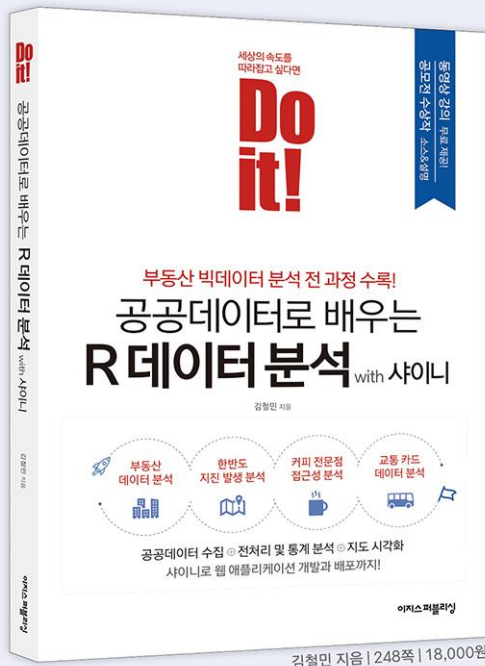


R로 공공데이터를 분석하는 전 과정 실습!
공모전 수상작으로 배우는 R 데이터 분석



공모전
수상작
소스 & 설명

동영상
강의
무료 제공

김철민 지음 | 248쪽 | 18,000원

02

자료 수집 전에 알아야 할 내용

02-1 자료는 어디서 구할까?

02-2 API 인증키 얻기

02-3 API에 자료 요청하기

02-4 API 응답 확인하기

자료 수집에 앞서 API 크롤링이 무엇인지 알아보고 공공데이터포털에서 API 인증키를 발급받는 방법을 알아봅니다. 또한 API를 이용해 자료를 요청하는 방법과 이에 따른 응답 결과가 어떤 구조인지 살펴봅니다.

02-1 자료는 어디서 구할까?

자료 수집은 데이터 분석 과정에서 맨 처음 만나는 관문입니다. 이때 어떤 자료를 어떻게 수집할지는 분석에 소요되는 시간과 비용 그리고 결과물의 수준을 결정합니다.

크롤러를 직접 만드는 일은 까다롭지만 공공데이터포털 같은 곳에서는 다양한 데이터를 이용할 수 있도록 API(application programming interface)를 제공합니다. 이 책에서 구하고자 하는 아파트 실거래 자료 역시 공공데이터포털에서 API 형식으로 제공합니다.

공공데이터포털이란?



DATA

- 1 공공데이터란?**
공공기관이 법령에 따라 생성, 수집하는 전자적 형태의 정보로써 개방가능한 모든 데이터
- 2 공공데이터포털이란?**
공공데이터를 한 곳에 모아 제공하는 통합 플랫폼

공공데이터포털 주요 서비스

- 1 데이터 찾기**
: 사용자가 원하는 데이터 (파일 및 오픈API 등)를 누구나 쉽고 빠르게 검색, 이용
(데이터목록, 국가통계데이터, 이슈데이터, 국가데이터랩)
- 2 데이터 요청**
: 현재 공공데이터포털에서 제공되지 않는 데이터를 직접 요청 가능
(데이터번거, 공공데이터 제공신청, 분쟁조정신청)
- 3 데이터 활용**
: 공공데이터를 활용하여 데이터를 분석 및 시각화
(공공데이터 시각화, 공공데이터 분석서비스, 국민참여지도)

