

미리 읽기

<https://tech.socarcorp.kr/data/2021/03/24/what-socar-data-engineering-team-does.html>

What is BigQuery

구글 빅쿼리란 구글에서 제공하는 **클라우드 데이터 웨어하우스**입니다. 부분적으로 무료로 사용할 수 있으며, 무료 제공되는 스토리지, 쿼리량에 상한이 정해져있고 그 이상으로 활용하시는 경우에는 과금이 될 수 있습니다.

신규 고객에게는 처음 90일 동안 사용할 수 있는 300달러 크레딧이 제공되고, 매월 10GB 스토리지, 최대 1TB의 쿼리가 무료로 제공됩니다.

특히 구글 애널리틱스를 이용해 데이터를 쌓고 있는 회사의 경우에는 GA 데이터를 빅쿼리에 연동해서 보는 경우가 많습니다. GA 보고서에서는 데이터의 요약값만 제공하기 때문에 데이터를 개별적으로 뜯어보기가 어려운데요. **SQL 쿼리를 이용하여 GA의 개별 데이터를 분석하고 싶을 때 빅쿼리를 연동해 사용하면 개별 데이터를 추출해서 볼 수 있습니다.**

물론 구글 애널리틱스를 사용하지 않아도 빅쿼리만 단독으로 사용할 수 있습니다. BigQuery에 CSV 데이터를 업로드하거나, 구글에서 제공하는 데이터셋을 연결하여 쿼리 해보고 분석 해 볼 수도 있습니다.

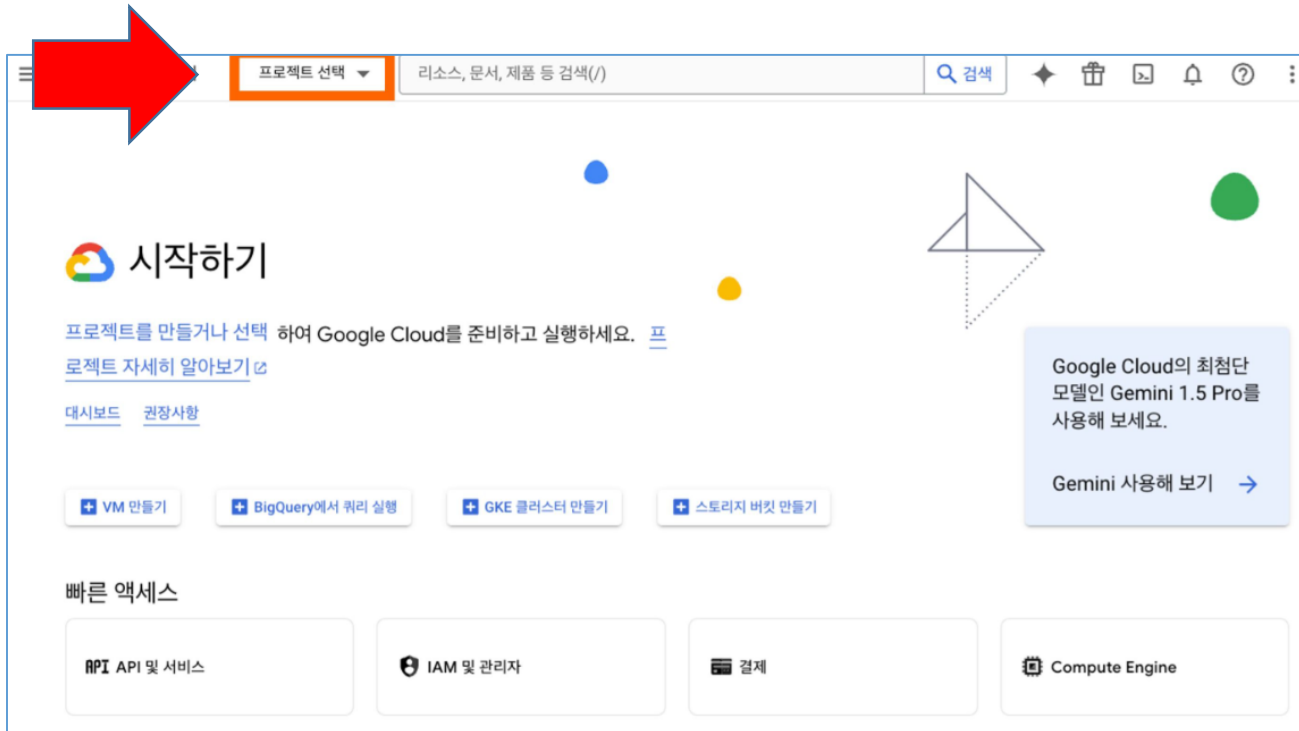
BigQuery 프로젝트 만들기

빅쿼리에서 데이터를 보기 위해서는 먼저 프로젝트를 만들어야 한다.

1. 구글클라우드콘솔 접속

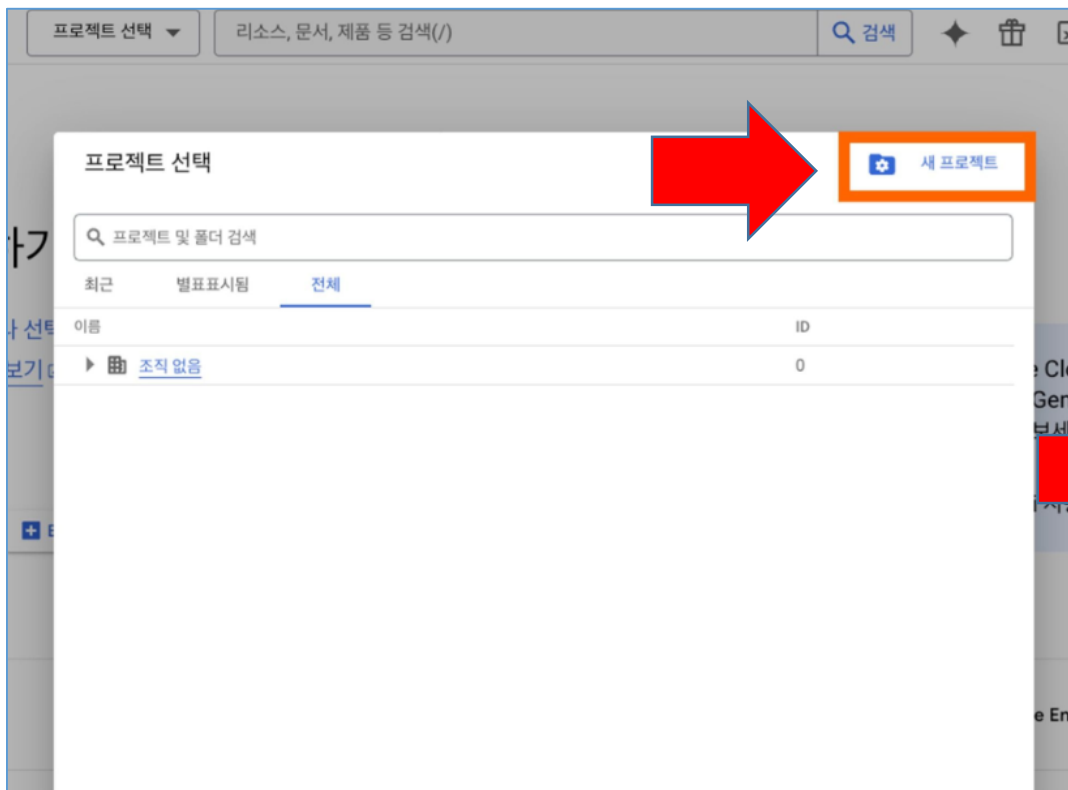
<https://console.cloud.google.com/welcome/new?hl=ko&inv=1&inv=Abouxg>

