

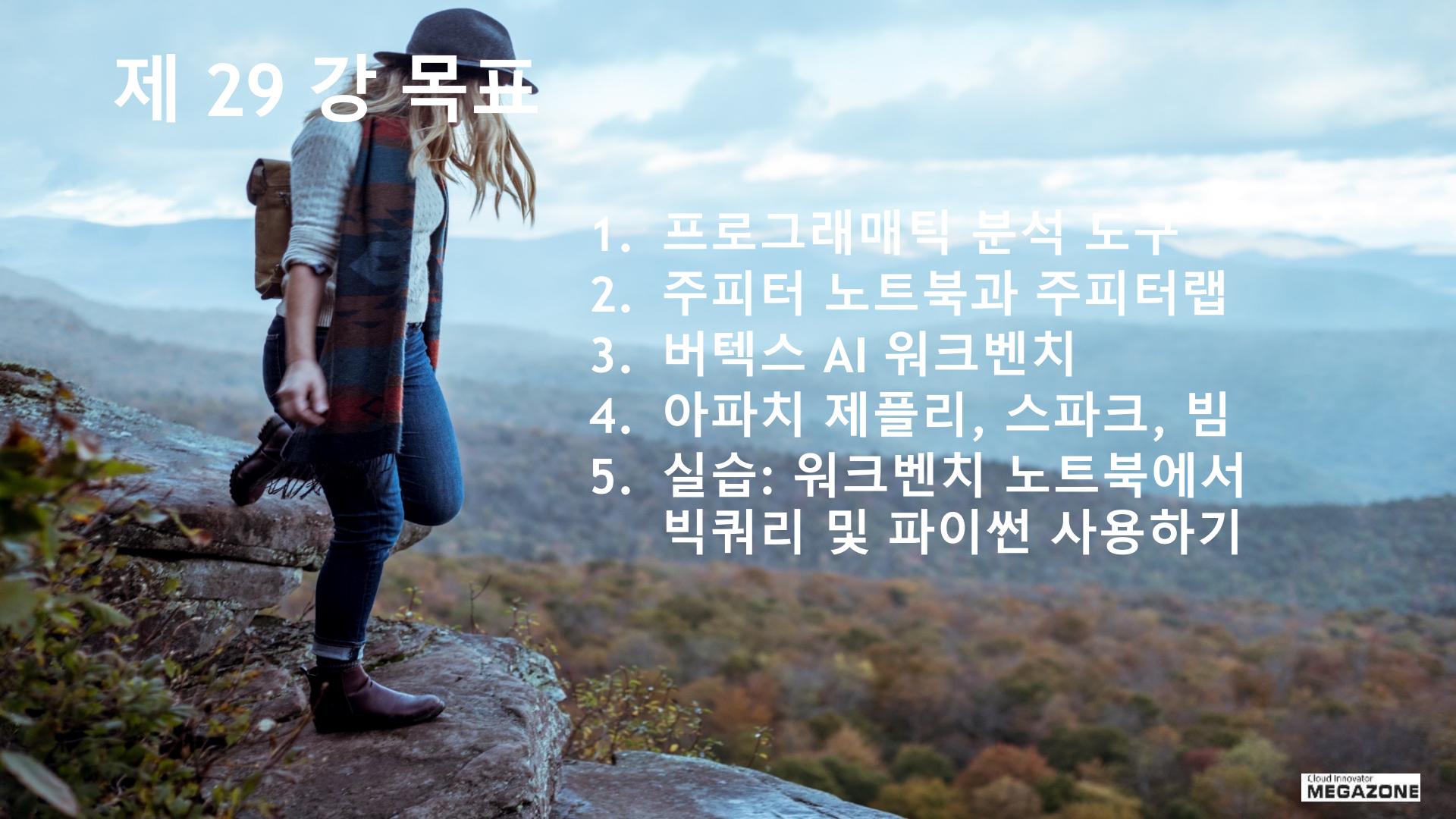


빅쿼리를 활용한 빅데이터 분석 (29)

서진호

자료 다운로드: <https://www.github.com/synabreu/BigQuery>

제 29 강 목표

- 
1. 프로그래매틱 분석 도구
 2. 주피터 노트북과 주피터랩
 3. 버텍스 AI 워크벤치
 4. 아파치 제플리, 스파크, 빔
 5. 실습: 워크벤치 노트북에서
빅쿼리 및 파이썬 사용하기

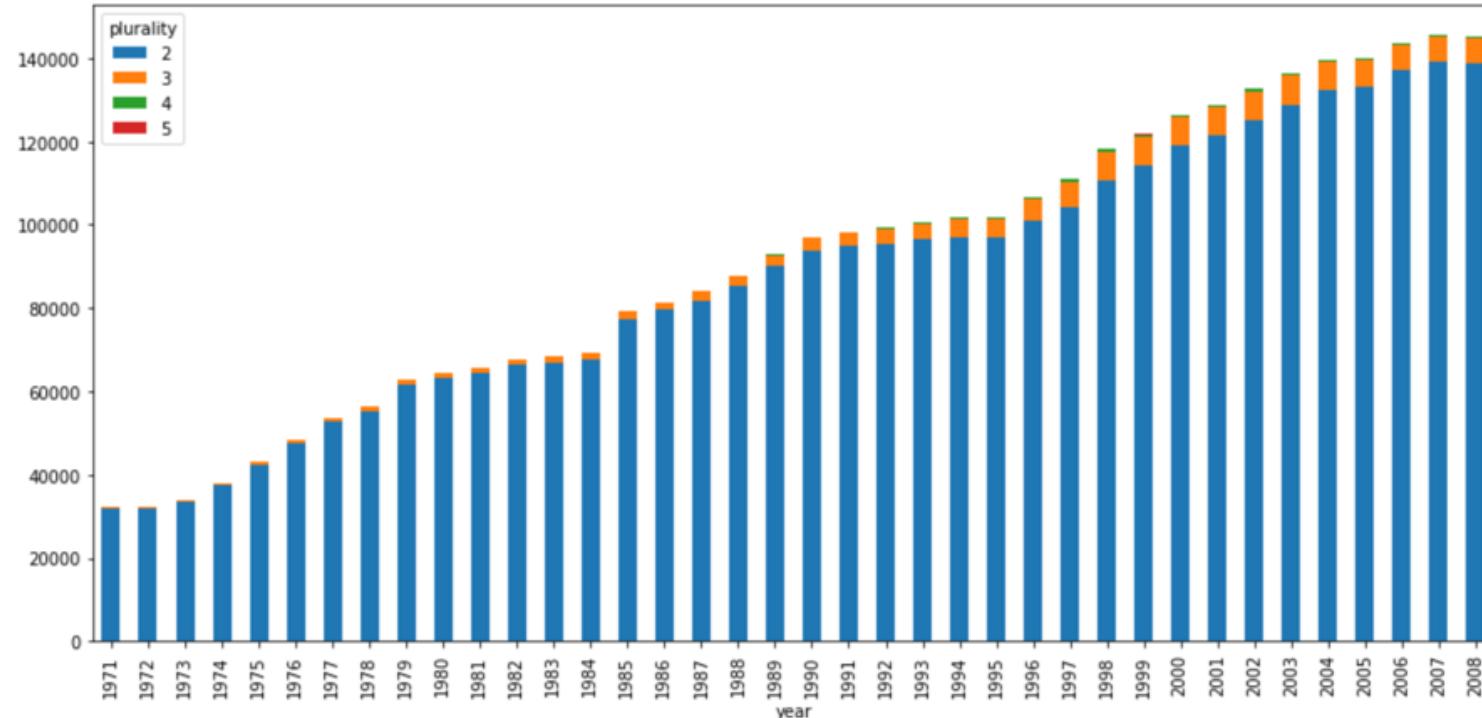
프로그래밍 분석 도구

- 주피터 노트북(Jupyter Notebooks)
- 주피터랩(JupyterLab)
- 버텍스 AI 워크벤치(Vertex AI Workbench)
- 아파치 제플린(Apache Zeppelin)
- 아파치 하둡의 아파치 스파크(Apache Hadoop/Apache Spark)
- 아파치 빔(Apache Beam)



주피터 노트북

```
pivot_table = df.pivot(index='year', columns='plurality', values='count')
pivot_table.plot(kind='bar', stacked=True, figsize=(15, 7));
```



주피터랩

The screenshot shows a Jupyter Notebook interface. On the left, a file browser displays a directory structure under 'tutorials > bigquery'. The current file, 'Visualizing BigQuery public data.ipynb', is selected and highlighted in blue. The main area contains a notebook cell with the following content:

Vizualizing BigQuery data in a Jupyter notebook

BigQuery is a petabyte-scale analytics data warehouse that you can use to run SQL queries over vast amounts of data in near realtime.

