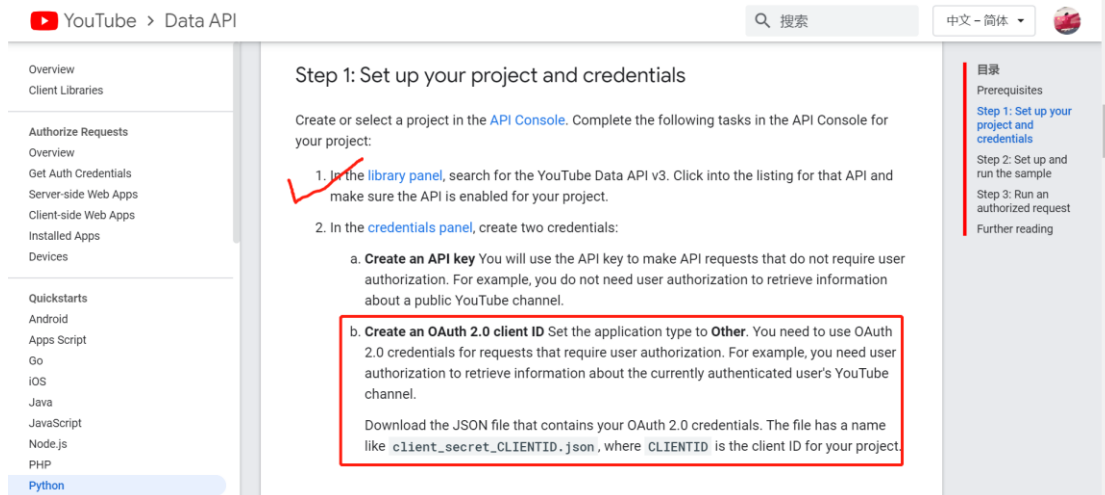
(9) Then you'll see your own API Key. (**Save your own API Key**, it is important, and you'll use it in the scraping code!!)



(10) For specific setting part, select “no” for Application Restrictions, select “YouTube Data API v3” for API Restrictions. Last save it!



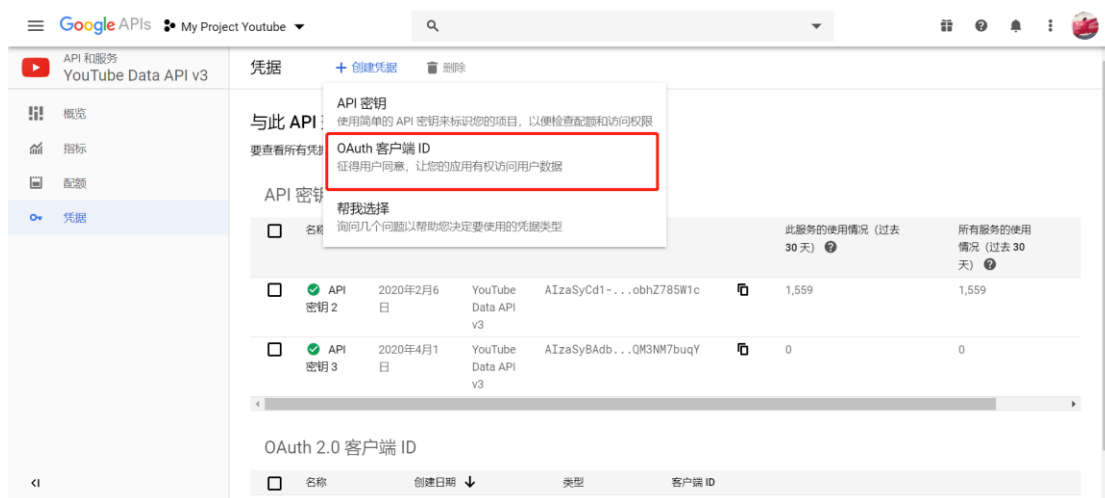
Now we finished creating API Key, next preparation step is creating an OAuth 2.0 client ID