2. '프로젝트 선택' 클릭



BigQuery 프로젝트 만들기

3. 프로젝트 선택 창에서, '새 프로젝트' 클릭

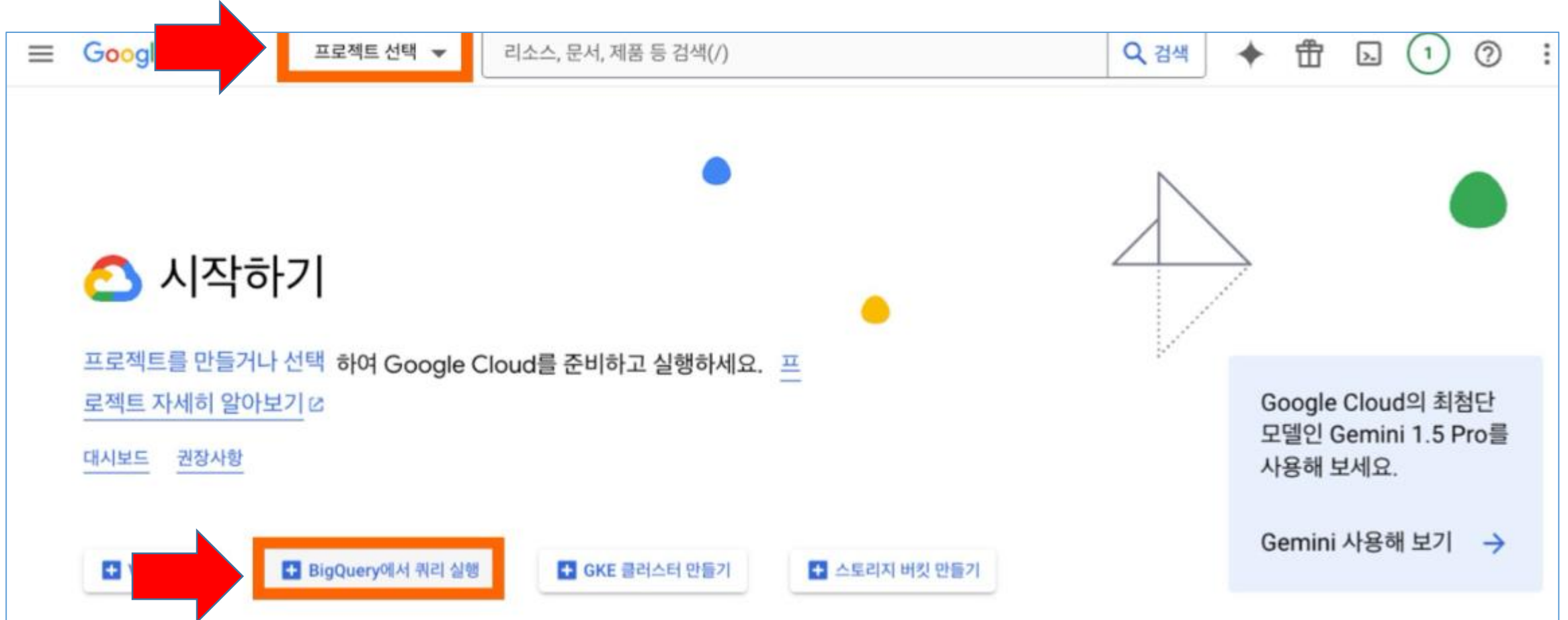


4. '프로젝트 이름' 설정



BigQuery 프로젝트 만들기

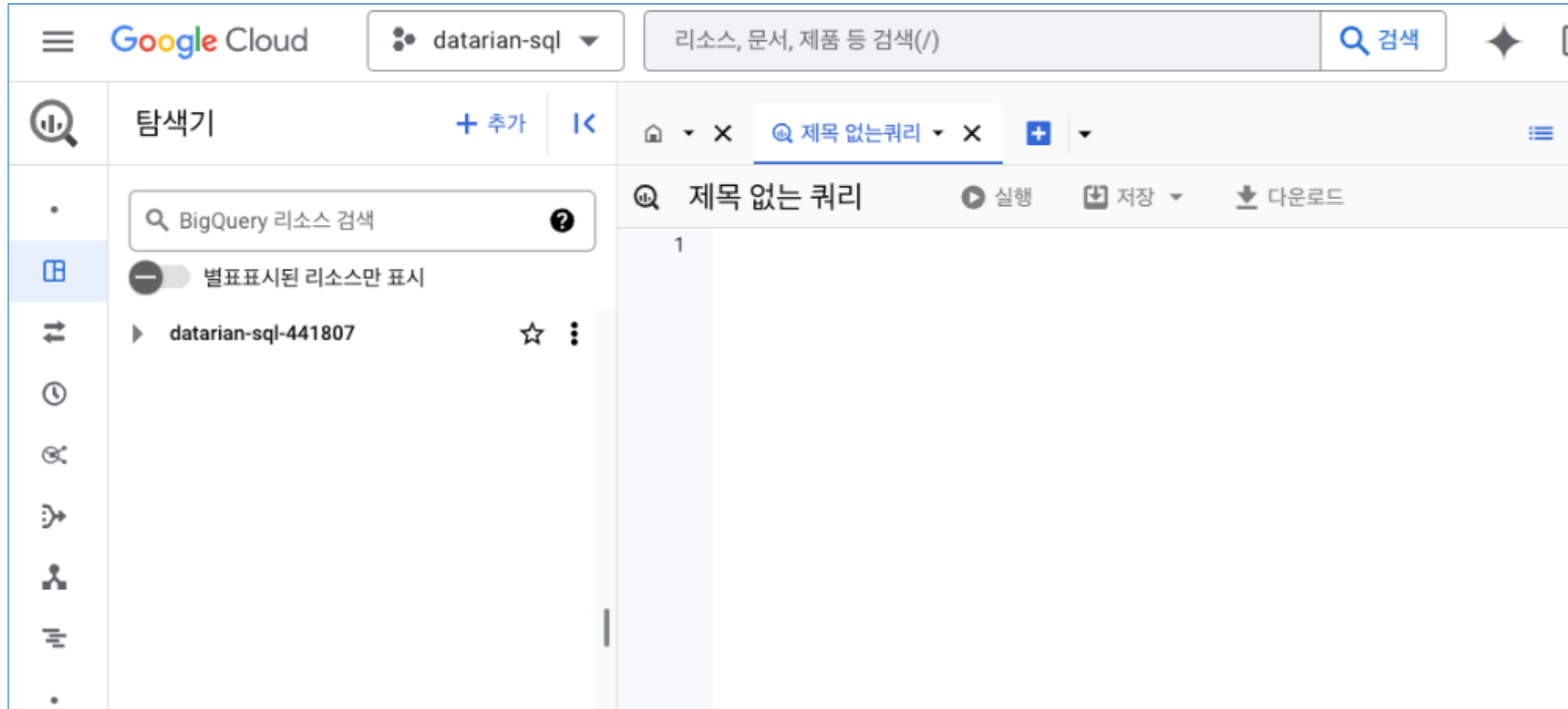
5.. 상단의 '프로젝트 선택' 클릭 후, 만든 프로젝트를 선택.
'BigQuery에서 쿼리 실행' 눌러 빅쿼리 콘솔 실행