02-2 API 인증키 얻기

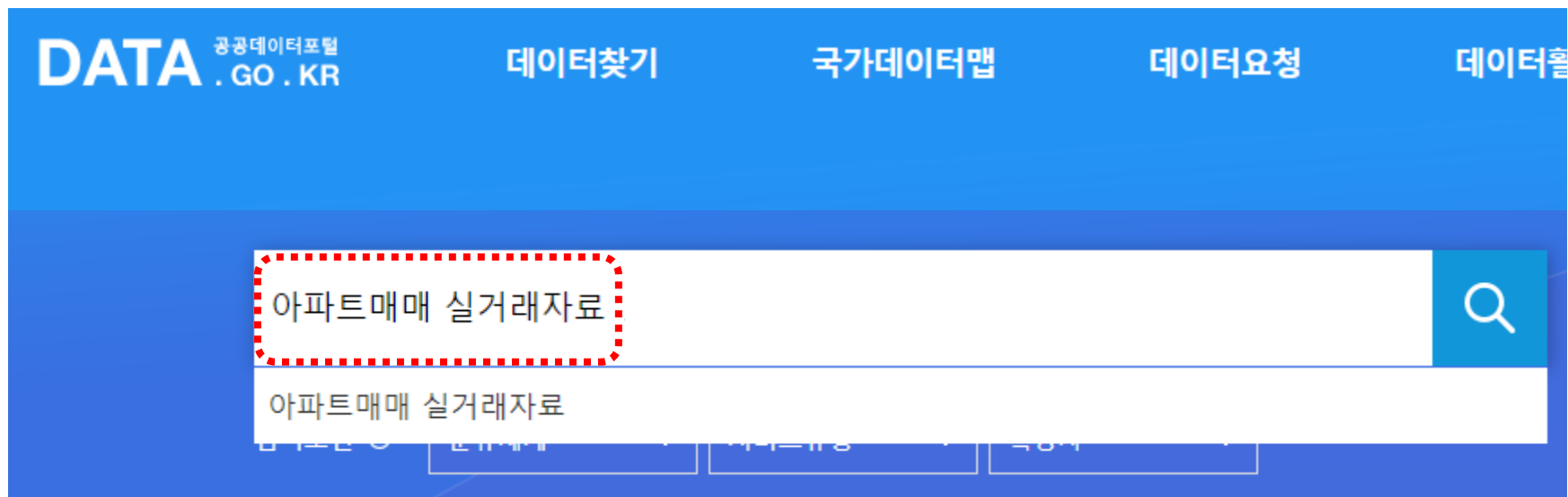
공공데이터포털에서 인증키를 발급받으려면 우선 웹 브라우저를 열고 공공데이터포털 (data.go.kr)에 접속합니다. 그리고 회원가입 후 로그인합니다.



02-2 API 인증키 얻기

1단계 API 활용 신청하기

공공데이터포털에서는 공공기관이 생성하고 수집한 데이터 가운데 개방할 수 있는 모든 데이터를 다양한 형식으로 제공합니다. 이 책에서는 국토교통부에서 제공하는 아파트매매 실거래자료가 필요하므로 이 키워드를 공공데이터포털 첫 화면에 있는 검색란에 입력하여 찾습니다.



02-2 API 인증키 얻기

1단계 API 활용 신청하기

검색 결과 화면에서 [오픈 API] 탭을 클릭한 후 목록에서 '국토교통부_아파트매매 실거래자료'가 보이면 오른쪽에서 <활용신청>을 클릭합니다.

전체(643건)

파일데이터(564건)

오픈 API(78건)

표준데이터셋1개(1건)

정확도순 ▼10개씩 ▼

오픈 API (78건)

공공행정

국가행정기관

국가중점

미리보기

XML

국토교통부_아파트매매 실거래자료

부동산 거래신고에 관한 법률에 따라 신고된 주택의 실거래 자료를 제공

제공기관 국토교통부 수정일 2021-11-12 조회수 38912 활용신청 12353 키워드 주택,아파트,실거래가

활용신청

02-2 API 인증키 얻기

1단계 API 활용 신청하기

그러면 다음처럼 'OpenAPI 개발계정 신청' 화면이 나옵니다. 이 화면에서 심의 여부를 확인해 보면 '자동승인'으로 되어 있습니다. 자동 승인은 따로 심사하지 않고 신청하면 곧바로 사용할 수 있음을 의미합니다.

www.data.go.kr 내용:

신청이 완료되었습니다. 신청된 API는 1~2시간 후 호출 가능합니다.