Data visualization tools can help you make sense of your BigQuery data and help you analyze the data interactively. You can use visualization tools to help you identify trends, respond to them, and make predictions using your data. In this tutorial, you use the BigQuery Python client library and pandas in a Jupyter notebook to visualize data in the BigQuery natality sample table.

Using Jupyter magics to query BigQuery data

The BigQuery Python client library provides a magic command that allows you to run queries with minimal code.

The BigQuery client library provides a cell magic, `%bigquery`. The `%bigquery` magic runs a SQL query and returns the results as a pandas `DataFrame`. The following cell executes a query of the BigQuery natality public dataset and returns the total births by year.

```
[1]: %bigquery
SELECT
    source_year AS year,
    COUNT(is_male) AS birth_count
FROM `bigquery-public-data.samples.natality`
GROUP BY year
ORDER BY year DESC
LIMIT 15
```

	year	birth_count
0	2008	4255156
1	2007	4324008
2	2006	4273225
3	2005	4145619
4	2004	4118907
5	2003	4096092

버텍스 AI 워크벤치

The screenshot shows a browser window with three tabs open: "Workbench - Vertex AI - My NB" (active), "BigQuery com... - JupyterLab", and "BigQuery - My NB Project - Go". The main content area is a JupyterLab interface titled "managed-notebook-1662472026".

File Explorer: Shows a directory structure under "/ ... / python-docs-samples / bigquery /". The files listed are:

- resources (21 minutes ago)
- BigQuery basics.ipynb (21 minutes ago)
- BigQuery command-line tool.ipynb (a minute ago)
- BigQuery query magic.ipynb (21 minutes ago)
- Getting started with BigQuery M... (21 minutes ago)
- Getting started with BigQuery M... (21 minutes ago)
- Visualizing BigQuery public data... (21 minutes ago)

Launcher: Shows tabs for "Launcher", "Query Editor 2", and "BigQuery command-line tool" (selected).

Query Editor 2: Displays the following text:

Create a new dataset

A dataset is contained within a specific [project](#). Datasets are top-level containers that are used to organize and control access to your [tables](#) and [views](#). A table or view must belong to a dataset. You need to create at least one dataset before [loading data into BigQuery](#).

First, name your new dataset:

```
[3]: dataset_id = "myDataSet"
```

The following command creates a new dataset in the US using the ID defined above.

NOTE: In the examples in this notebook, the `dataset_id` variable is referenced in the commands using both `{}` and `$`. To avoid creating and using variables, replace these interpolated variables with literal values and remove the `{}` and `$` characters.

```
[4]: !bq --location=US mk --dataset $dataset_id
```

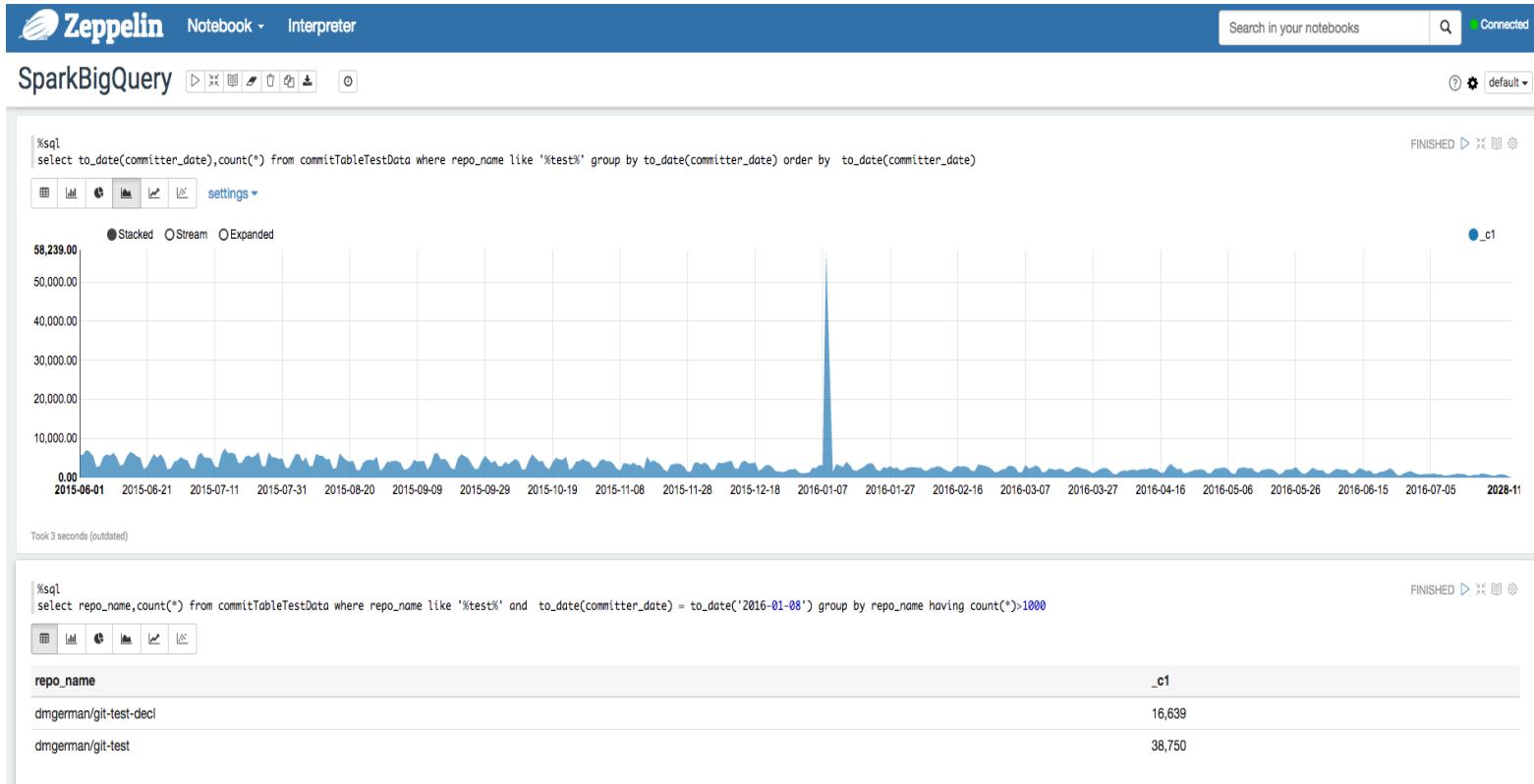
Dataset 'my-nb-project-361711:myDataSet' successfully created.

The response should look like the following:

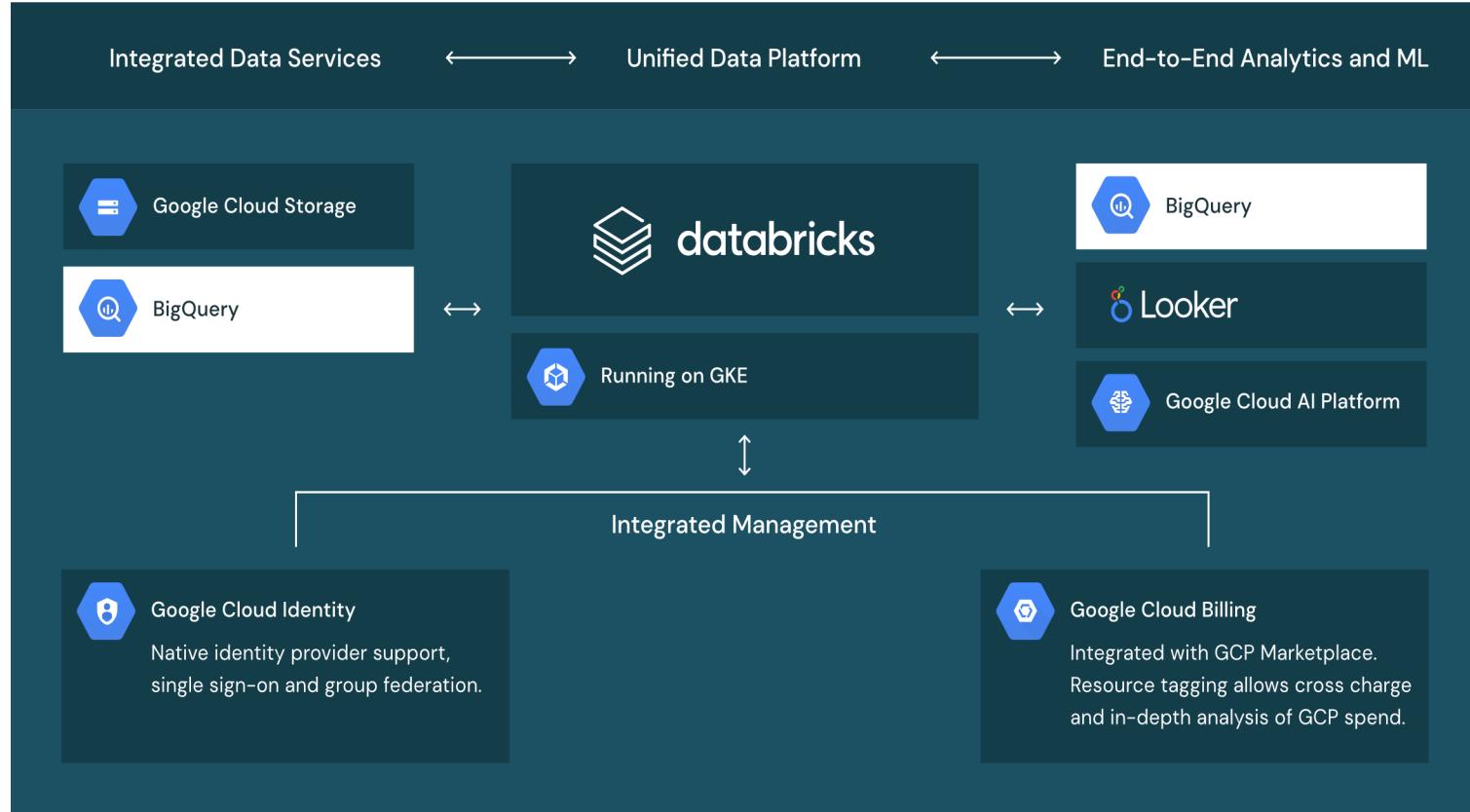
```
[4]: Dataset 'your-project-id:your_new_dataset' successfully created.
```

BigQuery command-line tool.ipynb

아파치 제플린



아파치 하둡의 아파치 스파크

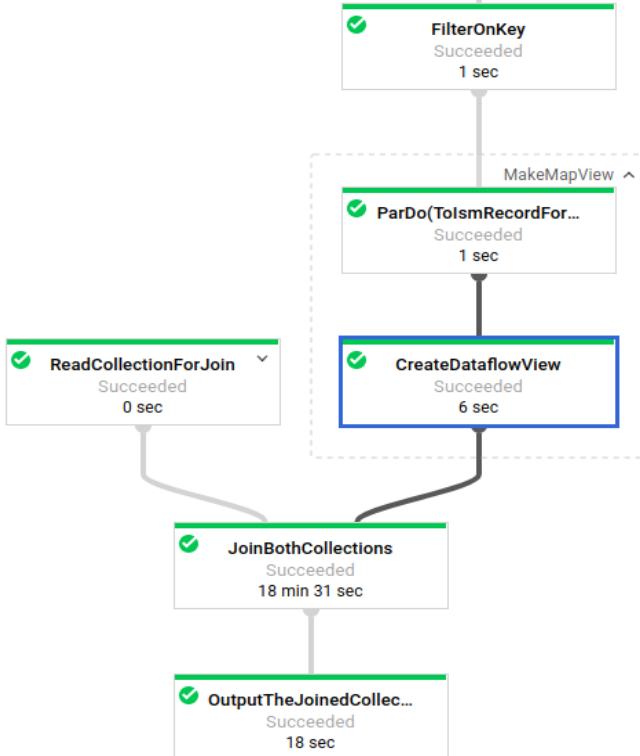


아파치 빙 - 구글 데이터플로

← df2

LOGS

X Step



Step summary

Step name	MakeMapView/CreateDataflowView
Wall time	6 min 3 sec

Input collections

MakeMapView/ParDo(TolsmRecordForGlobalWindow).out0	
Elements added	100,000
Estimated size	8.68 GB

Output collections

MakeMapView/CreateDataflowView.out0	
Elements added	-
Estimated size	-

Side Input Metrics

MakeMapView/CreateDataflowView.out0	
Time spent writing	6 min 3 sec
Bytes written	8.68 GB

Time & bytes read from side input

Side input consumer	Time spent reading	Bytes read
JoinBothCollections	10 min 3 sec	10.14 GB

실습1 – 워크벤치 노트북 사용하기 위한 새 프로젝트 생성



프로젝트 이동

Google Cloud My PC Project

console.cloud.google.com/home/dashboard?authuser=1&project=my-pc-project-357506

My PC Project

대시보드 활동 권장사항

프로젝트 정보

- 프로젝트 이름: My PC Project
- 프로젝트 번호: 985086557691
- 프로젝트 ID: my-pc-project-357506

이 프로젝트에 사용자 추가

프로젝트 설정으로 이동

리소스

- BigQuery
- SQL
- Compute Engine
- Storage
- Cloud 함수

API API

Google Cloud Platform 상태

결제

모니터링

Google Cloud Platform 상태

No data is available for the selected time frame.

Cloud 상태 대시보드로 이동

예상 요금: KRW ₩2,277.92
결제 기간: 2022. 9. 1. ~ 2022. 9. 6.