BigQuery 프로젝트 만들기

5. 프로젝트 만들어 졌는지 확인

-> 자세한 내용 <https://cloud.google.com/bigquery/docs/sandbox?hl=ko>



BigQuery에 데이터를 연결하는 3가지 방법

구글 빅쿼리에서 데이터를 연결하는 방법은 크게 세 가지가 있습니다.

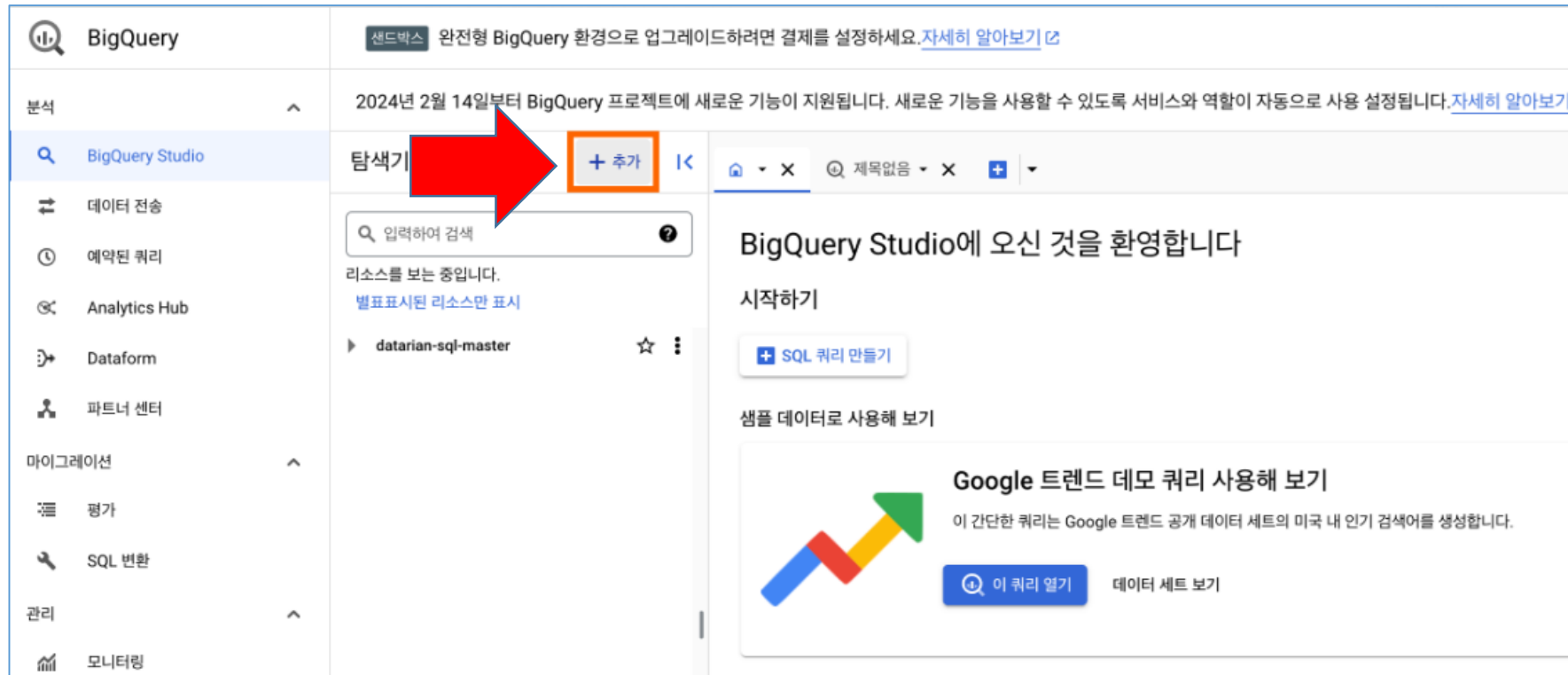
1. Google에서 제공하는 공개 데이터(Google Merchandise Store) 연결
2. CSV 데이터 BigQuery에 직접 업로드
3. GA4 데이터 BigQuery에 연결

설명할 부분은 많이 쓰이는 1번과 2번

Google Merchandise Store 연결

[Google Merchandise Store](#)는 Google 브랜드의 상품을 판매하는 전자상거래 사이트입니다. Analytics Hub에서 제공하는 Google Merchandise Store 데이터 셋에서는 트래픽 소스 데이터, 콘텐츠 데이터, 거래 데이터를 볼 수 있습니다.

1. 탐색기에서 '+ 추가' 버튼 클릭



The screenshot shows the BigQuery Studio interface. On the left is the '탐색기' (Explorer) panel, which contains a search bar and a list of resources. A red arrow points to the '+ 추가' (Add) button in the Explorer panel. The main workspace on the right displays a welcome message and a button to 'SQL 쿼리 만들기' (Create SQL Query).

BigQuery

샌드박스 완전형 BigQuery 환경으로 업그레이드하려면 결제를 설정하세요. [자세히 알아보기](#)

분석

2024년 2월 14일부터 BigQuery 프로젝트에 새로운 기능이 지원됩니다. 새로운 기능을 사용할 수 있도록 서비스와 역할이 자동으로 사용 설정됩니다. [자세히 알아보기](#)

BigQuery Studio

탐색기

+ 추가

입력하여 검색

리소스를 보는 중입니다.

[별표 표시된 리소스만 표시](#)

datarian-sql-master

BigQuery Studio에 오신 것을 환영합니다

시작하기

+ SQL 쿼리 만들기

샘플 데이터로 사용해 보기

Google 트렌드 데모 쿼리 사용해 보기

이 간단한 쿼리는 Google 트렌드 공개 데이터 세트의 미국 내 인기 검색어를 생성합니다.

