M5. API를 활용한 데이터 수집



커리큘럼

과정 목표

- 01 온라인(웹) 상의 다양한 데이터를 수집하기 위해 데이터의 형태와 웹의 기본구조를 이해한다.
- 02 크롤링을 위해 파이썬의 다양한 라이브러리의 기능과 역할을 학습한다.
- 03 실제 온라인에서 제공되는 다양한 형태의 데이터들을 원하는 방식으로 수집할 수 있도록 코드를 작성한다.

모듈 명	주요 내용
M1. 데이터의 형태	 데이터를 정의하고, 강의에서 활용할 데이터에 대한 유형을 소개 JSON, CSV, TSV, XLSX, 등 데이터의 형태를 소개하고 간단한 활용법 학습 Database에 대한 소개 및 SQLite를 활용한 간단한 활용법 숙지
M2. 웹의 기본 구 조	 웹의 개념과 간단한 역사 및 작동원리에 대한 내용을 파악 실제 다양한 웹페이지를 통해 웹을 구성하고 있는 요소(HTML, Javascript, JSON 등)들을 학습
M3. 크롤링을 위한 파이썬 라이브러리	 크롤링을 위한 다양한 파이썬 라이브러리에 대한 소개 주피터노트북을 활용하여 각 라이브러리에 대한 실습 진행
M4. 웹페이지를 활용한 데이터 수집	 실제 동작하는 웹페이지를 분석하고 수집하여, 원하는 형태의 데이터로 변환하는 실습 진행 대상: 네이버금융, 네이버뉴스, DART, 증권사이트 등
M5. API를 활용한 데이터 수집	 공공기관이나 민간기업에서 제공하고있는 OpenAPI를 활용하여 데이터를 수집하는 실습 진행 대상: 공공데이터(부동산,금융), SNS API, 쇼핑몰 API 등
M6. 데이터 수집 시스템 만들기	• 위에서 작성된 파이썬 코드를 활용하여, 정기적인 수집 시스템을 만드는 실습 진행

크롤링: 서버에 무리를 줄 수 있음

>> 데이터를 가져오기 위해 불필요한 서비스 요청이 많음

API: 서버에 무리를 주지 않음

<u>>> 원하는 데이터만 가져오기에 서버에 필요한 요청만 함</u>



1. Youtube API

2. 공공 데이터 API

3. FinHub(NASDAQ)



DATA API

Youtube에 공개된 동영상 및 채널의 정보 - 조회수, 동영상 좋아요 수 등

Analytics & Reporting

Youtube에 공개되지 않은 채널 소유자만 볼 수 있는 정보 - 지역별 구독자 분포, 구독자/비구독자의 조회 비율 등



DATA API

Youtube에 공개된 동영상 및 채널의 정보 - 조회수, 동영상 좋아요 수 등



리소스 및 리소스 유형

리소스는 고유한 식별자를 가진 개별 데이터 항목입니다. 아래 표는 API를 사용하여 상호작용할 수 있는 다양한 리소스 유형에 대해 설명합니다.