결제 살펴보기

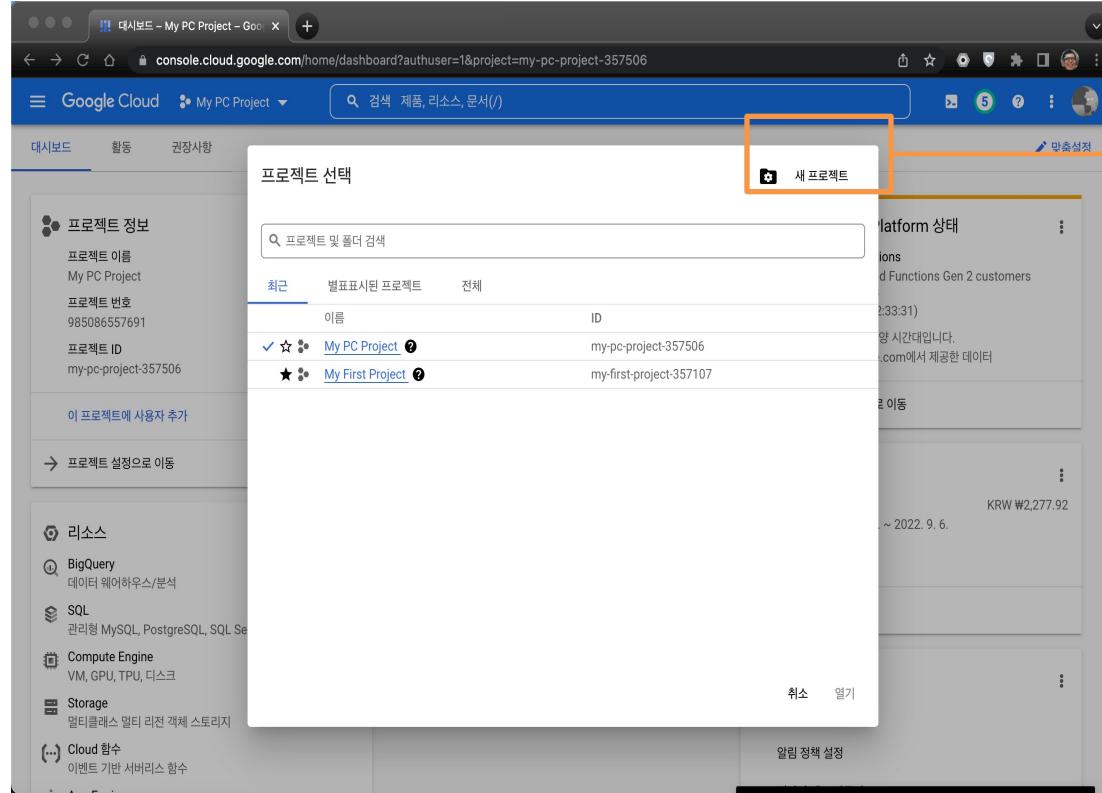
청구 세부정보 보기

내 대시보드 만들기

알림 정책 설정

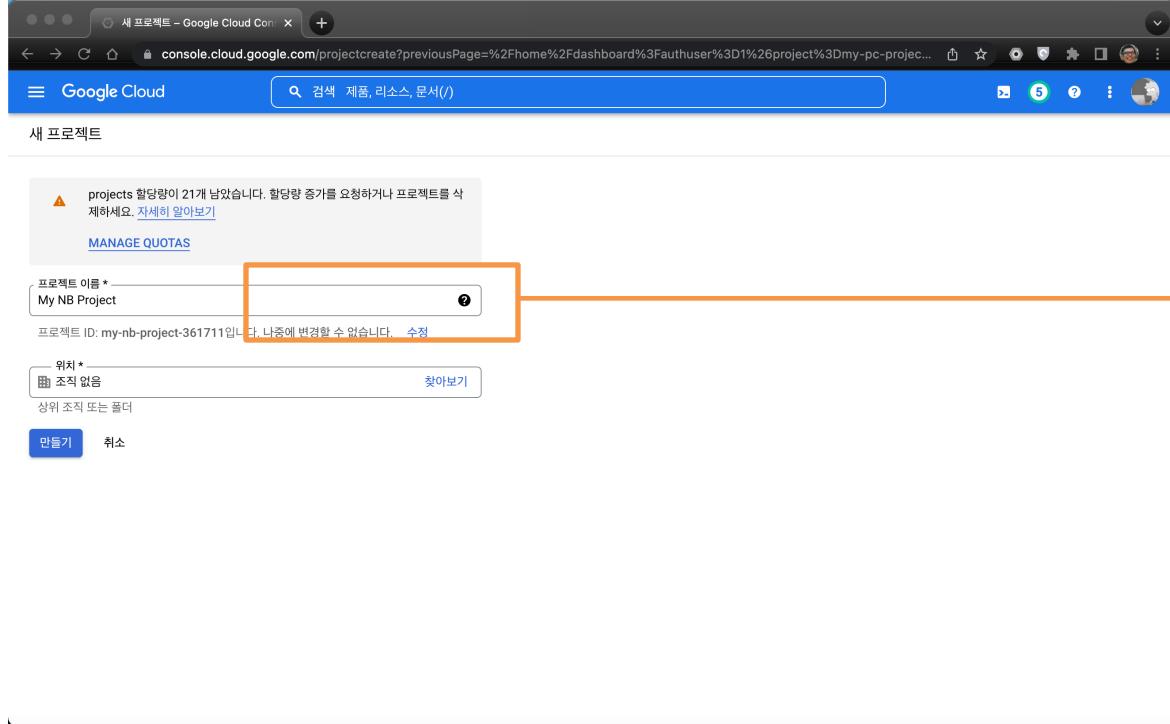
My PC Project 클릭

새 프로젝트 선택



새 프로젝트 클릭

새 프로젝트 입력



- 프로젝트명: My NB Project
- 만들기 클릭

프로젝트 선택

The screenshot shows the Google Cloud Platform dashboard for the project 'My PC Project'. On the left, there's a sidebar with sections for 'Project Information', 'APIs & Services', 'Resources', and 'Cloud Functions'. The main area displays a timeline of recent actions under the heading 'Recent Activity'. One specific action is highlighted with an orange box: '프로젝트 My NB Project 만들기' (Create project My NB Project) with a timestamp of '방금' (Just now). An orange arrow points from this highlighted item to the right, where the text '프로젝트 선택' (Project Selection) is displayed.

프로젝트 My NB Project 만들기
방금

버킷 1개 삭제 중 4시간 전

객체 1개 및 폴더 0개 삭제 중 4시간 전

버킷 1개 삭제 중 1일 전

객체 1개 및 폴더 0개 삭제 중 1일 전

객체 2개 및 폴더 0개 삭제 중 2일 전

버킷 1개 삭제 중 2일 전

객체 1개 및 폴더 0개 삭제 중 2일 전

객체 3개 및 폴더 0개 삭제 중 2일 전

서비스 사용 설정:
bigqueryconnection.googleapis.com
My PC Project 2일 전

서비스 사용 설정:
bigqueryreservation.googleapis.com
My DS Project 4일 전

프로젝트 선택

BigQuery API 검색

A screenshot of a web browser showing search results for 'BigQuery API' on the Google Cloud search interface. The search bar at the top contains the query '검색 BigQuery API'. The results page has a sidebar on the left with filters for products, documents, Marketplace, organizations, folders, projects, and locations. The main content area shows search results for 'BigQuery API'.

검색 결과

'BigQuery API'의 검색 결과 136개 중 30개가 표시됩니다.

API BigQuery API
A data platform for customers to create, manage, share and query data.
유형: API
제작자: Google Enterprise API

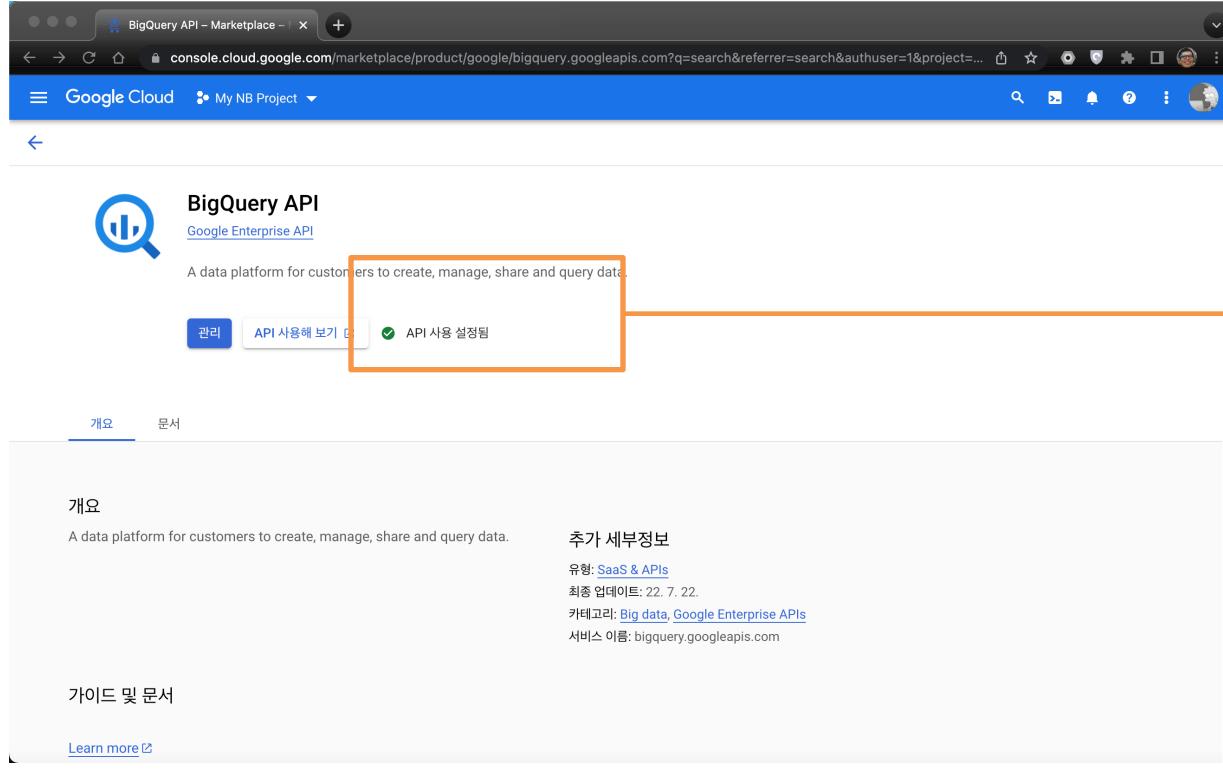
API Analytics Hub API
Exchange data and analytics assets securely and efficiently.
유형: API
제작자: Google

API BigQuery Connection API
Allows users to manage BigQuery connections to external data sources.
유형: API
제작자: Google Enterprise API

BigQuery API 클라이언트 라이브러리 - Google Cloud
이 페이지에서는 BigQuery API용 Cloud 클라이언트 라이브러리를 시작하는 방법을 설명합니다. 이전 Google API 클라이언트 라이브러리를 포함한 Cloud APIs용 ...
유형: 문서

BigQuery API 클릭

BigQuery API 사용 확인



BigQuery API – Marketplace

Google Cloud My NB Project

BigQuery API

Google Enterprise API

A data platform for customers to create, manage, share and query data.