이 쿼리 열기

데이터 세트 보기

Google Merchandise Store [연결](#)

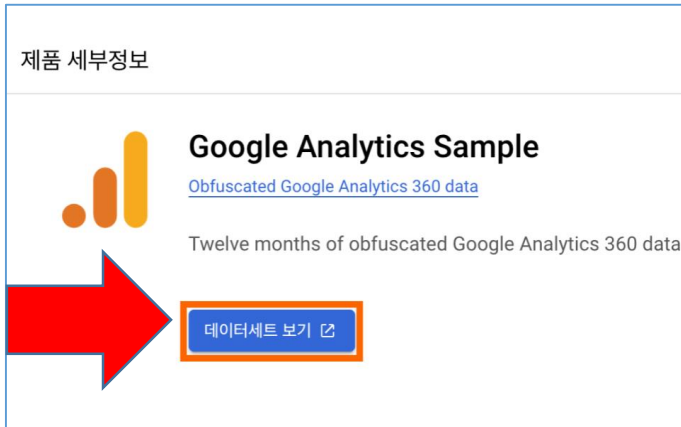
2. '공개 데이터 세트' 메뉴 클릭

- ▶ Data Transfer Service를 통한 Amazon 객체 스토리지 서비스
- ▶ Azure Blob Storage(및 Azure Data Lake Storage Gen2) - 데이터 전송
Data Transfer Service를 통한 Microsoft 객체 스토리지 서비스(및 데이터 레이크 스토리지)
- ▶ 공개 데이터 세트
Google Cloud 공개 데이터 세트 프로그램을 통해 제공되는 BigQuery 공개 데이터 세트
- ▶ Datastream
MySQL, PostgreSQL, Oracle과 같은 관계형 데이터베이스에서 실시간 CDC 복제
- ▶ Informatica 데이터 로더
코드 없이 로드합니다. 35개 이상의 소스가 100% 무료로 제공되며 용량 제한이 없습니다.