※ 일부 API에서 24시간 후 인증이 가능한 API가 있습니다.
빠른 시일내에 모든 API가 1시간 이내 인증 가능하도록 개선하겠습니다

확인

OpenAPI 개발계정 신청

XML 국토교통부_아파트매매 실거래자료

| | | | |
|------|-------|-------|--------------------|
| 제공기관 | 국토교통부 | 서비스유형 | REST |
| 심의여부 | 자동승인 | 신청유형 | 개발계정 활용신청 |
| 처리상태 | 신청 | 활용기간 | 승인일로부터 24개월 간 활용가능 |

공공데이터 제공제도

- * 공공데이터중 위치정보를 포함한 서비스를 사용하고자 하는 사업자는 '위치정보의 보호 및 이용 등에 관한 법률'에 따라 방송통신위원회에 '위치정보서비스 허가'를 받거나 '위치기반 서비스사업 신고'를 하여야 합니다.
- * 이에 해당하는 사업자인 경우에는 첨부파일에 '위치기반서비스사업신고필증'을 첨부해 주시기 바랍니다.
- * 활용신청 시 '위치기반서비스사업신고필증'이 등록되지 않으면 반려가 될 수 있으니 참고하시기 바랍니다.

활용목적 선택

*표시는 필수 입력항목입니다.

*활용목적

☒ 웹 사이트 개발 ☐ 앱개발 (모바일,솔루션등) ☐ 기타 ☐ 참고자료 ☐ 연구(논문 등)

테스트

3/250

확인

02-2 API 인증키 얻기

2단계 승인 및 인증키 확인하기

API 활용 신청을 마치고 승인이 났는지 확인하려면 마이페이지를 보면 됩니다. 공공데이터 포털 사이트 위쪽 메뉴에서 [마이페이지]를 클릭하면 다음과 같은 개발 계정 화면이 나옵니다. 이 화면에서 신청과 활용, 중지 건수를 확인할 수 있습니다. 신청한 자료가 승인되면 활용 건수로 바뀌고 아래 목록에서 해당 자료명 앞에 [승인]이 붙습니다.

DATA 공공데이터포털 .GO.KR

데이터찾기 국가데이터맵 데이터요청 데이터활용 정보공유 이용안내

마이페이지

오픈API >

DATA

나의 문의 >

나의 관심

공공행정 국토교통부

연장신청 [승인] 국토교통부_아파트매매 실거래 상세 자료

신청일 2021-04-18 만료예정일 2023-04-18

02-2 API 인증키 얻기

2단계 승인 및 인증키 확인하기

기본정보

| | | | |
|-------|-------------------------------------|------|------|
| 데이터명 | 국토교통부_아파트매매 실거래자료 상세설명 | | |
| 서비스유형 | REST | 심의여부 | 자동승인 |
| 신청유형 | 개발계정 활용신청 | 처리상태 | 승인 |
| 활용기간 | 2022-01-18 ~ 2024-01-18 | | |

승인이 완료된 자료를 클릭하면 상세보기 화면이 나타납니다. 상세보기 화면에는 여러 가지 정보가 나타나는데 이 가운데 일반 인증키(Encoding) 항목의 데이터가 이 API를 이용할 때 필요한 인증키입니다. 각자 발급받은 이 고유한 인증키로 서버에 자료를 요청합니다. 이 키를 따로 복사해 두었다가 나중에 코드를 작성할 때 활용할 수 있도록 합니다.

서비스정보

| | |
|---|--|
| 참고문서 | 아파트 매매 신고정보 조회 기술문서.hwp |
| 데이터포맷 | XML |
| End Point | http://openapi.molit.go.kr:8081/OpenAPI_ToolInstallPackage/service/rest/RTMSOBJSvc/getRTMSDataSvcAptTrade?_wddl&type=xml |
| API 환경 또는 API 호출 조건에 따라 인증키가 적용되는 방식이 다를 수 있습니다. 포털에서 제공되는 Encoding/Decoding 된 인증키를 적용하면서 구동되는 키를 사용하시기 바랍니다. * 향후 포털에서 더 명확한 정보를 제공하기 위해 노력하겠습니다. | |
| 일반 인증키 (Encoding) | R5rAjPY3n8V%2FBZ7jg0nF5v61b6vEiXN7k9KAW1ornFaqhrgNI91phc49FuHbC5%2Bra%2B2Q%2B3ec%2FBCDb8gqyhw1lA%3D%3D |
| 일반 인증키 (Decoding) | R5rAjPY3n8V/BZ7jg0nF5v61b6vEiXN7k9KAW1ornFaqhrgNI91phc49FuHbC5+ra+2Q+3ec/BCDb8gqyhw1lA== |