리소스 유형						
activity	특정 사용자가 YouTube 사이트에서 실행한 작업의 정보를 포함합니다. 활동 피드에 보고되는 사용자 작업에 는 동영상 평가, 동영상 공유, 동영상을 즐겨찾기에 추가, 채널 게시판에 게시 등이 포함됩니다.					
channel	단일 YouTube 채널에 대한 정보를 포함합니다.					
channelBanner	새로 업로드한 이미지를 채널의 배너 이미지로 설정하는 데 사용할 URL을 식별합니다.					
guideCategory	YouTube가 채널의 콘텐츠 또는 기타 지표(예: 인기도)를 기반으로 채널에 연결하는 카테고리를 식별합니다. 가이드 카테고리는 YouTube 사용자가 원하는 콘텐츠를 더 쉽게 찾을 수 있는 방법으로 채널을 구성하기 위한 방법을 모색합니다. 채널은 한 개 이상의 가이드 카테고리에 연결될 수 있지만 가이드 카테고리에 모두 포함된 다고는 할 수 없습니다.					
playlist	단일 YouTube 재생목록을 표시합니다. 재생목록은 순서대로 감상하거나 다른 사용자와 공유할 수 있는 동영 상의 모음입니다.					
playlistItem	재생목록에 포함된 동영상과 같은 리소스를 식별합니다. playlistItem 리소스에는 포함된 리소스가 재생목록에 서 사용되는 방식을 설명하는 세부정보도 포함되어 있습니다.					
search result	API 요청에 지정된 검색 매개변수와 일치하는 YouTube 동영상, 채널 또는 재생목록의 정보를 포함합니다. 검색 결과는 동영상과 같이 고유하게 식별할 수 있는 리소스를 보여주지만, 자체적으로는 영구적인 데이터를 가지지 않습니다.					
subscription	YouTube 사용자의 구독 정보를 포함합니다. 구독정보는 채널에 새 동영상이 추가되거나 다른 사용자가 YouTube에서 동영상 업로드, 동영상 평가 또는 동영상 추천 등의 작업 중 하나를 수행할 때 이를 알려줍니다.					
thumbnail	하나의 리소스에 연결된 미리보기 이미지를 식별합니다.					
video	단일 YouTube 동영상을 표시합니다.					
videoCategory	업로드된 동영상과 연결되었거나 연결할 수 있는 카테고리를 식별합니다.					



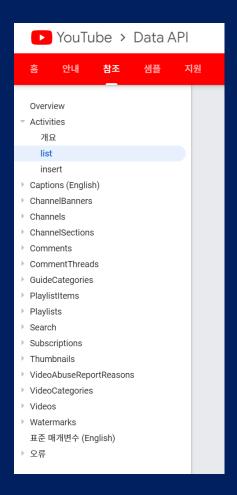
리소스 및 리소스 유형

리소스는 고유한 식별자를 가진 개별 데이터 항목입니다. 아래 표는 API를 사용하여 상호작용할 수 있는 다양한 리소스 유형에 대해 설명합니다.

리소스 유형	
activity	특정 사용자가 YouTube 사이트에서 실행한 작업의 정보를 포함합니다. 활동 피드에 보고되는 사용자 작업에 는 동영상 평가, 동영상 공유, 동영상을 즐겨찾기에 추가, 채널 게시판에 게시 등이 포함됩니다.
channel	단일 YouTube 채널에 대한 정보를 포함합니다.
channelBanner	새로 업로드한 이미지를 채널의 배너 이미지로 설정하는 데 사용할 URL을 식별합니다.
guideCategory	YouTube가 채널의 콘텐츠 또는 기타 지표(예: 인기도)를 기반으로 채널에 연결하는 카테고리를 식별합니다. 가이드 카테고리는 YouTube 사용자가 원하는 콘텐츠를 더 쉽게 찾을 수 있는 방법으로 채널을 구성하기 위한 방법을 모색합니다. 채널은 한 개 이상의 가이드 카테고리에 연결될 수 있지만 가이드 카테고리에 모두 포함된 다고는 할 수 없습니다.
playlist	단일 YouTube 재생목록을 표시합니다. 재생목록은 순서대로 감상하거나 다른 사용자와 공유할 수 있는 동영 상의 모음입니다.
playlistItem	재생목록에 포함된 동영상과 같은 리소스를 식별합니다. playlistItem 리소스에는 포함된 리소스가 재생목록에 서 사용되는 방식을 설명하는 세부정보도 포함되어 있습니다.
search result	API 요청에 지정된 검색 매개변수와 일치하는 YouTube 동영상, 채널 또는 재생목록의 정보를 포함합니다. 검색 결과는 동영상과 같이 고유하게 식별할 수 있는 리소스를 보여주지만, 자체적으로는 영구적인 데이터를 가지지 않습니다.
subscription	YouTube 사용자의 구독 정보를 포함합니다. 구독정보는 채널에 새 동영상이 추가되거나 다른 사용자가 YouTube에서 동영상 업로드, 동영상 평가 또는 동영상 추천 등의 작업 중 하나를 수행할 때 이를 알려줍니다.
thumbnail	하나의 리소스에 연결된 미리보기 이미지를 식별합니다.
video	단일 YouTube 동영상을 표시합니다.
videoCategory	업로드된 동영상과 연결되었거나 연결할 수 있는 카테고리를 식별합니다.