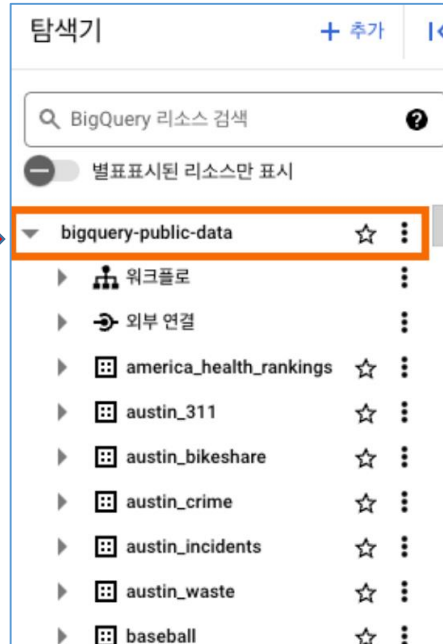
3. 검색창에서 'Merchandise' 검색

Google Merchandise Store 연결

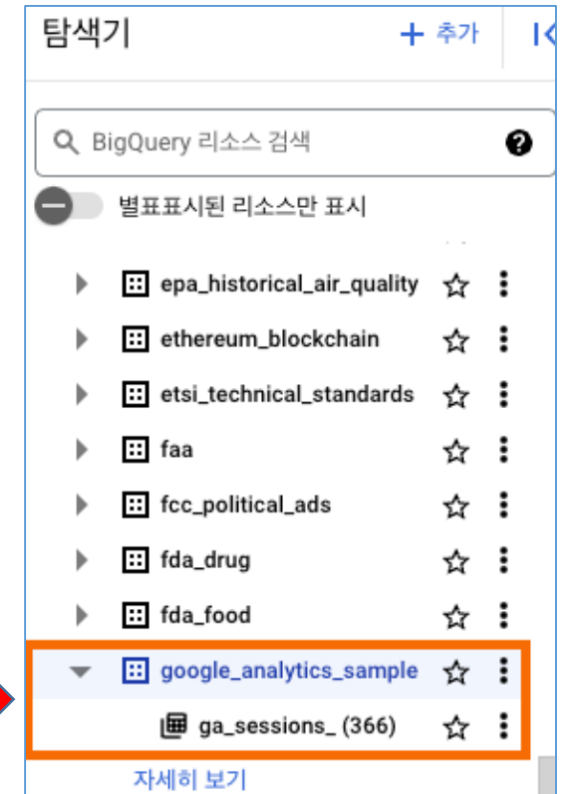
4. '데이터세트 보기' 클릭



5. 'bigquery-public-data' 프로젝트가 탐색기에 추가되었는지 확인

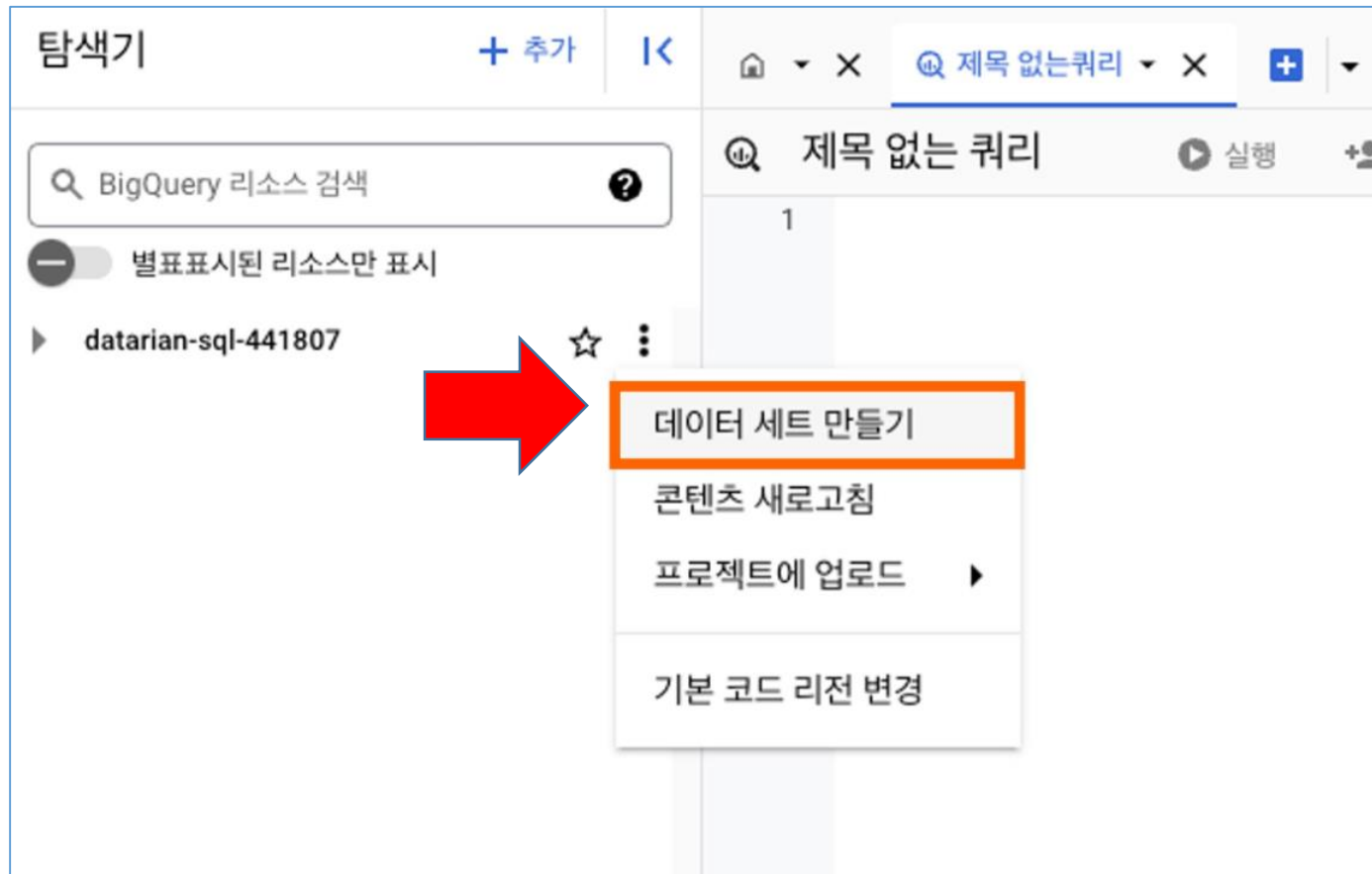


6. 왼쪽의 ▶ 아이콘을 클릭해 메뉴를 펼치면 'google_analytics_sample' 데이터 세트 확인



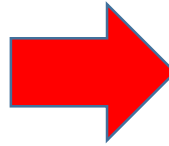
CSV 데이터 BigQuery에 업로드

1. 빅쿼리에 업로드할 CSV 데이터 준비
2. 프로젝트 오른쪽에 있는 더 보기 버튼을 클릭하여 '데이터 세트 만들기' 선택



CSV 데이터 BigQuery에 업로드

3. 데이터 세트 만들기 창에서 '데이터 세트 ID'를 지정해 주고
'데이터세트 만들기' 클릭



데이터세트 만들기

프로젝트 ID *
datarian-sql-441807 [변경](#)

데이터세트 ID *
waiters_tips
문자, 숫자, 밑줄이 허용됩니다.

위치 유형 ?
☐ 리전
데이터 세트를 다른 Google Cloud 서비스와 함께 배치할 리전을 지정합니다.
☒ 멀티 리전
더 높은 할당량 한도를 달성하기 위해 BigQuery에서 그룹 내의 리전을 선택하도록 허용합니다.

멀티 리전 *
US (미국 내 여러 리전) ▼

외부 데이터 세트
The selected region supports the following external dataset types: Cloud Spanner

☐ 외부 데이터 세트 링크 ?

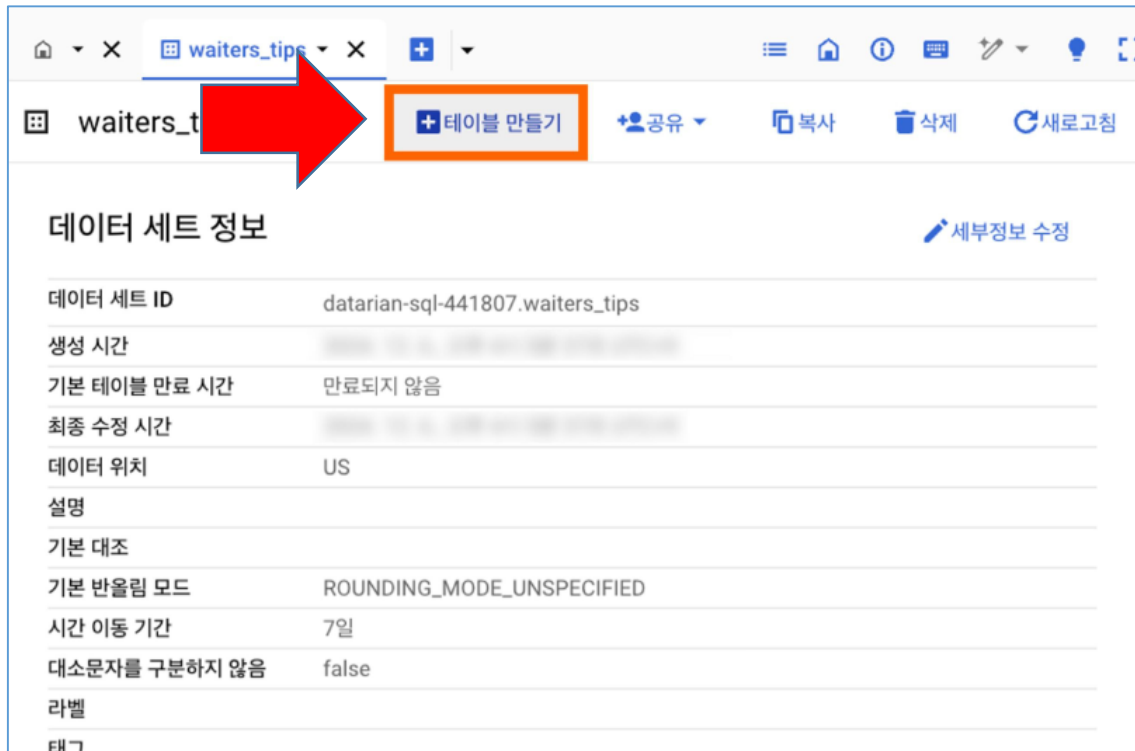
기본 테이블 만료
☐ 테이블 만료 사용 설정 ?