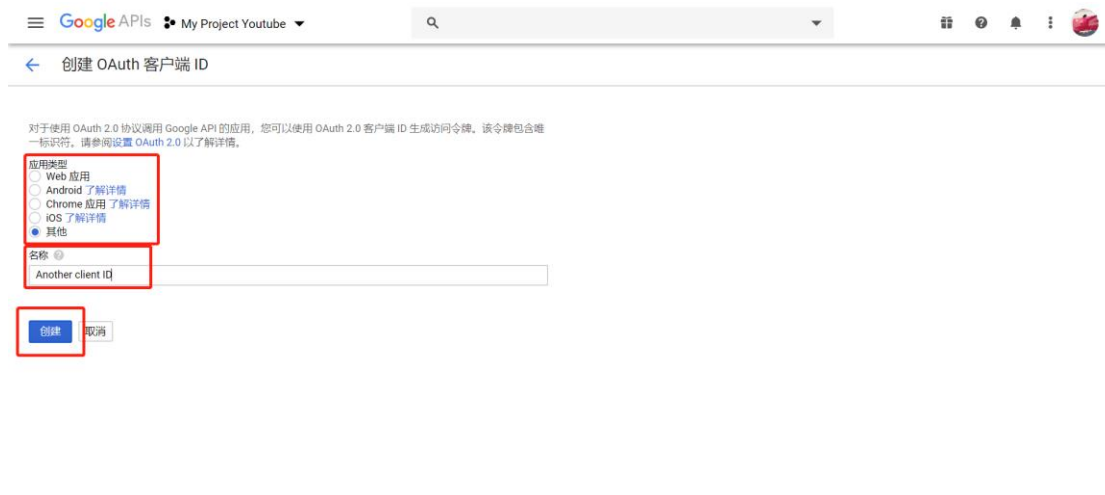
The picture above shows our process.

(11) Now, back to this google API page as below:

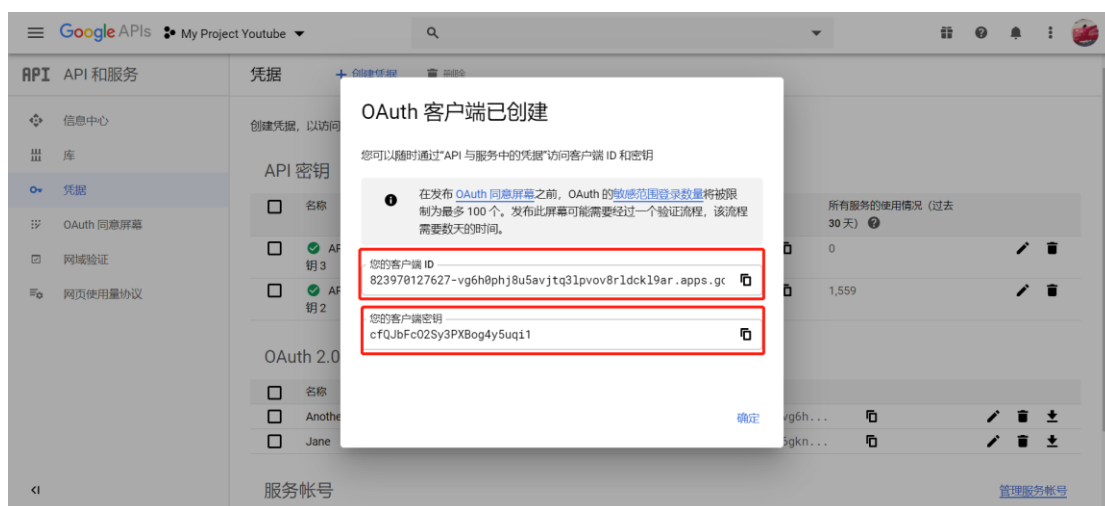
Similar to what we did for API Key, now click “create a credential” and select “OAuth client ID”