Youtube API - 사용해보기



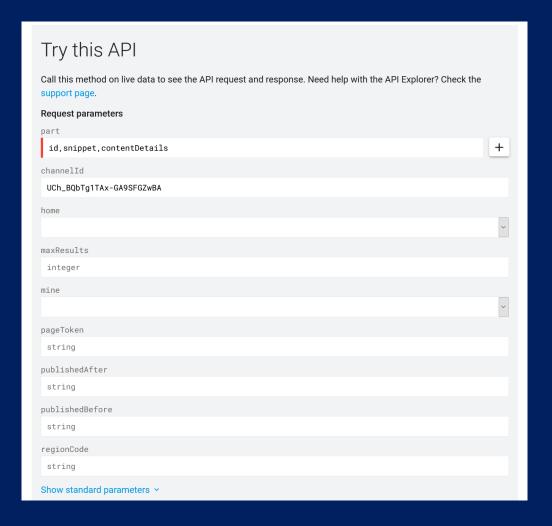
1)좌측 사진의 메뉴로 접속

https://developers.google.com/youtube/v3/docs/activities/list?hl=ko

- 2) 스크롤 하단 사용해보기
- 3) 필요한 데이터 넣기

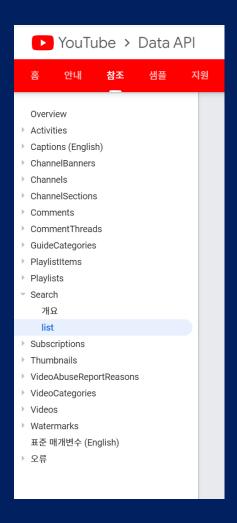


Youtube API - **사용해보기**





Youtube API - 사용해보기



1)좌측 사진의 메뉴로 접속

https://developers.google.com/youtube/v3/docs/search/list?hl=ko

- 2) 스크롤 하단 사용해보기
- 3) 필요한 데이터 넣기



직접 사용해 보세요! 🖘						
API Explorer를 사용하여 실시간 데이터에서 이 메소드를 호출하고 API 요청 및 응답을 확인해 보세요.						
Try this API						
Call this method on live data to see the API request and response. Need help with the API Explorer? Check the support page.						
Request parameters						
part						
id,snippet +						
channelId						
string						
channelType						
publishedBefore						
string						
q						
금융위원회						
regionCode						
string						