기본 최대 테이블 기간 일

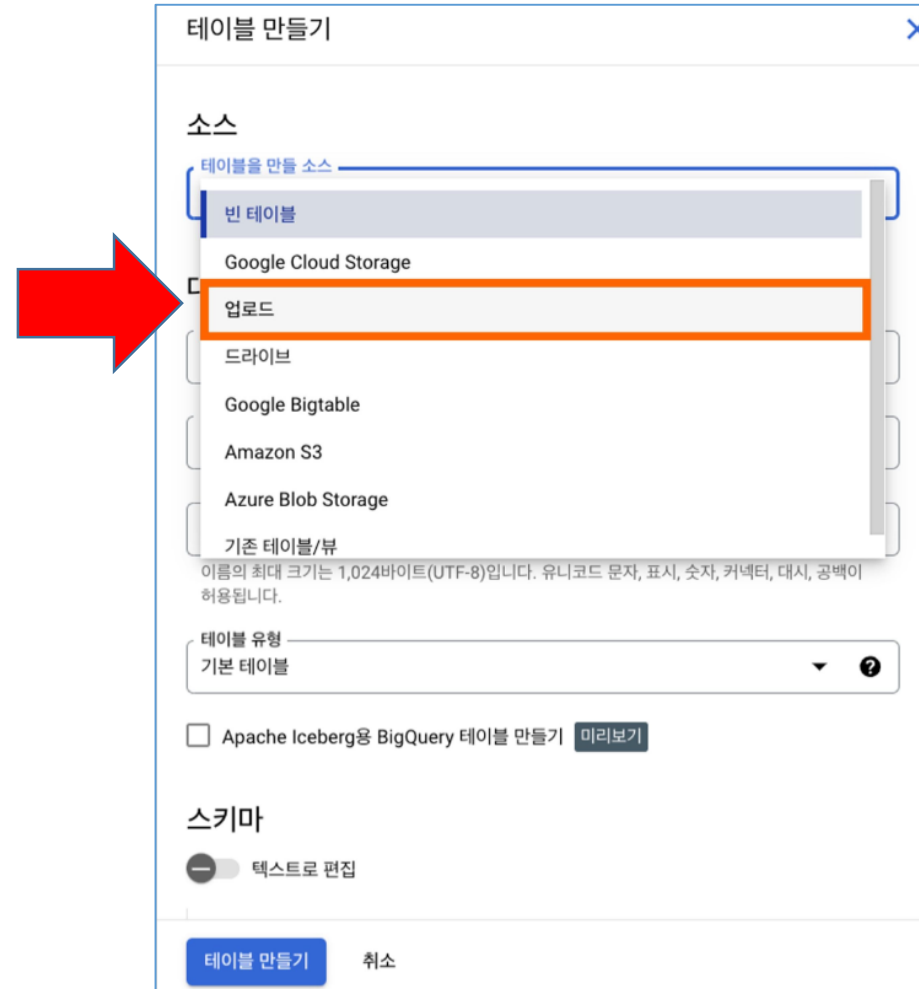
[데이터세트 만들기](#) [취소](#)

CSV 데이터 BigQuery에 업로드

4. 데이터 세트 정보 열람 창이 뜨면 오른쪽 상단에 위치한 '테이블 만들기' 버튼 클릭

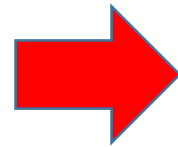


5. 상단에 테이블 만들 소스란에서 '업로드' 선택



CSV 데이터 BigQuery에 업로드

6. 프로젝트, 데이터 세트가 잘 설정되었는지 확인한 후 데이터를 저장할 테이블명 작성



테이블 만들기

소스

테이블을 만들 소스
업로드

파일 선택 *
tips_sample.csv X [찾아보기](#) ?

파일 형식
CSV

대상 위치

프로젝트 *
datarian-sql-441807 [찾아보기](#)

데이터 세트 *
waiters_tips

테이블 *

tips_sample

이름의 최대 크기는 1,024바이트(UTF-8)입니다. 유니코드 문자, 표시, 숫자, 커넥터, 대시, 공백이 허용됩니다.

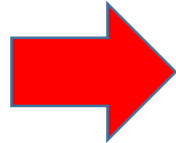
테이블 유형

기본 테이블

CSV 데이터 BigQuery에 업로드

7. 데이터 스키마를 설정. 특별히 컬럼명, 데이터 타입 수정이 필요한 것이 아니라면 **자동 감지**를 체크

컬럼명, 데이터 타입 수정이 필요한 경우에는 스키마 수동지정 기능을 이용



테이블 만들기

대상 위치

프로젝트 *
datarian-sql-441807
찾아보기

데이터 세트 *
waiters_tips

테이블 *
tips_sample
이름의 최대 크기는 1,024바이트(UTF-8)입니다. 유니코드 문자, 표시, 숫자, 커넥터, 대시, 공백이 허용됩니다.

테이블 유형
기본 테이블

스키마

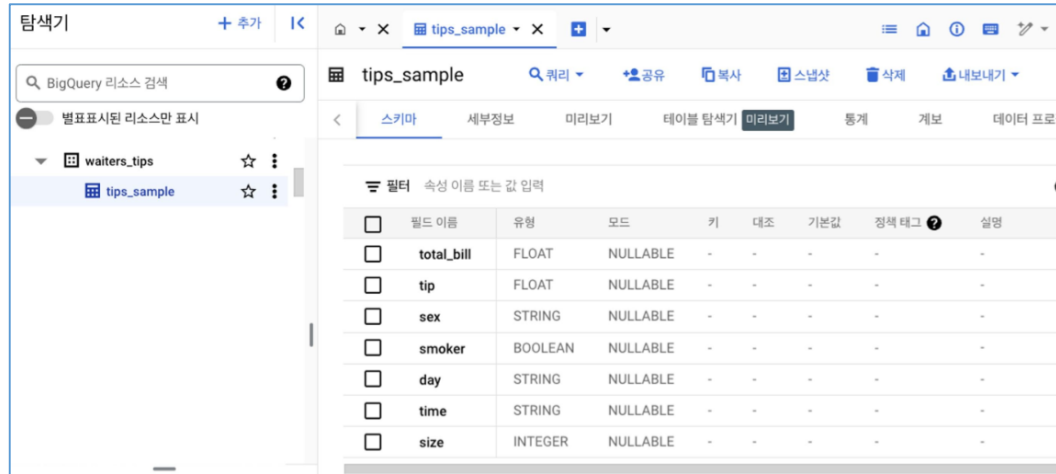
☒ 자동 감지

i

스키마가 자동으로 생성됩니다.

CSV 데이터 BigQuery에 업로드

8. 하단의 '테이블 만들기' 클릭. 데이터 확인하기



구글 빅쿼리에 자료 업로드시 문제점들

<https://blog.naver.com/gangfeel2>

구글빅쿼리 사용법 (SQL과 같음)

<https://zzsza.github.io/bigquery/basic/select.html>

*간단히 구글 스프레드 시트와 빅쿼리를 연결하여 SQL 문 사용

<https://blog.naver.com/bohee9410/223337228801>

<https://blog.naver.com/bohee9410/223337266356>

<https://www.autooffice.co.kr/blog/2024/02/05/%EC%8A%A4%ED%94%84%EB%A0%88%EB%93%9C%EC%8B%9C%ED%8A%B8-%EB%8D%B0%EC%9D%B4%ED%84%B0%EB%A5%BC-%EB%B9%85%EC%BF%BC%EB%A6%AC%EC%99%80-%EB%A3%A8%EC%BB%A4-%EC%8A%A4%ED%8A%9C%EB%94%94%EC%98%A4%EB%A1%9C/>