관리 API 사용해 보기 API 사용 설정됨

개요 문서

개요

A data platform for customers to create, manage, share and query data.

추가 세부정보

유형: SaaS & APIs
최종 업데이트: 22. 7. 22.
카테고리: Big data, Google Enterprise APIs
서비스 이름: bigquery.googleapis.com

가이드 및 문서

[Learn more](#)

API 사용 설정됨

Notebooks API 사용 검색

The screenshot shows a Google Cloud search interface. The search bar at the top contains the query '검색 Notebooks API'. Below the search bar, the results are displayed under the heading '검색 결과' (Search results). The first result is 'Notebooks API', which is highlighted with an orange box and an arrow pointing to it from the right side of the slide. This result includes a brief description: 'Notebooks API is used to manage notebook resources in Google Cloud.' Below this, there are three more items: 'Databricks', 'Incorta Direct Data Platform', and a document titled '사용자 관리형 노트북 인스턴스의 환경 업그레이드'.

검색 결과

'Notebooks API'의 검색결과 14개 중 14개가 표시됩니다.

Notebooks API

Notebooks API is used to manage notebook resources in Google Cloud.

유형: API
제작자: Google Enterprise API

Databricks

All your data, analytics and AI on one Lakehouse platform

유형: Marketplace 제품
제작자: Databricks

Incorta Direct Data Platform

A unified platform for data ingest, inquiry, and insight.

유형: Marketplace 제품
제작자: Incorta, Inc.

사용자 관리형 노트북 인스턴스의 환경 업그레이드

Notebooks API가 인스턴스의 Google Cloud 프로젝트에서 사용 설정되어 있어야 합니다. ...
사용자 관리형 노트북 인스턴스는 Notebooks API를 사용하여 생성되었거나 ...

유형: 문서

Notebooks API
클릭

Notebooks API 사용

The screenshot shows a web browser window for the Google Cloud Marketplace. The URL is console.cloud.google.com/marketplace/product/google/notebooks.googleapis.com?q=search&referrer=search&authuser=1&project=My%20NB%20Project. The page displays the 'Notebooks API' product from 'Google Enterprise API'. A large orange arrow points from the '사용 클릭' (Click to use) button to the button itself, which is highlighted with an orange border.

Notebooks API
Google Enterprise API

Notebooks API is used to manage notebook resources in Google Cloud.

사용 클릭

개요

Notebooks API is used to manage notebook resources in Google Cloud.

추가 세부정보

유형: [SaaS & APIs](#)
최종 업데이트: 22. 7. 22.
카테고리: [Google Enterprise APIs](#)
서비스 이름: notebooks.googleapis.com

가이드 및 문서

[Learn more ↗](#)

Notebooks API 사용 확인

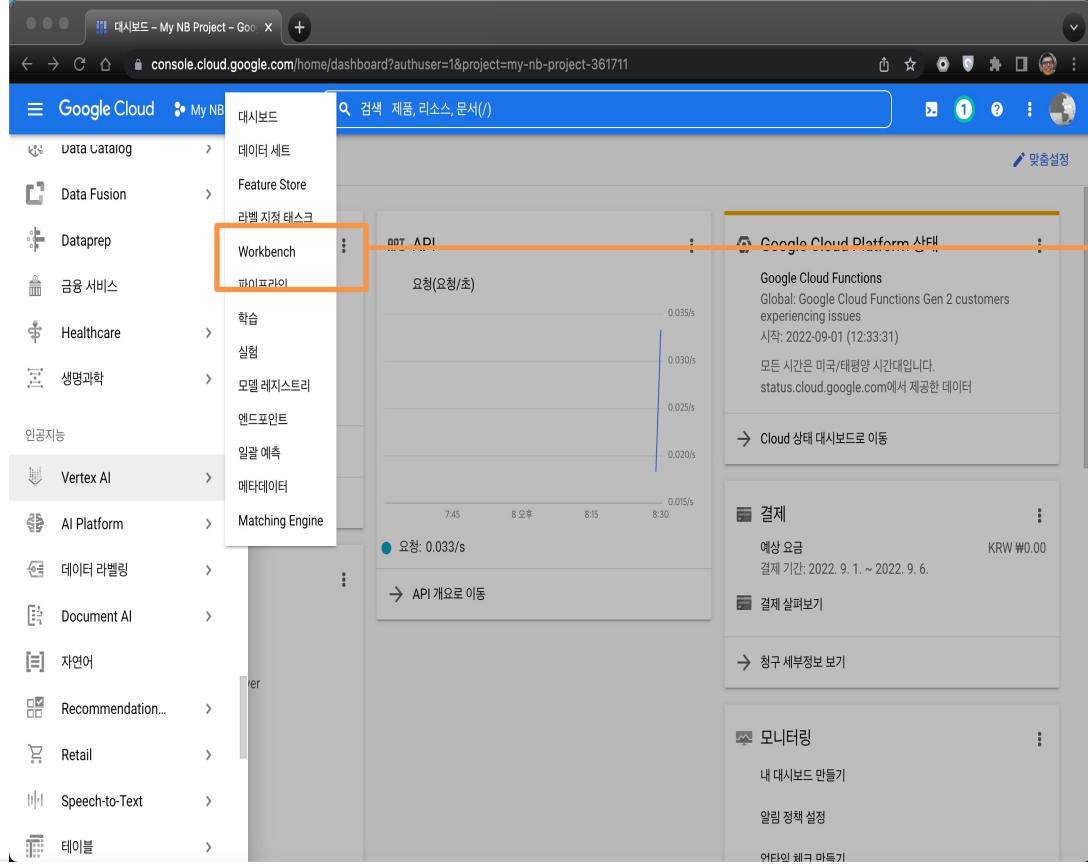
The screenshot shows the Google Cloud Platform API Services library interface. On the left, a sidebar lists various services: API 및 서비스 (selected), 사용 설정된 API 및 서비스 (Notebooks API selected), 라이브러리, 사용자 인증 정보, OAuth 동의 화면, 도메인 확인, and 페이지 사용 등의. The main content area displays the 'API/서비스 세부정보' (Notebooks API) page. A large orange arrow points from the text 'API 사용 설정 확인' to the 'Notebooks API' section. The 'Notebooks API' section includes a brief description: 'Notebooks API is used to manage notebook resources in Google Cloud.', a provider note: '제공: Google Enterprise API', and details: '서비스 이름: notebooks.googleapis.com', '유형: 공개 API', and '상태: 사용 설정됨'. Below this, there are tabs for '측정 항목' (Metrics), '할당량' (Quotas), '사용자 인증 정보' (Authentication), and '비용' (Cost). A search bar at the bottom left says '응답 코드별 트래픽' (Traffic by response code) and a note at the bottom right says '▲ 선택한 기간에는 데이터가 없습니다.' (No data for the selected period).

API 사용
설정 확인

실습2 – 워크벤치 노트북에서 빅쿼리 및 파이썬 사용하기



Vertex AI 의 Workbench 실행



Workbench 클릭

새 노트북 실행

The screenshot shows the Google Cloud Vertex AI Workbench interface. On the left, there's a sidebar with various options like Vertex AI, Workbench, and Marketplace. The main area has a search bar at the top. Below it, there's a toolbar with buttons for 'New Notebook' (highlighted with an orange box), 'Revert', 'Start', 'Stop', 'Delete', 'Upgrade', and 'Schedule'. A message in Korean says 'Notebooks에는 JupyterLab 3이 사전 설치되어 있으며 GPU 지원 머신러닝 프레임워크로 구성되어 있습니다.' followed by a link '자세히 알아보기'. There's also a filter section and a large brain icon. At the bottom, there's a message 'This project does not have any notebooks.' followed by 'Create a notebook instance' and 'Quick start' buttons.