- 1) Youtube 용 프로젝트 생성
- 2) API Key 발급

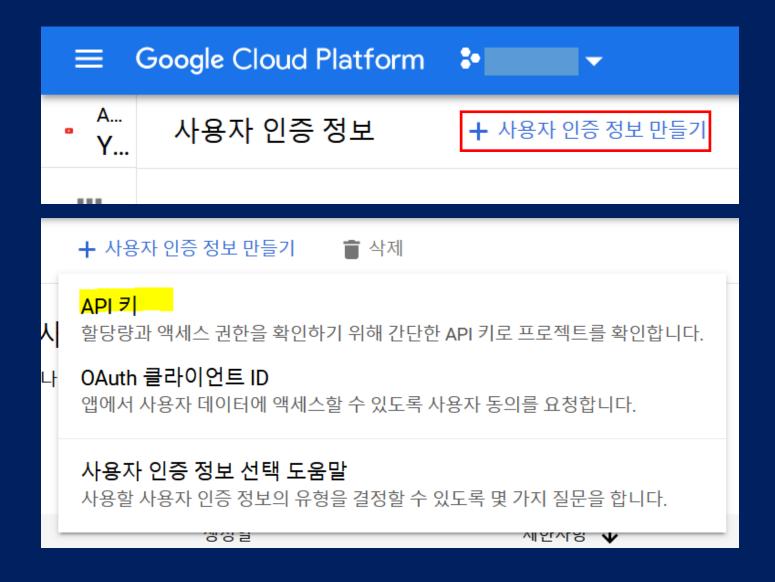
: 영구적이며 타인에게 노출해서는 안되는 중요한 정보



- 1) Google Cloud Platform 접속
- 2) 프로젝트 생성
- 3) Youtube Data API 추가



Youtube API - API 키 발급





공공데이터란 공공기관이 만들어내는 모든 자료나 정보, 국민 모두의 소통과 협력을 이끌어내는 공적인 정보를 말합니다. 각 공공기관이 보유한 공 공데이터 목록과 국민에게 개방할 수 있는 공공데이터를 포털에 등록하면 모두가 공유할 수 있는 양질의 공공데이터로 재탄생하게 됩니다.

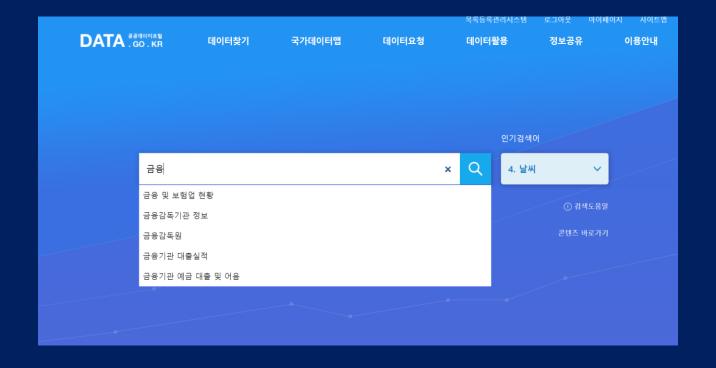


공공데이터 포털

https://www.data.go.kr/



공공데이터 종류 및 활용방법



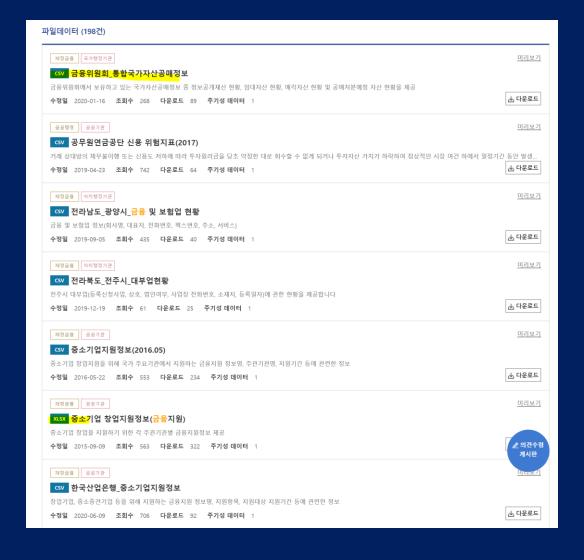


공공데이터 종류 및 활용방법





공공데이터 종류 - 1) 파일데이터





금융위원회 통합국가자산공매정보

https://www.data.go.kr/data/15042593/fileData.do

다운로드 및 실습

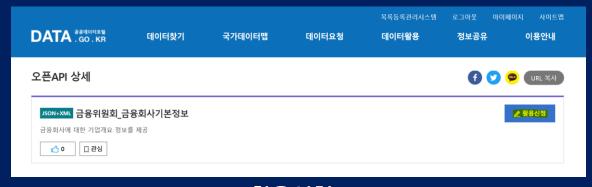


금융위원회_금융회사기본정보

https://www.data.go.kr/data/15043232/openapi.do

DATA ಪ್ರತಿಕಾಣಕ್ಕೆ	데이터찾기	국가데이터맵	데이터요청	목록등록관리시스템 데이터활용	로그아웃 마이 [‡] 정보공유	네이지 사이트맵 이용안내
오픈API 상세					6 🔾	WRL 복사
JSON+XML 금융위원회_급 금융회사에 대한 기업개요 정보 소 0 □ 관심					ı	☑ 활용신청