다양한 예
제 있음

*활용 사례 보기

<https://blog.bizspring.co.kr/%ED%99%9C%EC%9A%A9-%EB%B0%A9%EB%B2%95-%EC%82%AC%EB%A1%80/%EB%84%A4%EC%9D%B4%EB%B2%84-%EA%B4%91%EA%B3%A0-%EB%8D%B0%EC%9D%B4%ED%84%B0-%EA%B5%AC%EA%B8%80-%EB%B9%85%EC%BF%BC%EB%A6%AC-%EC%9E%90%EB%8F%99%EC%9C%BC%EB%A1%9C-%EC%8C%93%EB%8A%94-%EB%B0%A9%EB%B2%95/>

구글 GA 자격증 미취득자는 GA자격증과 관련한 유튜브 영상 보면서 퍼포먼스 마케팅 한번 살펴보세요.
GA와 빅쿼리 연결법임

<https://osoma.kr/blog/ga4-bigquery-connect/>

<https://osoma.kr/blog/why-connect-ga4-bigquery/>

구글 빅쿼리에서 데이터 추출 (CSV로 저장)

GA하신분들은 이 사이트의 내용으로 공부



<https://zzsza.github.io/bigquery/useful/extract-data.html>

연결된 시트(Connected Sheets)_구글 스프레드시트로 저장

<https://zzsza.github.io/bigquery/useful/connected-sheets.html>

아래 내용은 대강 읽기만 하기

<https://zzsza.github.io/bigquery/useful/sampling.html#farm-fingerprint>

<https://zzsza.github.io/bigquery/useful/schedule-query.html>

<https://zzsza.github.io/bigquery/useful/new-feature.html>

코랩과 빅쿼리 연결

<https://velog.io/@dlgkdis801/Node-03.-GCP-%EC%86%8C%EA%B0%9C>

콘솔을 사용한 신용카드 거래 사기 감지를 위한 BigQuery ML

<https://codelabs.developers.google.com/codelabs/frauddetection-using-console?hl=ko#0>

구글 빅쿼리에서 모델 학습하기

<https://suwani.tistory.com/173>

BigQuery ML 초매개변수 조정을 사용하여 모델 성능 향상

<https://www.cloudskillsboost.google/focuses/33355?locale=ko&parent=catalog>

읽기만 합니다. (클라우드 회사에서 제공하는
업무에서 바로 활용 가능한 BigQuery 기능 살펴보기)

https://gcp.cloocus.com/tech-037-bigquery_function/

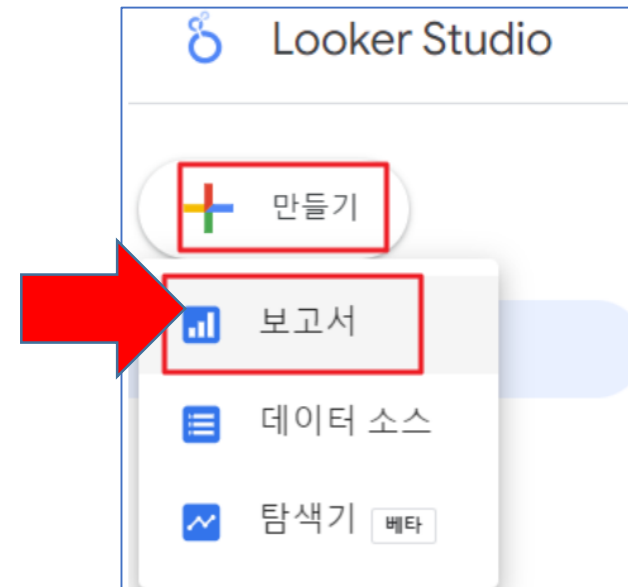
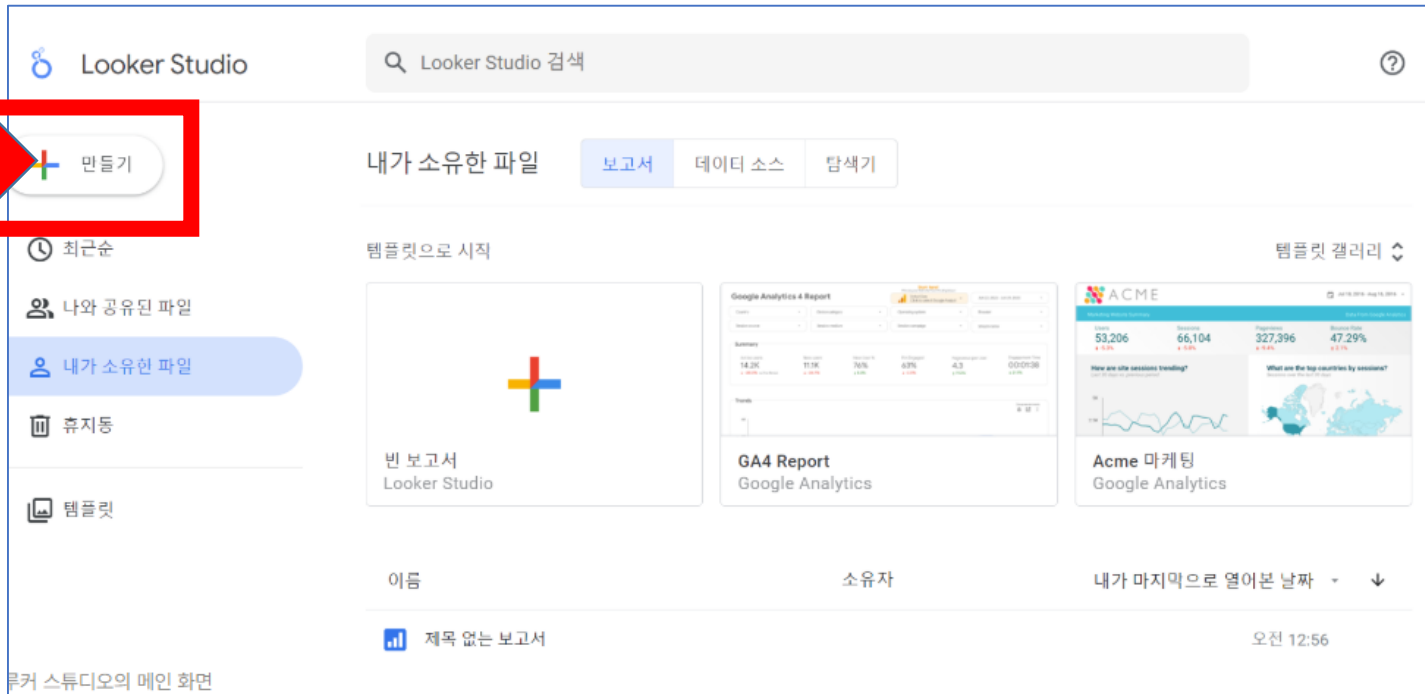
Lookerstudio에서 BigQuery 테이블 가져오기

Looker Studio는 Google의 데이터 시각화 및 보고 도구로, 다양한 데이터 소스를 연결하여 대시보드를 생성할 수 있음. BigQuery와의 연동이 편리한 이유는 실시간 데이터 쿼리 지원, 원활한 Google 생태계 통합, 코드 없이 직관적인 UI로 데이터 분석 가능하기 때문.