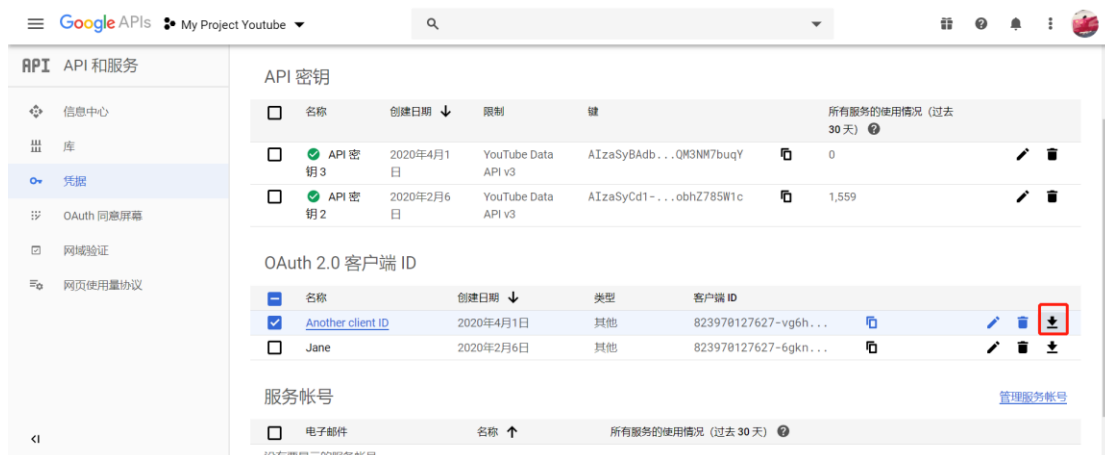
(12) Just click “other”, then name it and create it.



(13) And now you have client ID and client Key.



(14) Click download bar and download your client information json file. **(Downloading it!!!)**





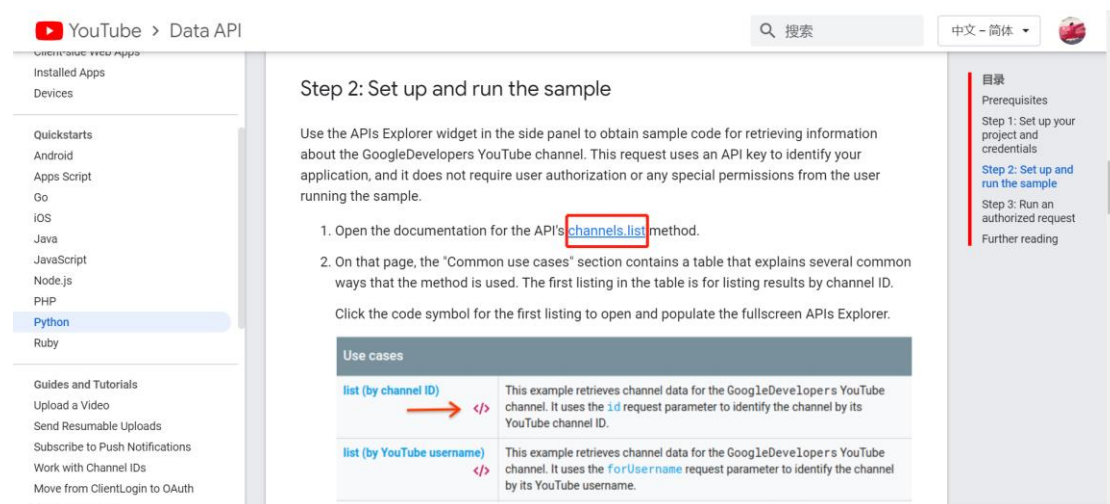
(15) Put it in the same folder with the python code file. It's better to rename it as

“client\_secret.json”. It depends on you; you do not have to rename it like this.