02-3 API에 자료 요청하기

1단계 서비스 URL 확인하기

기술 문서 4쪽을 보면 상세 기능 정보가 나오는데
여기서 'call Back URL' 항목의 주소가 바로 서비스 URL입니다.

표 2-1 상세 기능 정보

| 상세 기능 번호 | 1 | 상세 기능 유형 | 조회 (자료) |
|---------------|---|------------|----------|
| 상세 기능명(국문) | 아파트 매매 신고 정보 | | |
| 상세 기능 설명 | 행정표준코드관리시스템(www.code.go.kr)의 법정동 코드 중 앞 5자리(예시: 서울 종로구 - 11110), 거래 연월(예시: 201801)로 해당 지역, 해당 기간의 아파트 매매 신고 정보 조회 | | |
| Call Back URL | http://openapi.molit.go.kr:8081/OpenAPI_ToolInstallPackage/service/rest/RTMSOBJSvc/getRTMSDataSvcAptTrade | | |
| 최대 메시지 크기 | [1000 bytes] | | |
| 평균 응답 시간 | [500] ms | 초당 최대 트랜잭션 | [30] tps |

02-3 API에 자료 요청하기

2단계 요청 변수 확인하기

자료를 어디에 요청해야 할지 알았다면 그다음은 어떻게 요청해야 할지도 알아야 합니다. 기술 문서 4쪽에 나오는 요청 메시지 명세를 보면 아파트 실거래 자료를 요청할 때 다음 3가지 정보가 필요하다고 합니다. 이 정보를 조합해 하나의 요청 URL로 만들어서 API에 요청하면 서버가 해당 자료를 찾아서 보내 줍니다.

표 2-2 요청 메시지 명세

| 항목명(영어) | 항목명(한글) | 크기 | 구분 | 샘플 데이터 | 설명 |
|------------|---------|-----|----|-----------------|---------------------------|
| LAWD_CD | 지역 코드* | 5 | 1 | 11110 | 지역별 코드 (10자리 중 앞 5자리)* |
| DEAL_YMD | 거래 연월 | 6 | 1 | 201512 | 실거래 자료의 거래 연월 |
| serviceKey | 인증키 | 100 | 1 | 인증키(URL Encode) | 공공데이터포털에서 발급받는 인증키 |

* 지역 코드는 행정안전부의 행정표준코드관리시스템(www.code.go.kr)을 따릅니다.

02-3 API에 자료 요청하기

3단계 요청 URL 만들기

다음은 2015년 6월에 서울특별시 종로구에 등록된 아파트 매매 실거래 자료를 요청하는 예입니다. 지역 코드(11110)와 거래 연월(201506) 그리고 「02-2」절에서 발급받은 인증키를 채워 넣습니다. 이렇게 만든 URL을 웹 브라우저 주소 창에 입력하고 **Enter**를 누르면 서버로부터 XML 형태로 응답받은 데이터를 확인할 수 있습니다.

데이터 요청을 위한 URL 작성 예시

서비스URL?LAWD_CD=지역코드&DEAL_YMD=거래연월&serviceKey=인증키

02-4 API 응답 확인하기

1단계 응답 내역 알아보기

XML^{extensible markup language}은 컴퓨터끼리 데이터를 원활하게 전달하는 것을 목적으로 만든 언어입니다. 모든 XML 데이터는 루트 노드^{root node}에서 출발하며 부모-자식^{Parent-Child} 관계로 연결되어 있습니다.