새 노트북 클릭

관리형 노트북 만들기

관리형 노트북 만들기 - Vertex AI · BigQuery - My NB Project - Google Cloud · console.cloud.google.com

Google Cloud My NB Project 검색 제품, 리소스, 문서(/)

Vertex AI 관리형 노트북 만들기 알아보기

노트북 이름 * managed-notebook-1662472026
글자 수 제한은 63자(영문 기준)입니다. 소문자, 숫자, '-'만 사용할 수 있으며 문자로 시작해야 하며 '.'로 끝날 수 없습니다.

리전 us-central1 (Iowa)

월 US\$202.09 예상
시간당 약 \$0.277
사용한 만큼만 비용 지불: 선불 비용 없이 초당 청구
네트워킹 비용도 적용됩니다. 자세히 알아보기

권한

JupyterLab 액세스 모드는 노트북 인스턴스를 사용할 수 있는 사용자 및 Google API를 호출하는 데 사용되는 사용자 인증 정보를 결정합니다. 노트북을 만든 후에는 모드를 변경할 수 없습니다.

Service account
서비스 계정 모드는 사용자는 자신의 서비스 계정에 `serviceAccounts.actAs` 권한을 부여받은 모든 사용자가 JupyterLab UI에 액세스할 수 있습니다. 자세히 알아보기

Single user only
단일 사용자 전용 모드는 아래에 지정된 사용자만 액세스하도록 제한합니다. 자세히 알아보기

Compute Engine 기본 서비스 계정 사용
Compute Engine 기본 서비스 계정을 변경했다면 충분한 API 권한을 보유하고 있는지 확인하시기 바랍니다. 자세히 알아보기

고급 설정

만들기 취소

Service account 선택

고급 설정

관리형 노트북 만들기 – Vertex AI – BigQuery – My NB Project – console.cloud.google.com/vertex-ai/workbench/create-managed?authuser=0&project=my-nb-project-361711

Google Cloud My NB Project 검색 제품, 리소스, 문서(/)

Vertex AI 관리형 노트북 만들기 알아보기

대시보드 고급 설정 환경 월 US\$202.09 예상

데이터 세트 시간당 약 \$0.277

Feature Store 사용한 만큼만 비용 지불: 선불 비용 없이 초당 청구

라벨 지정 태스크 네트워킹 비용도 적용됩니다. 자세히 알아보기

Workbench

커스텀 Docker 이미지 자체 커스텀 Docker 이미지를 제공하여 추가 커스텀 Jupyter 커널을 만들고 액세스합니다. 컨테이너의 사용 가능한 모든 Jupyter 커널을 가져옵니다.

커스텀 Docker 이미지 제공

하드웨어 구성

머신 유형 * n1-standard-4 (4 vCPUs, 15 GB RAM)

GPU 유형 None

위에서 선택한 영역, 환경, 머신 유형에 따라 선택할 수 있는 GPU 유형 및 최소 GPU 수가 달라질 수 있습니다. 자세히 알아보기

Data 디스크 유형 Standard Persistent Disk Data 디스크 크기(GB) * 100

휴지통에 삭제하시겠어요? ?

디스크 암호화

Google 관리 암호화 키 구성이 필요하지 않음

고객 관리 암호화 키(CMEK) Google Cloud Key Management Service를 통해 관리

Marketplace

아무것도 선택하지
말고 그대로 보고
아래로 커서를 진행

만들기 선택

관리형 노트북 만들기 – Vertex AI – BigQuery – My NB Project – G – +

console.cloud.google.com/vertex-ai/workbench/create-managed?authuser=0&project=my-nb-project-361711

Google Cloud My NB Project 검색 제품, 리소스, 문서(/)

Vertex AI 관리형 노트북 만들기 알아보기

대시보드 Google 관리 암호화 키 구성이 필요하지 않음 월 US\$202.09 예상

데이터 세트 고객 관리 암호화 키(CMEK) Google Cloud Key Management Service를 통해 관리 시간당 약 \$0.277

Feature Store 유축 상태 종료 사용한 만큼만 비용 지불: 선불 비용 없이 초당 청구

라벨 지정 태스크 유축 상태 종료 사용 설정 네트워킹 비용도 적용됩니다. 자세히 알아보기

Workbench 종료 전 비활성 시간(분)* 180 세부정보

파이프라인 범위: 10분~1일(1,440분)

학습 아래 네트워킹 옵션 자세히 알아보기

실험 Google 관리 네트워크 구성이 필요하지 않습니다.

모델 레지스트리 이 프로젝트의 네트워크 비공개 서비스 액세스가 필요합니다.

엔드포인트 나와 공유된 네트워크 비공개 서비스 액세스가 필요합니다.

일괄 예측 보안

Matching Engine nbconvert 사용 설정 노트북 UI에서 파일 다운로드 사용 설정 터미널 사용 설정

Marketplace 만들기 취소

만들기 클릭

JUPYTERLAB 열기

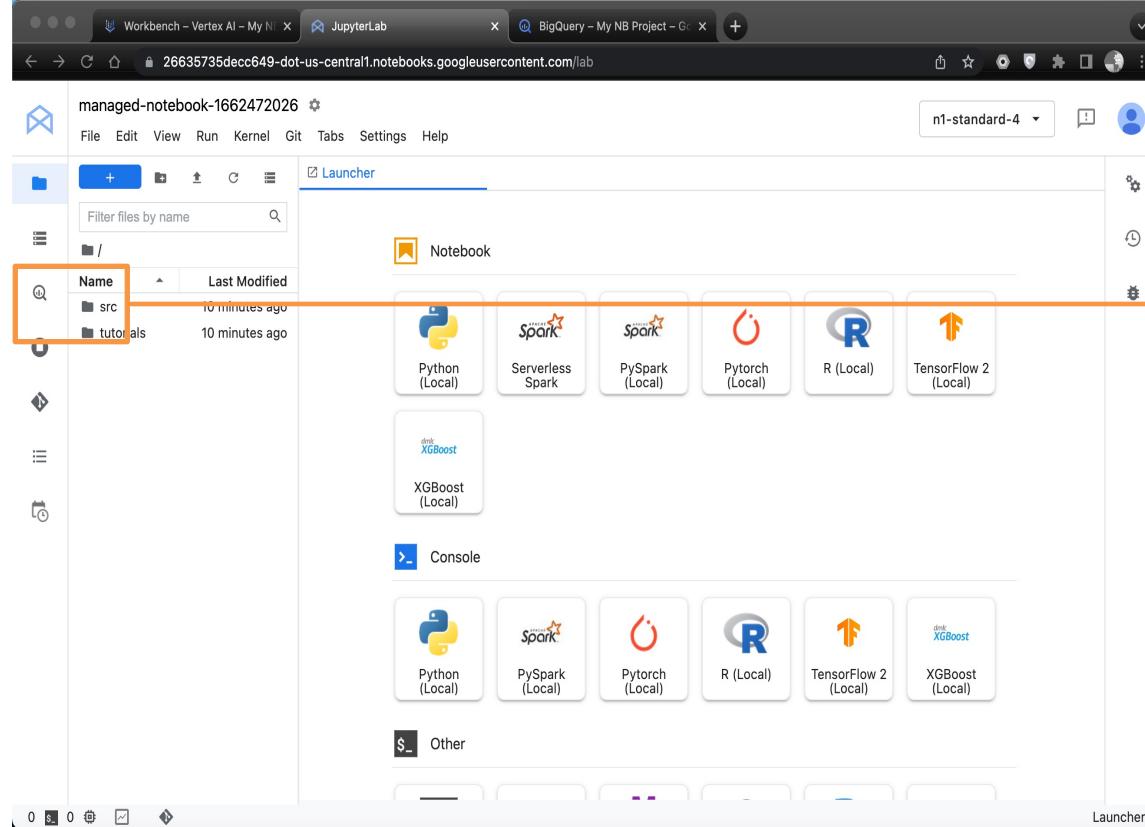
The screenshot shows the Google Cloud Platform Workbench interface. On the left, there is a sidebar with various service icons: Vertex AI, 대시보드 (Dashboard), 데이터 세트 (Data Sets), Feature Store, 라벨 지정 테스크 (Labeling Task), Workbench (selected), 파이프라인 (Pipeline), 학습 (Learning), 실험 (Experiment), 모델 레지스트리 (Model Registry), 엔드포인트 (Endpoint), 일괄 예측 (Batch Prediction), 메타데이터 (Metadata), Matching Engine, Marketplace, and Cloud Functions.