활용신청



공공데이터 종류 - 2) 오픈 API

JSON+XML 금융위원호	I_금융회사기본정보			시스템위	유형 선택			
제공기관	금융위원회	서비스유형	REST	LI A FIL	o න්			
심의여부	자동승인	신청유형	개발계정 활용신청	● 일반 ○ 서버구축 [안내] 서버 구축 여부에 따라 시스템 유형을 입력하실 수 있습니다 일반 : OPEN API 서비스를 호출하여 응답 받은 결과값을 서버에 저장하지 않고 사용할 겨우 (서버 미구				
처리상태	신청	활용기간	승인일로부터 24개월 간 활용가능				겨우 (서버 미구축)	
공공데이터 제공제도							API 서비스를 호출하여 응답 받은 결과값을 서버에 저장하거나 DB화	
* 공공데이터중 위치정보를 포함한 서비스를 사용하고자 하는 사업자는 ' 위치정보의 보호 및 이용 등에 관한 법률 '에 따라 방송통신위원회에 ' 위치정보서 비스 허가'를 받거나 '위치기반 서비스사업 신고'를 하여야 합니다. * 이에 해당하는 사업자인 경우에는 점부파일에 '위치기반서비스사업신고필증'을 점부해 주시기 바랍니다. * 활용신청 시 '위치기반서비스사업신고필증'이 등록되지 않으면 반려가 될 수 있으니 참고 하시기 바랍니다.		상세기능정보 선택						
		~	상세기	기능	설명	일일 트래픽		
활용목적 선택 *표시는 필수 입력항목입니다. *활용목적 웹 사이트 개발 ○ 앱개발 (모바임/솔루션등) ● 기타 ○ 참고자료 ○ 연구(논문 등)		· 🔽	금융회사기업개요조회		기준일자, 법인등록번호, 금융회사명을 통하여 사업자등록번호, 금융회사주소, 금융회사설립일자, 금융회사종업원수 등을조회하는 금융회사기업개요조회 기능	10000		
공공데이터 연습용			라이선스 표시					
9/250		*이용허락범위 이용허락범위 제한 없음 * 동의합니다.						
첨부파일	파일 선택							
	Drag & Drop으로 파일을 선택 가능	합니다.					취소 활용신청	

내용 작성 후 활용 신청



공공데이터 종류 - 2) 오픈 API





공공데이터 종류 - 2) 오픈 API: 활용 - 금융회사 정보

기준일자, 법인등록번호, 금융회사명을 통하여 사업자등록번호, 금융회사주소, 금융회사설립일자, 금융회사종업원수 등을 조회하는 금융회사기업개요조회 기 능

- 활용승인 절차 개발단계 : 허용 / 운영단계 : 허용
- · 신청가능 트래픽 1000000 / 운영계정은 활용사례 등록시 신청하면 트래픽 증가 가능
- · 요청주소 http://apis.data.go.kr/1160100/service/GetFnCoBasiInfoService/getFnCoOutl
- ・서비스URL http://apis.data.go.kr/1160100/service/GetFnCoBasiInfoService

<u></u> 활용신청

요청변수(Request Parameter)

항목명(국문)	항목명(영문)	항목크기	항목구분	샘플데이터	항목설명
서비스키	ServiceKey	4	필수	-	공공데이터포털에서 받은 인증키
페이지 번호	pageNo	4	필수	1	페이지번호
한 페이지 결과 수	numOfRows	4	필수	10	한 페이지 결과 수
결과형식	resultType	4	필수	xml	결과형식(xml/json)
기준일자	basDt	8	옵션	20200322	작업 또는 거래의 기준이 되는 일자(년월일)
법인등록번호	crno	13	옵션	1101113892240	법인등록번호
금융회사명	fncoNm	1000	옵션	메리츠자산운용	금융회사의 이름