우리가 받고자 하는 실거래 자료는 응답 순번이 저장된 루트 노드인 <response>에서 출발합니다. 이는 다시 <header>와 <body>라는 노드를 포함합니다. <header> 노드는 상태 정보를 알려 주는 <resultCode>와 상태 메시지가 담긴 <resultMsg>로 구성되어 있습니다. <body> 노드 아래에 있는 <items> 노드에는 세부 응답 내역이 있습니다.

예를 들어 다음 그림은 2015년 12월 서울특별시 종로구의 실거래 자료를 요청할 때 얻을 수 있는 XML 데이터를 보여 줍니다. 응답 상태와 응답 내역이 어떻게 구성돼 있는지 확인해 보세요.

표 2-3 실거래 API 응답 내역

| 항목명(영문) | 항목명(국문) | 항목 설명 | 크기 | 구분 | 샘플 데이터 |
|------------------------|-----------|-----------------------|-----|----|-----------------|
| resultCode | 결과 코드 | 결과 코드 | 2 | 1 | 00 |
| resultMsg | 결과 메시지 | 결과 메시지 | 50 | 1 | NORMAL SERVICE. |
| Deal Amount | 금액 | 거래 금액(만원) | 40 | 1 | 82,500 |
| Build Year | 건축 연도 | 건축 연도 | 4 | 1 | 2015 |
| Deal Year | 연도 | 거래 연도 | 4 | 1 | 2015 |
| Dong | 법정동 | 법정동 | 40 | 1 | 사직동 |
| Apartment Name | 아파트 | 아파트 이름 | 40 | 1 | 광화문 푸름스페이스본 |
| Deal Month | 월 | 거래 월 | 2 | 1 | 12 |
| Deal Day | 일 | 일 | 6 | 1 | 1 |
| Area for Exclusive Use | 전용면적 | 전용면적(m ²) | 20 | 1 | 94.51 |
| Jibun | 지번 | 지번 | 10 | 1 | 9 |
| Regional Code | 지역 코드 | 지역 코드 | 5 | 1 | 11110 |
| Floor | 층 | 층 | 4 | 1 | 11 |
| Cancel Deal Type | 해제 여부 | 해제 여부 | 1 | 0 | 0 |
| Cancel Deal Day | 해제 사유 발생일 | 해제 사유 발생일 | 8 | 0 | 21.01.27 |
| REQ GBN | 거래 유형 | 중개 및 직거래 여부 | 10 | 1 | 중개거래 |
| Rdealer Lawdnm | 중개업소 주소 | 시군구 단위 | 150 | 1 | 서울 서초구 |

* 2021년부터 허위매물거래 문제로 실거래 신고 해제여부(Cancel Deal Type), 해제사유발생일(Cancel Deal Day), 거래유형(REQ GBN), 중개업소주소(Rdealer Lawdnm) 같은 정보들이 추가되었으나 이번 자료 수집에서는 사용하지 않습니다.

02-3 API에 자료 요청하기

2단계 XML 형태의 응답 내역 확인하기

| 요청 URL | 응답 결과 |
|--|---|
| <code>http://openapi.molit.go.kr:8081/ OpenAPI_ToolInstallPackage/ service/rest/RTMSOBJSvc/ getRTMSDataSvcAptTrade? serviceKey=###인증키### &LAWD_CD=11110&DEAL_YMD=201507</code> | <pre>▼<response> ▼<header> <resultCode>00</resultCode> <resultMsg>NORMAL SERVICE.</resultMsg> </header> ▼<body> ▼<items> ▼<item> <거래금액> 65,000</거래금액> <거래유형> </거래유형> <건축년도>2000</건축년도> <년>2015</년> <법정동> 청운동 </법정동> <아파트>청운현대 </아파트> <월>7</월> <일>20</일> <전용면적>84.51</전용면적> <중개사소재지> </중개사소재지> <지번>56-45</지번> <지역코드>11110</지역코드> <층>5</층> <해제사유발생일> </해제사유발생일> <해제여부> </해제여부> </item> ▼<item> <거래금액> 100,000</거래금액> <거래유형> </거래유형></pre> |

그림 2-11 요청 URL과 응답 결과 예시