The main area displays a list of managed notebooks. A specific row is highlighted with an orange border:

노트북 이름 ↑	소유자	기준 Docker 이미지	최종 수정 일자	
<input type="checkbox"/> ● managed-notebook-1662472026	JUPYTERLAB 열기	Service account	No custom images	2022. 9. 6. 오후 10:50:09

JUPYTERLAB
열기 클릭

BigQuery 실행



BigQuery 아이콘 클릭

공개 데이터셋에서 테이블 쿼리

The screenshot shows the Google Cloud Vertex AI Workbench JupyterLab interface. The left sidebar displays a list of public datasets from BigQuery, including 'austin_311', 'austin_bikeshare', 'austin_crime', 'austin_incidents', 'austin_waste', 'baseball', 'bitcoin_blockchain', 'blackhole_database', 'bikeshare_stations', 'bikeshare_trips', 'bls', 'bls_qcwe', 'breathe', 'broadstreet_adb', 'catalanian_mobile_coverage', 'census_bureau_acs', and 'census_bureau_construct'. A context menu is open over the 'bikeshare_trips' table, with the 'Query table' option highlighted and a red box drawn around it. An orange arrow points from this highlighted option to the text 'Query Table 클릭' located on the right side of the slide.

Query Table
클릭

테이블 조회 쿼리 서밋

The screenshot shows the Google Cloud BigQuery JupyterLab interface. On the left, there's a sidebar with a tree view of datasets and tables, including 'bigquery-public-data' and its sub-tables like 'austin_311', 'austin_bikeshare', 'baseball', etc. The main area has two tabs: 'Launcher' and 'Query Editor 2'. The 'Query Editor 2' tab is active, displaying a query: `SELECT * FROM `bigquery-public-data.austin_bikeshare.bikeshare_stations` LIMIT 1000;`. A red arrow points to the 'Submit Query' button, which is highlighted with a blue border. Above the button, a tooltip says 'This query will process 12.0 KB when run.' The top navigation bar shows tabs for 'Workbench - Vertex AI - My NB Project' and 'JupyterLab'.

Submit Query 클릭

테이블 쿼리 결과

The screenshot shows a JupyterLab interface with a BigQuery notebook open. The left sidebar displays a tree view of datasets and tables, including 'bigquery-public-data.austin_bikeshare.bikeshare_stations'. The main area shows a 'Query Editor 2' tab with the following SQL query:

```
1 SELECT * FROM `bigquery-public-data.austin_bikeshare.bikeshare_stations` LIMIT 1000;
```

The results table has the following columns: Row, station_id, name, status, and address. The first two rows are highlighted with an orange box.

Row	station_id	name	status	address
1	2576	Rainey @ River St	closed	64 Rainey St
2	2712	Toomey Rd @ South Lamar	closed	1301 Toomey Road
3	3381	East 7th & Pleasant Valley	closed	2772 E 7th
4	3464	Pease Park	closed	1155 Kingsbury St
5	1002	6th & Navasota St.	closed	1308 W. 6th St.
6	1003	8th & Guadalupe	closed	800 Guadalupe St.
7	1004	Red River & LBJ Library	closed	2322 Red River Street
8	1005	State Parking Garage @ Brazos & 18th	closed	1789 Brazos St.

Rows per page: 100 | 1-100 of 102 | < > >>

Query results 확인

예제 소스 코드 폴더로 이동

The screenshot shows a JupyterLab interface with the following components:

- File Browser:** On the left, there's a sidebar titled "managed-notebook-1662472026" containing a file tree. A folder icon is highlighted with an orange box. The tree shows a root folder with "src" and "tutorials" subfolders, both updated 17 minutes ago.
- Query Editor:** The main area has tabs for "Launcher" and "Query Editor 2". The "Query Editor 2" tab is active, showing a query code block:

```
SELECT * FROM `bigquery-public-data.austin_bikeshare.bikeshare_stations` LIMIT 1000;
```
- Results:** Below the query editor, the results are displayed in a table format. The columns are "Row", "station_id", "name", "status", and "address". The first 8 rows of data are shown:| Row | station_id | name | status | address |
| --- | --- | --- | --- | --- |
| 1 | 2576 | Rainey @ River St | closed | 64 Rainey St |
| 2 | 2712 | Toomey Rd @ South Lamar | closed | 1301 Toomey Road |
| 3 | 3381 | East 7th & Pleasant Valley | closed | 2772 E 7th |
| 4 | 3464 | Pease Park | closed | 1155 Kingsbury St |
| 5 | 1002 | 6th & Navasota St. | closed | 1308 W. 6th St. |
| 6 | 1003 | 8th & Guadalupe | closed | 800 Guadalupe St. |
| 7 | 1004 | Red River & LBJ Library | closed | 2322 Red River Street |
| 8 | 1005 | State Parking Garage @ Brazos & 18th | closed | 1789 Brazos St. |

Folder 클릭 후
리소스 예제
코드로 tutorials
클릭

BigQuery command-line tool 파일 선택

The screenshot shows a JupyterLab environment. On the left is a file browser for a notebook named 'managed-notebook-1662472026'. The browser lists several files, with three specific files highlighted by a red box: 'bigQuery basics.ipynb', 'bigQuery command-line tool.ipynb', and 'bigQuery query magic.ipynb'. An orange arrow points from these highlighted files to the text 'BigQuery command-line tool.ipynb 클릭' on the right. The central part of the interface is a 'Query Editor 2' window. It contains a query: 'SELECT * FROM `bigquery-public-data.austin_bikeshare.bikeshare_stations` LIMIT 1000;'. Below the query, the results are displayed in a table:

Row	station_id	name	status	address
1	2576	Rainey @ River St	closed	64 Rainey
2	2712	Toomey Rd @ South Lamar	closed	1301 Too
3	3381	East 7th & Pleasant Valley	closed	2772 E 7t
4	3464	Pease Park	closed	1155 King
5	1002	6th & Navasota St.	closed	1308 W. 6t
6	1003	8th & Guadalupe	closed	800 Guadalupe
7	1004	Red River & LBJ Library	closed	2322 Red Rive

At the bottom of the Query Editor, there are navigation controls: 'Rows per page: 100', '1-100 of 102', and arrows for navigating through the results.

BigQuery command-line
tool.ipynb 클릭

Select Kernel

The screenshot shows a Jupyter Notebook interface within a browser window. The title bar indicates the URL is 26635735decc649-dot-us-central1.notebooks.googleusercontent.com/lab/tree/tutorials/python-docs-samples/bigquery/BigQuery... . The notebook sidebar lists several files, including 'BigQuery basics.ipynb', 'BigQuery command-line tool.ipynb', and 'Getting started with BigQuery M...'. The main area displays the 'BigQuery command-line tool' documentation. A modal dialog box titled 'Select Kernel' is open, showing the dropdown menu with 'Python (Local)' selected. An orange arrow points from the text 'Python (Local)' in the dialog to the 'Select' button, which is highlighted with a blue border.

managed-notebook-1662472026

File Edit View Run Kernel Git Tabs Settings Help

n1-standard-4

Launcher Query Editor 2 BigQuery command-line tool

No Kernel

BigQuery command-line tool

The BigQuery command-line tool is installed as part of the Cloud SDK and can be used to interact with BigQuery. When you use CLI commands in a notebook, the command must be prepended with a ! .

Select Kernel

Select kernel for: "BigQuery command-line tool.ipynb"

Python (Local)

No Kernel Select

Create a new dataset

A dataset is contained within a specific project. Datasets are top-level containers that are used to organize and control access to your tables and views. A table or view must belong to a dataset. You need to create at least one dataset before loading data into BigQuery.