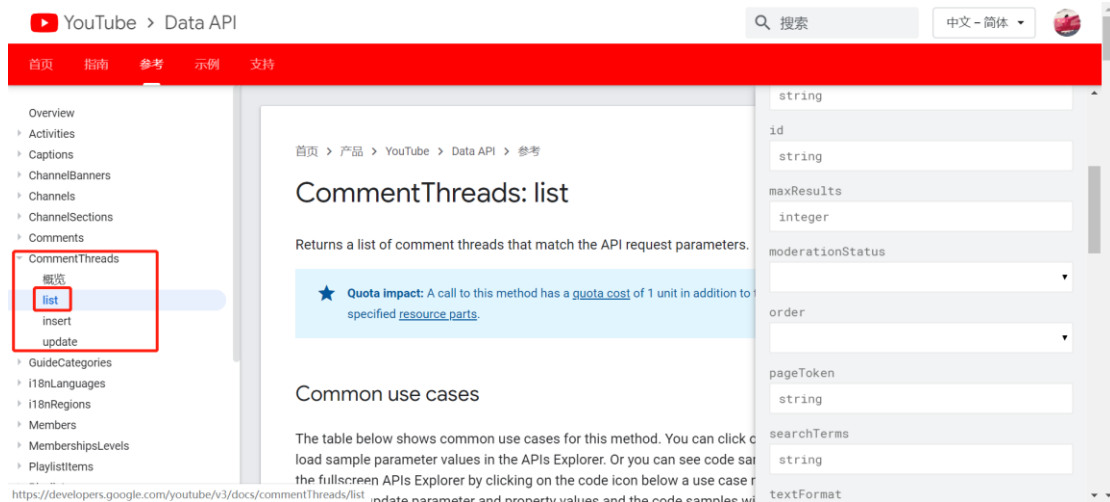
Overall, now you've got all you need. Next step is testing and running the sample. (Still need to follow this link: <https://developers.google.com/youtube/v3/quickstart/python.>)

## 2. Testing and Connecting:

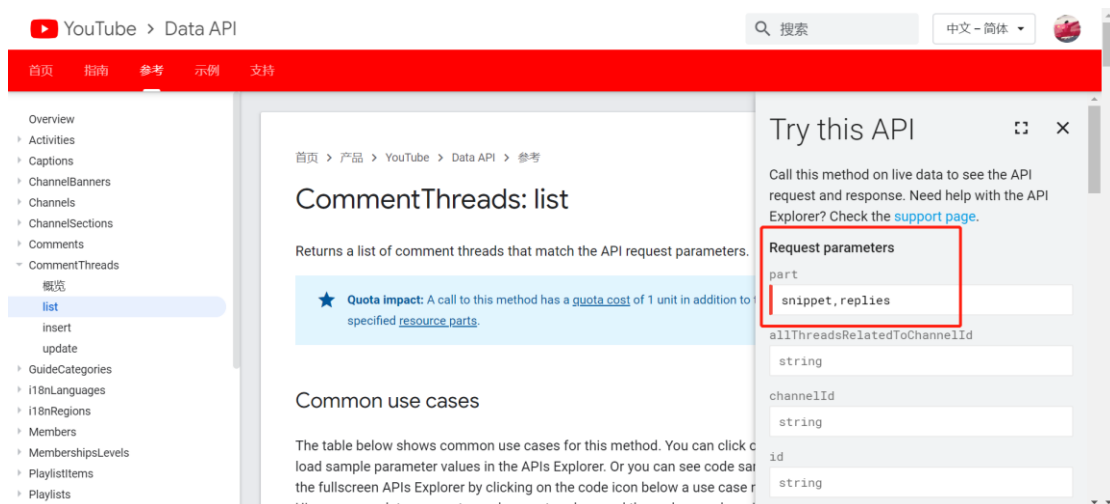
(1) Click “Channel id” and go another link.