공공데이터 종류 - 2) 오픈 API: 활용 - 단기금융증권 발행정보

금융위원회_단기금융증권발행정보

https://www.data.go.kr/tcs/dss/selectApiDataDetailView.do?publicDataPk=15059591

오픈API 상세	f y 😕 (URL 복사
JSON+XML 금융위원회_단기금융증권발행정보 단기기업어음발행기본정보, CD발행기본정보, 전자단기사체발행기본정보, ABCP발행기본정보등을 제공	∠ 활용신청



코드 실습

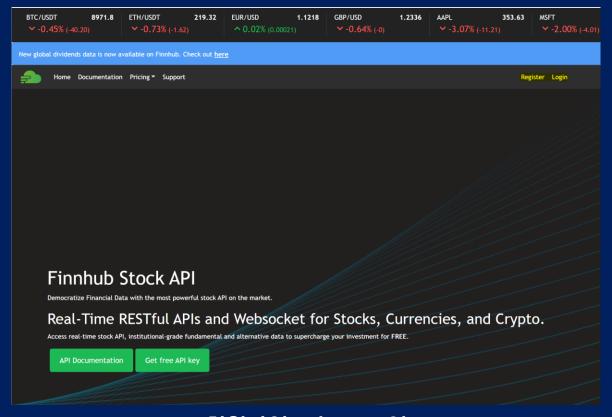


미국 나스닥(NASDAQ) 관련 정보를 받아올 수 있는 API

- 무료 제공(하지만, API 요청 제한 있음)
- 회사 정보 및 뉴스 등도 함께 볼 수 있음



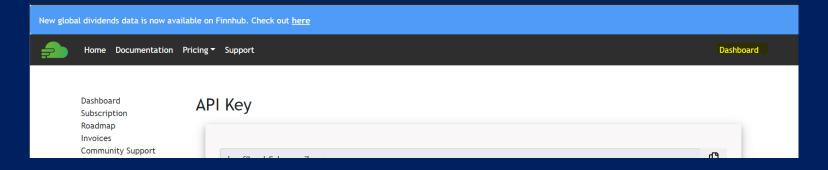
Finnhub API 사전 준비



회원가입 또는 로그인



Finnhub API 사전 준비



로그인 후 DashBoard에서 API Key 확인 가능



API 문서 링크

https://finnhub.io/docs/api#introduction



Finnhub API

Basic Financials Get company basic financials such as margin, P/E ratio, 52-week high/low etc. Method: GET **Examples:** /stock/metric?symbol=AAPL&metric=all /stock/metric?symbol=AMZN&metric=price **Arguments:** symbol REQUIRED Symbol of the company: AAPL. metric REQUIRED Metric type. Can be 1 of the following values [all, price, valuation, margin] **Response Attributes:** symbol Symbol of the company. metricType Metric type. metric Map key-value pair of metrics. Unit of metric type growth, and margin is (%). For other metric types, unit is reported currency which can be obtained with

Company Profile

요청 예시

```
Sample response

1 {
2     "metric": {
3         "10DayAverageTradingVolume": 3.87966,
4         "13WeekPriceReturnDaily": 9.87114,
5         "26WeekPriceReturnDaily": -6.21787,
6         "3MonthAverageTradingVolume": 64.24552,
7         "52WeekHigh": 2035.8,
8         "52WeekHighDate": "2019-07-11",
9         "52WeekLow": 1566.756,
10         "52WeekLowDate": "2019-02-08",
11         "52WeekLowDate": "2019-02-08",
12         "50ayPriceReturnDaily": 14.2094,
13         "beta": 1.60236,
14         "marketCapitalization": 938033.4,
15         },
16         "metricType": "price",
17         "symbol": "AMZN"
18     }
```

응답 예시



Finnhub API 활용



NASDAQ 종목 검색 방법

- 1) 구글 검색
- 2) 좌측 정보에서 확인