First, name your new dataset:

```
[ ]: dataset_id = "your_new_dataset"
```

The following command creates a new dataset in the US using the ID defined above

Mode: Command Ln 1, Col 1 BigQuery command-line tool.ipynb

Python (Local)
선택 후 Select
클릭

BigQuery 명령 실행

The screenshot shows a Jupyter Notebook interface with the title "managed-notebook-1662472026". The left sidebar displays a file tree with several notebooks listed, including "BigQuery basics.ipynb", "BigQuery command-line tool.ipynb" (which is currently selected), and "Getting started with BigQuery M...". The main content area is titled "Create a new dataset". It contains instructions about datasets being top-level containers for tables and views. A code cell at the bottom of the section contains the command `[3]: dataset_id = "myDataSet"`. An orange box highlights this line, and an orange arrow points from it to the explanatory text on the right.

A dataset is contained within a specific project. Datasets are top-level containers that are used to organize and control access to your tables and views. A table or view must belong to a dataset. You need to create at least one dataset before loading data into BigQuery.

First, name your new dataset:

```
[3]: dataset_id = "myDataSet"
```

The following command creates a new dataset in the GCS using the ID defined above.

NOTE: In the examples in this notebook, the `dataset_id` variable is referenced in the commands using both `{}` and `$`. To avoid creating and using variables, replace these interpolated variables with literal values and remove the `{}` and `$` characters.

```
[4]: !bq --location=US mk --dataset $dataset_id
```

Dataset 'my-nb-project-361711:myDataSet' successfully created.

The response should look like the following:

```
Dataset 'your-project-id:your_new_dataset' successfully created.
```

노트북에서
BigQuery 명령
실행

Select Kernel

The screenshot shows a Jupyter Notebook interface within a browser window. The title bar indicates the URL is 26635735decc649-dot-us-central1.notebooks.googleusercontent.com/lab/tree/tutorials/python-docs-samples/bigquery/BigQuery... . The notebook sidebar lists several files, including 'BigQuery basics.ipynb', 'BigQuery command-line tool.ipynb', and 'Getting started with BigQuery M...'. The main area displays the 'BigQuery command-line tool' documentation. A modal dialog box titled 'Select Kernel' is open, showing the dropdown menu with 'Python (Local)' selected. An orange arrow points from the text 'Python (Local)' in the dialog to the right.

managed-notebook-1662472026

File Edit View Run Kernel Git Tabs Settings Help

n1-standard-4

Launcher Query Editor 2 BigQuery command-line tool

No Kernel

BigQuery command-line tool

The BigQuery command-line tool is installed as part of the Cloud SDK and can be used to interact with BigQuery. When you use CLI commands in a notebook, the command must be prepended with a ! .

Select Kernel

Select kernel for: "BigQuery command-line tool.ipynb"

Python (Local)

No Kernel Select

Create a new dataset

A dataset is contained within a specific project. Datasets are top-level containers that are used to organize and control access to your tables and views. A table or view must belong to a dataset. You need to create at least one dataset before loading data into BigQuery.

First, name your new dataset:

```
[ ]: dataset_id = "your_new_dataset"
```

The following command creates a new dataset in the US using the ID defined above

Mode: Command Ln 1, Col 1 BigQuery command-line tool.ipynb

Python (Local)
선택 후 Select
클릭

실습3 – 전체 삭제



노트북 파일 셧다운

The screenshot shows a Jupyter Notebook interface within a browser window. The left sidebar lists notebooks: 'managed-notebook.ipynb', 'BigQuery basics.ipynb', 'BigQuery command-line tool.ipynb', 'Getting started.ipynb', and 'Visualizing BigQ.ipynb'. The main area displays the 'BigQuery command-line tool' notebook. A context menu is open over the first notebook, with the 'Shut Down Kernel' option highlighted by an orange box. An arrow points from this highlighted option to the right text.

managed-notebook.ipynb

File Edit View

+ Open
Open With
+ Open in New Browser Tab

Rename

Git

Delete

Cut

Copy

Duplicate

Download

Shut Down Kernel

New Folder

New File

New Markdown File

Paste

Shift+Right Click for Browser Menu

n1-standard-4

command-line tool

BigQuery command-line tool

The BigQuery command-line tool is installed as part of the Cloud SDK and can be used to interact with BigQuery. When you use CLI commands in a notebook, the command must be prepended with a !.

View available commands

To view the available commands for the BigQuery command-line tool, use the help command.

bq help

Create a new dataset

A dataset is contained within a specific project. Datasets are top-level containers that are used to organize and control access to your tables and views. A table or view must belong to a dataset. You need to create at least one dataset before loading data into BigQuery.

First, name your new dataset:

[1]: dataset_id = "myDataSet"

The following command creates a new dataset in the US using the ID defined above.

Mode: Command L 1, Col 1 BigQuery command-line tool.ipynb

현재 실행하고 있는 노트북 파일 모두 Shut Down Kernel 클릭

노트북 삭제

The screenshot shows the Google Cloud Platform Workbench interface. On the left, there is a sidebar with various options like Vertex AI, Workbench, Pipelines, Experiments, Model Registry, Endpoints, Forecasting, Metadata, Matching Engine, and Marketplace. The 'Workbench' option is currently selected and highlighted in blue. In the main content area, under the 'Workbench' heading, there is a sub-section titled '관리형 노트북' (Managed Notebook). This section contains a table with one row of data:

노트북 이름	소유자	커스텀 Docker 이미지	최종 수정 날짜
managed-notebook-1662472026	JUPYTERLAB 열기	Service account	2022. 9. 6. 오후 10:50:09

To the right of the table, there is a horizontal toolbar with several icons: 새 노트북 (New Notebook), 새로고침 (Refresh), 시작 (Start), 중지 (Stop), 재설정 (Reset), 업그레이드 (Upgrade), and 삭제 (Delete). The '삭제' (Delete) button is highlighted with an orange rectangle and has an orange arrow pointing towards it from the text on the right.

노트북 선택
후 삭제 클릭

프로젝트 종료

설정 – IAM 및 관리자 – My NB Proje X

console.cloud.google.com/iam-admin/settings?project=my-nb-project-361711

Google Cloud My NB Project

검색 제품, 리소스, 문서(/)

설정 이전 종료

알아보기

IAM 및 관리자

IAM

ID 및 조직

정책 문제 해결 도구

정책 분석자

조직 정책

서비스 계정

워크로드 아이덴티티 제휴

라벨

태그

설정

개인정보 보호 및 보안

IAP(Identity-Aware Proxy)

리소스 관리

출시 노트

프로젝트 이름 * My NB Project

프로젝트 ID my-nb-project-361711

프로젝트 번호 274652386556

액세스 투명성

조직에 속하지 않은 프로젝트에는 액세스 투명성이 제공되지 않습니다. 단일 프로젝트에 액세스 투명성을 사용 설정하려면 영업팀이나 지원팀에 문의하세요.

종료 클릭

프로젝트 종료 확인

The screenshot shows the Google Cloud IAM settings page for the project "My NB Project". A modal dialog box is open, titled "My NB Project" 프로젝트 종료, containing information about the consequences of deactivating the project. The "프로젝트 ID" field is highlighted with an orange rectangle and an arrow points from the text "프로젝트 ID 입력 후 종료 클릭" to it.

설정 – IAM 및 관리자 – My NB Project

Google Cloud My NB Project

IAM 및 관리자 설정 이전 종료 알아보기

프로젝트 이름 * My NB Project

프로젝트 ID my-nb-project

프로젝트 번호: 2746523865

"My NB Project" 프로젝트 종료

프로젝트를 종료하면 즉시 다음과 같은 결과가 발생합니다.

모든 결제 및 트래픽 전달이 중지됩니다.
프로젝트에 액세스할 수 없게 됩니다.
프로젝트 소유자에게 알림이 전송되며 소유자는 30일 이내에 삭제를 중지할 수 있습니다.
프로젝트가 30일 후에 삭제되도록 예약됩니다. 하지만 일부 리소스는 활성 먼저 삭제될 수 있습니다.

프로젝트가 30일 후에 삭제되도록 예약됩니다. 하지만 일부 리소스는 활성 먼저 삭제될 수 있습니다.

"My NB Project" 프로젝트를 종료하려면 다음 프로젝트 ID를 입력하세요. my-nb-project-361711

프로젝트 ID * my-nb-project-361711

취소 종료

프로젝트 명
입력 후 종료
클릭