1. 루커스튜디오 접속

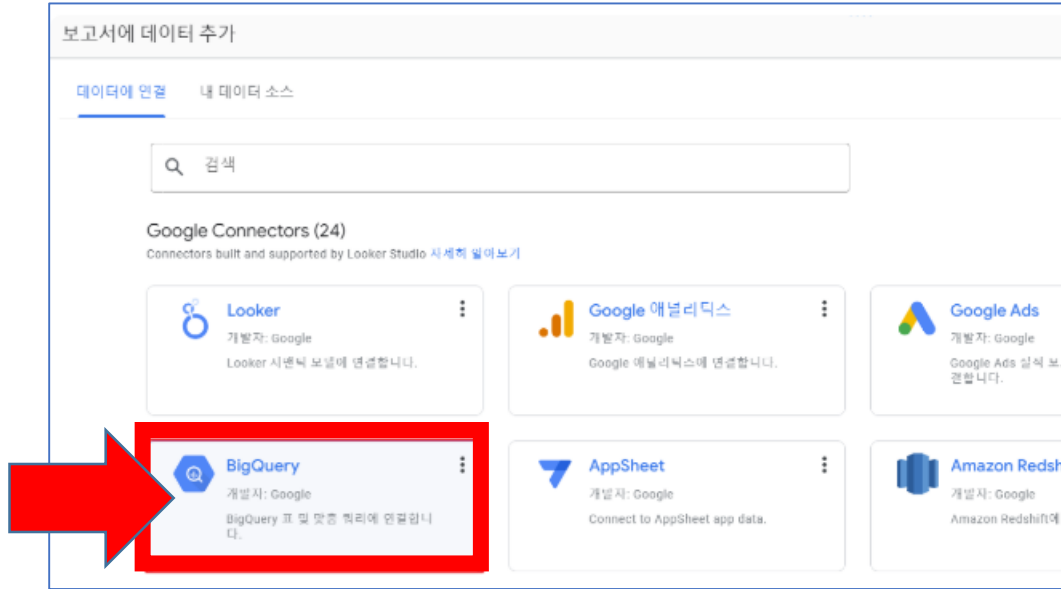
<https://lookerstudio.google.com/u/0/navigation/reporting>

2. '만들기' 버튼 클릭, '보고서' 클릭



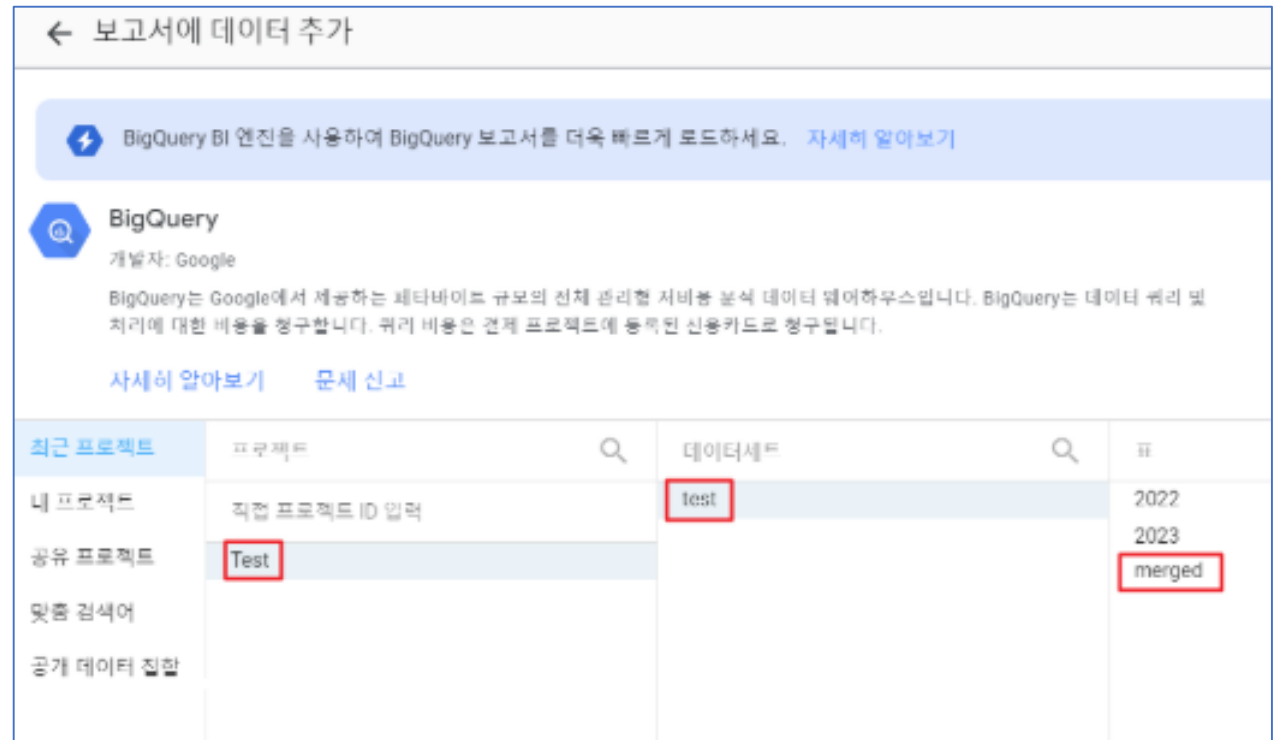
Lookerstudio에서 BigQuery 테이블 가져오기

3. 연결할 데이터의 종류로 **BigQuery(빅쿼리)** 선택



4. 빅쿼리에서 불러오길 원하는 테이블을 선택 후 '추가' 버튼 클릭.

test 데이터세트의 merged 테이블을 선택



Lookerstudio에서 BigQuery 테이블 가져오기

5. 원하는 차트 추가해서 분석

The screenshot displays the Looker Studio interface. At the top, the title is '제목 없는 보고서' (Untitled Report). Below the title bar, there are navigation icons and a toolbar. The main content area shows a table with the following data:

	client	Record Count
1.	D문구점	3
2.	B문구점	2
3.	A문구점	2
4.	C문구점	2

On the right side, the '차트' (Chart) configuration panel is open. It shows the '데이터 소스' (Data Source) as 'salesData'. The '기간 측정기준' (Time Dimension) is set to 'salesDate(날짜)'. The '측정기준' (Metric) is set to 'RBC client'. The '측정항목' (Measure) is set to 'AUT Record Count'. The '드릴다운' (Drilldown) is disabled. The '선택 측정항목' (Selected Measure) is 'Record Count'. The '측정항목 슬라이더' (Measure Slider) is also disabled.