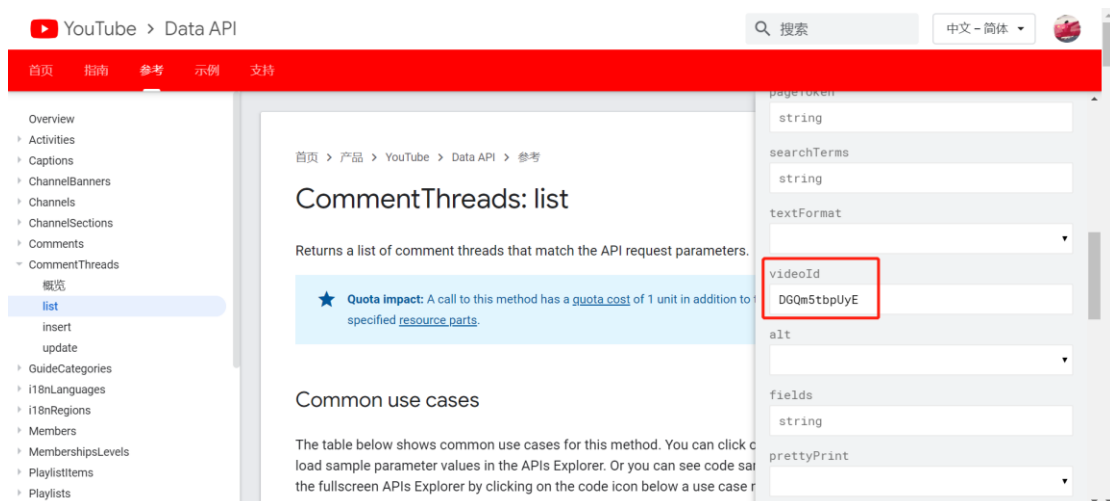
(2) At the new link, scrolled down and find “CommentThread” on the left side, after clicking that, you can see the drop-down box, then click “list”.



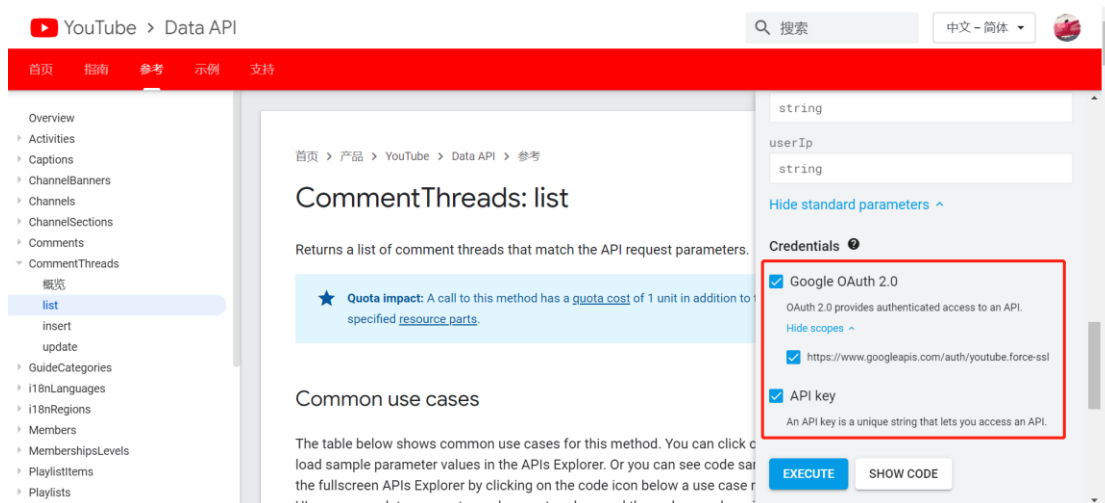
(3) For the right side of the page (Request parameters), type “snippet, replies” for “part”, then scroll down.



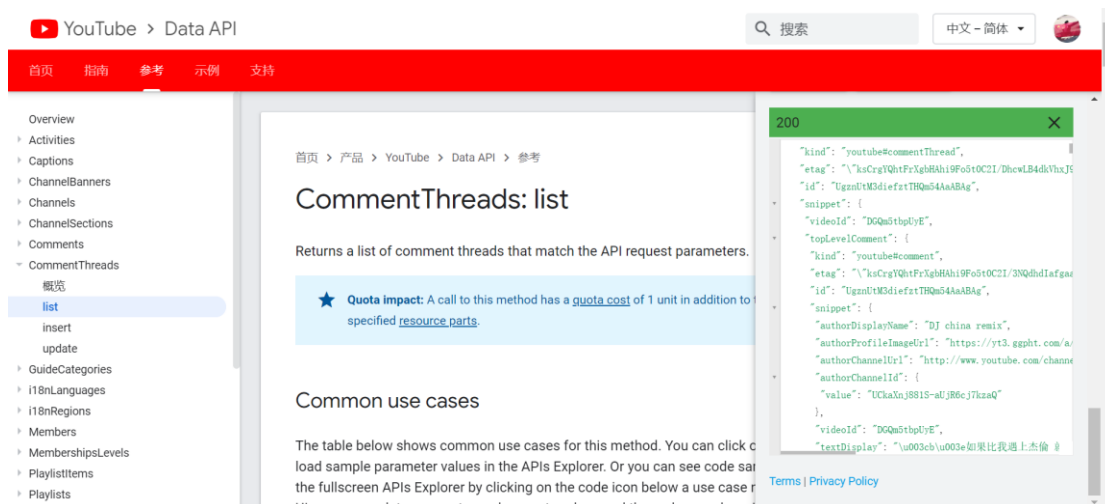
(4) Choose the video that you are interested in and type the videoID in this part.



(5) Make sure you mark all bars, then you can execute it.



(6) If there's no errors like this, then it means you've successfully done this!

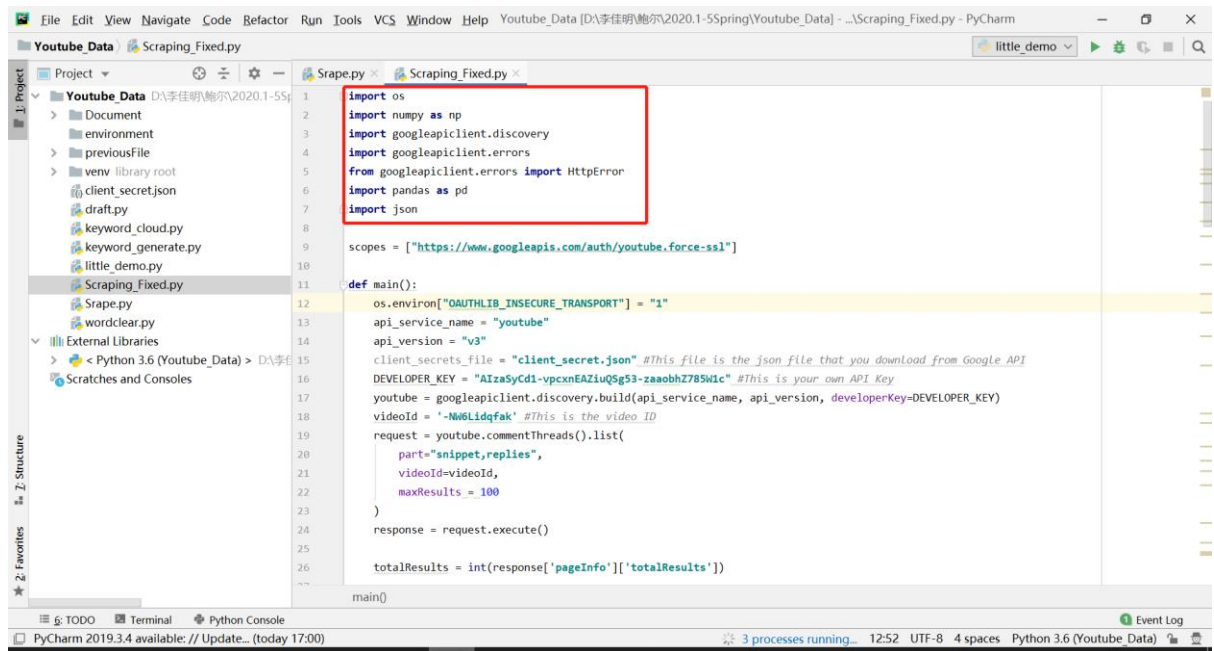


### 3. Writing Python Code:

There are several parts that need to pay attention to:

(1) Install Python3.6 as your PyCharm environment.

(2) Install packages of Python in this screenshot:



(3) Some parameters or information that need to be changed:

