구글 Genomics
2016.7.26
Install genomics tools 소개

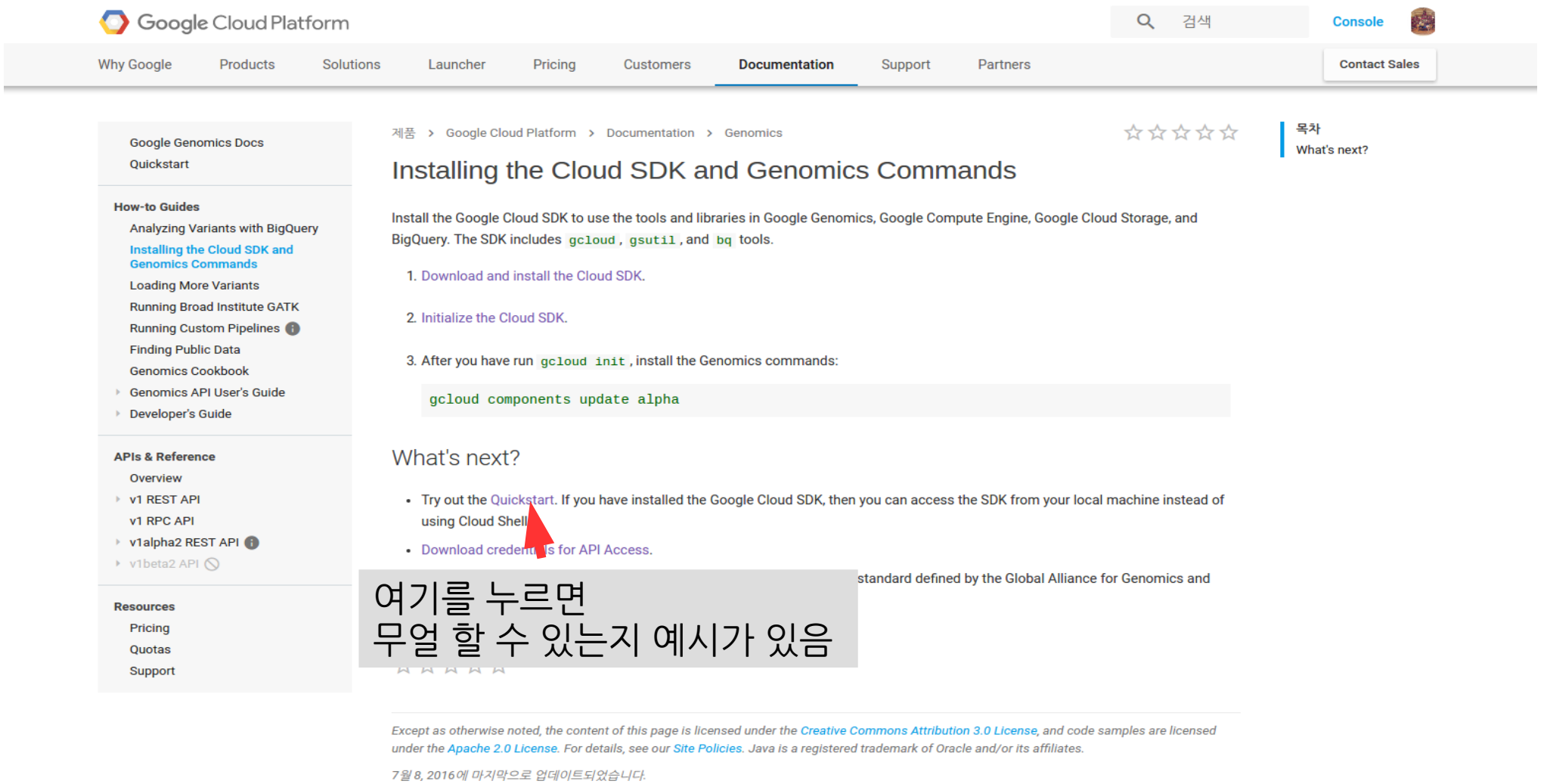
이병철

이전 발표 자료

https://github.com/biospin/BigBio/blob/master/part04/week01_160531/Getting%20Started%20With%20Google%20Genomics.pdf

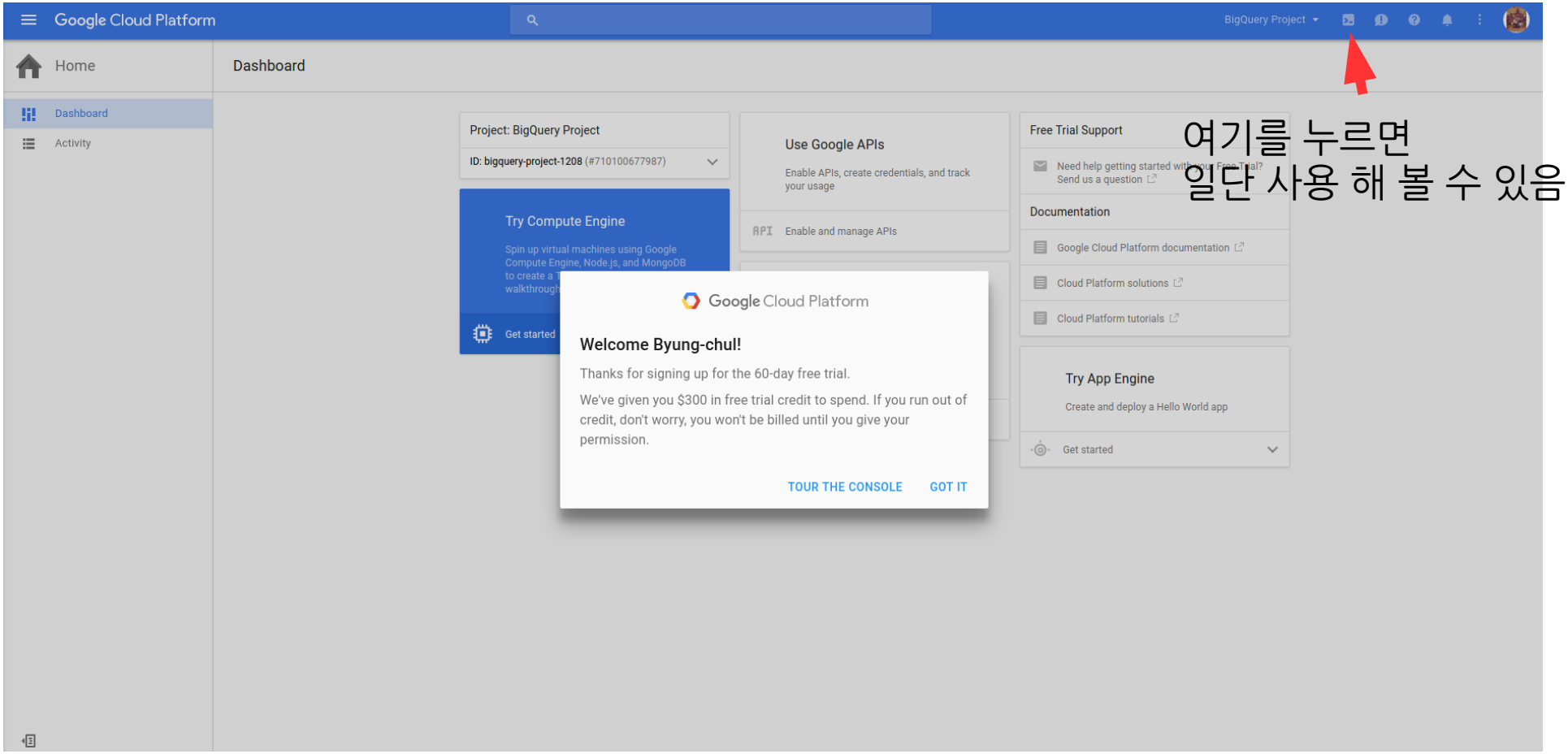
개요

주소: https://cloud.google.com/genomics/install-genomics-tools

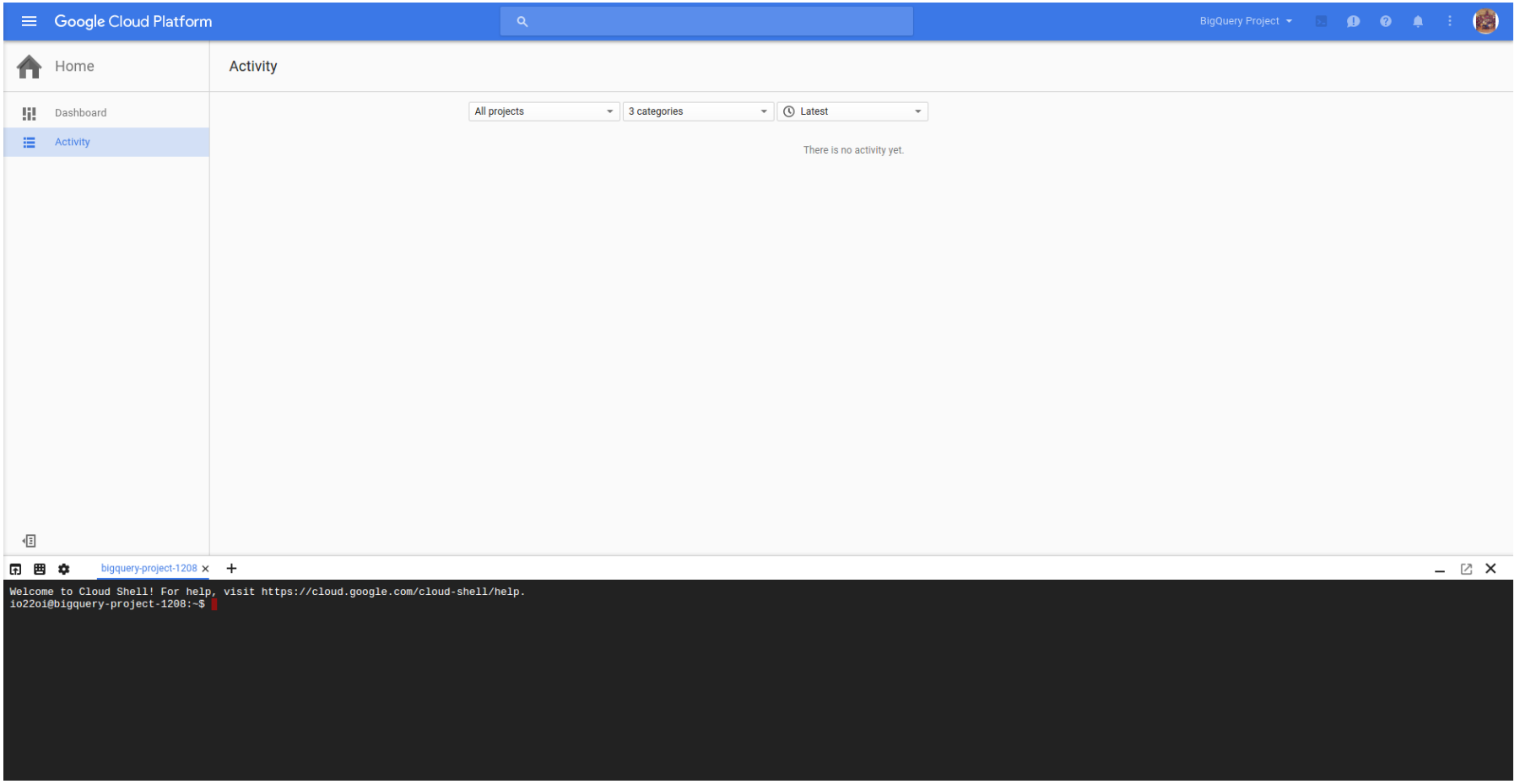


회원가입완료

우선 해야 할 일: 회원 가입



Google cloud shell



하나 씩 따라 하기

Installing the Cloud SDK and Genomics Commands

Install the Google Cloud SDK to use the tools and libraries in Google Genomics, Google Compute Engine, Google Cloud Storage, and BigQuery. The SDK includes `gcloud`, `gsutil`, and `bq` tools.

1. Download and install the Cloud SDK.

2. Initialize the Cloud SDK.

3. After you have run `gcloud init`, install the Genomics commands:

```
gcloud components update alpha
```

Download and install the cloud SDK

현재 3가지 설치 방법 선택 가능
1. versioned archive 이용
2. interactive installer 이용
3. (Debian/ubuntu) apt-get 이용

Apt-get 이용 예시

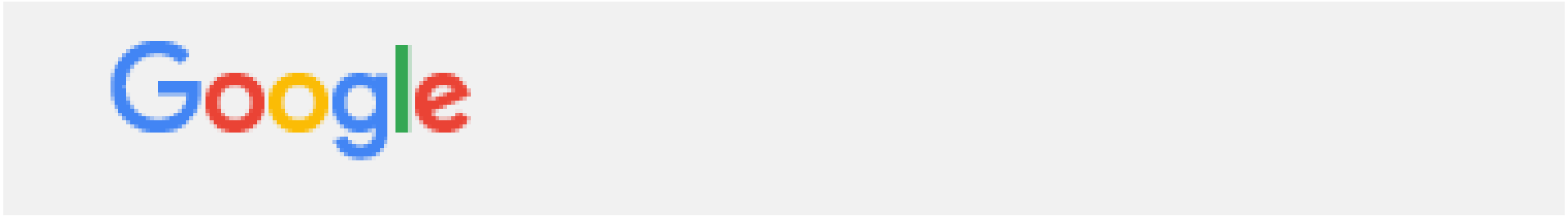
- # Create an environment variable for the correct distribution
- export CLOUD_SDK_REPO="cloud-sdk-$(lsb_release -c -s)"
-
- # Add the Cloud SDK distribution URI as a package source
- echo "deb http://packages.cloud.google.com/apt $CLOUD_SDK_REPO main" | sudo tee /etc/apt/sources.list.d/google-cloud-sdk.list
-
- # Import the Google Cloud public key
- curl https://packages.cloud.google.com/apt/doc/apt-key.gpg | sudo apt-key add -
-
- # Update and install the Cloud SDK
- sudo apt-get update && sudo apt-get install google-cloud-sdk
-
- # Run gcloud init to get started
- gcloud init

]$ gcloud init

UI가 없을 때는 --console-only를 붙여서 사용

```
ubuntu@ip-172-31-43-244:~$ gcloud init --console-only
Welcome! This command will take you through the configuration of gcloud.

Your current configuration has been set to: [default]

Network diagnostic detects and fixes local network connection issues.
Checking network connection...done.
Reachability Check passed.
Network diagnostic (1/1 check) passed.

You must log in to continue. Would you like to log in (Y/n)?  y

Go to the following link in your browser:
```

여기에 연결 주소가 나옴. 이것을 웹 브라우저에 복사 붙여넣기

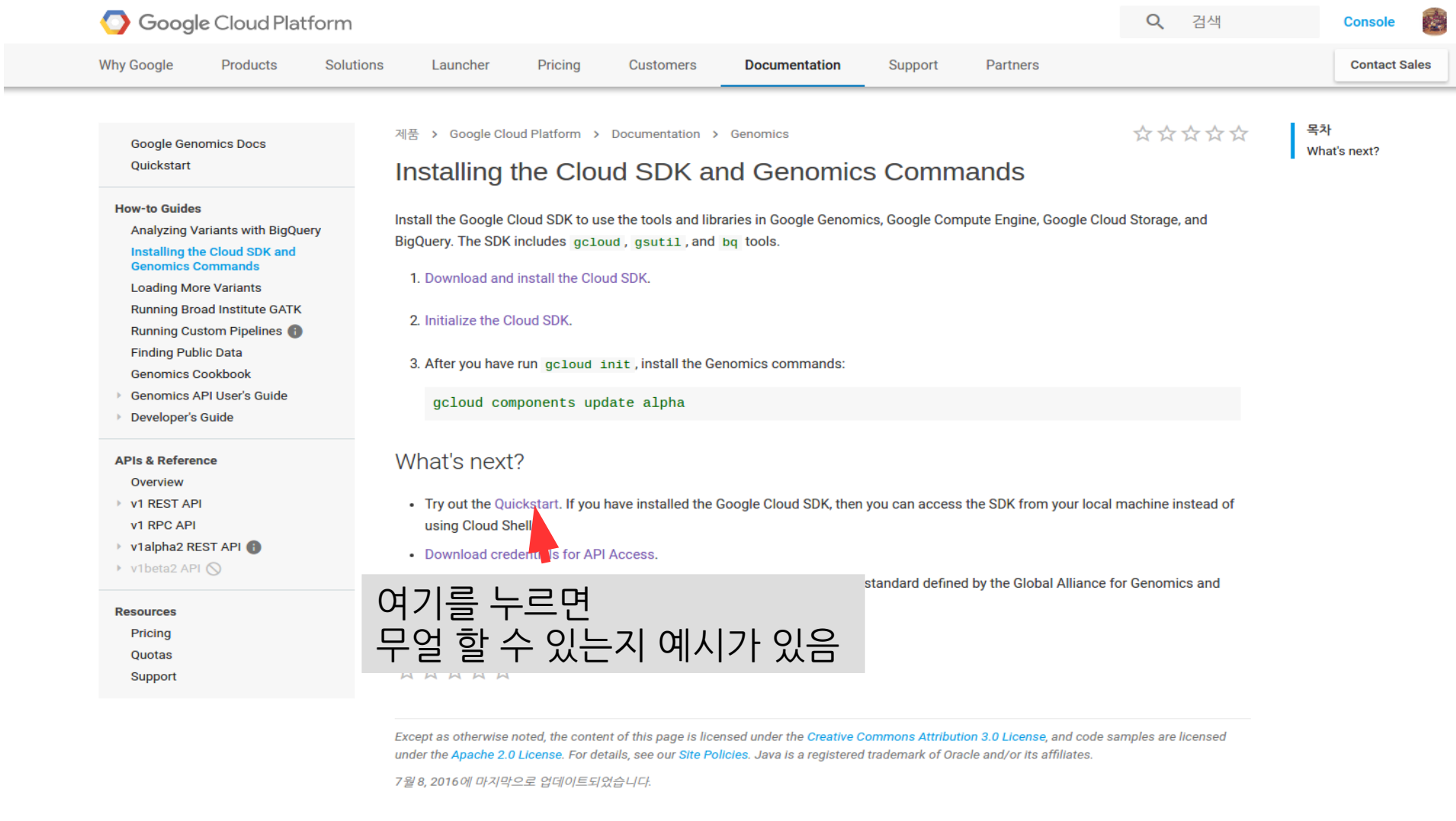


Please copy this code, switch to your application and paste it there:

여기에 key가 나옴. 이것을 다시 터미널에 붙여 넣기

```
Enter verification code: 4/[redacted]
You are logged in as: [[redacted]].

Your current project has been set to: [bigquery-project-1208].

Not setting default zone/region (this feature makes it easier to use
[gcloud compute] by setting an appropriate default value for the
--zone and --region flag).
See https://cloud.google.com/compute/docs/gcloud-compute section on how to set
default compute region and zone manually. If you would like [gcloud init] to be
able to do this for you the next time you run it, make sure the
Compute Engine API is enabled for your project on the
https://console.developers.google.com/apis page.

Created a default .boto configuration file at [/home/ubuntu/.boto]. See this fil
e and
[https://cloud.google.com/storage/docs/gsutil/commands/config] for more
information about configuring Google Cloud Storage.
Your Google Cloud SDK is configured and ready to use!

* Commands that require authentication will use io22oi@gmail.com by default
* Commands will reference project `bigquery-project-1208` by default
Run `gcloud help config` to learn how to change individual settings

This gcloud configuration is called [default]. You can create additional configu
rations if you work with multiple accounts and/or projects.
Run `gcloud topic configurations` to learn more.

Some things to try next:

* Run `gcloud --help` to see the Cloud Platform services you can interact with.
And run `gcloud help COMMAND` to get help on any gcloud command.
* Run `gcloud topic -h` to learn about advanced features of the SDK like arg fil
es and output formatting
```

이제 뭘 하지?


Google Cloud Platform
검색
Console
Why Google
Products
Solutions
Launcher
Pricing
Customers
Documentation
Support
Partners
Contact Sales
Google Genomics Docs
Quickstart
How-to Guides
Analyzing Variants with BigQuery
Installing the Cloud SDK and Genomics Commands
Loading More Variants
Running Broad Institute GATK
Running Custom Pipelines
Finding Public Data
Genomics Cookbook
Genomics API User's Guide
Developer's Guide
APIs & Reference
Overview
v1 REST API
v1 RPC API
v1alpha2 REST API
v1beta2 API
Resources
Pricing
Quotas
Support
제품 > Google Cloud Platform > Documentation > Genomics
Installing the Cloud SDK and Genomics Commands
Install the Google Cloud SDK to use the tools and libraries in Google Genomics, Google Compute Engine, Google Cloud Storage, and BigQuery. The SDK includes gcloud, gsutil, and bq tools.
1. Download and install the Cloud SDK.
2. Initialize the Cloud SDK.
3. After you have run gcloud init, install the Genomics commands:
gcloud components update alpha
What's next?
Try out the Quickstart. If you have installed the Google Cloud SDK, then you can access the SDK from your local machine instead of using Cloud Shell
Download credentials for API Access.
standard defined by the Global Alliance for Genomics and
목차
What's next?
여기를 누르면
무얼 할 수 있는지 예시가 있음
Except as otherwise noted, the content of this page is licensed under the Creative Commons Attribution 3.0 License, and code samples are licensed under the Apache 2.0 License. For details, see our Site Policies. Java is a registered trademark of Oracle and/or its affiliates.
7월 8, 2016에 마지막으로 업데이트되었습니다.

Quickstart

Quickstart

This page shows you how to set up and start using Google Genomics.

Before you begin

1. Select or create a Cloud Platform Console project.
   GO TO THE PROJECTS PAGE
2. Enable the Genomics, BigQuery, and Cloud Storage APIs.
   ENABLE THE APIS

Launch Cloud Shell to use the command line

You can use Cloud Shell to access the Google Cloud SDK, which includes tools and libraries that you need to create and manage resources on Google Cloud Platform, including Google Genomics, Google Compute Engine, Google Cloud Storage, and BigQuery.

To launch Cloud Shell:

1. Navigate to the project you want to use in the Cloud Platform Console.
2. Click the **Activate Google Cloud Shell** button at the top of the console window.

A Cloud Shell session opens inside a new frame at the bottom of the console and displays a command-line prompt.

Run a query

Query a dataset from the 1000 Genomes Project using the Genomics tools.

1. Search a variant set for variants at a specific location:

```
gcloud alpha genomics variants list --variant-set-id "10473108253681171589" --reference-name "22
```

This query returns the following variant:

```
VARIANT_SET_ID        REFERENCE_NAME  START     END       REFERENCE_BASES  ALTERNATE_BASES
10473108253681171589  22              51003835  51003836  A                [u'G']
```

2. Search callsets for individuals with calls (including reference calls) at the same location:

```
gcloud alpha genomics callsets list "10473108253681171589" --limit 10
```

This query returns the following individuals:

```
ID                      NAME     VARIANT_SET_IDS
10473108253681171589-0  HG00261  [u'10473108253681171589']
10473108253681171589-1  HG00593  [u'10473108253681171589']
10473108253681171589-2  NA12749  [u'10473108253681171589']
10473108253681171589-3  HG00150  [u'10473108253681171589']
10473108253681171589-4  NA19675  [u'10473108253681171589']
10473108253681171589-5  NA19651  [u'10473108253681171589']
10473108253681171589-6  NA19393  [u'10473108253681171589']
10473108253681171589-7  NA19207  [u'10473108253681171589']
10473108253681171589-8  HG00342  [u'10473108253681171589']
10473108253681171589-9  NA12546  [u'10473108253681171589']
```

외부 터미널에서도 가능

```
ubuntu@ip-172-31-43-244: ~
        "0.0003"
      ],
      "VT": [
        "SNP"
      ]
    },
    "names": [
      "rs131767",
      "rs131767"
    ],
    "quality": 100.0,
    "referenceBases": "A",
    "referenceName": "22",
    "start": "51003835",
    "variantSetId": "10473108253681171589"
  }
]
ubuntu@ip-172-31-43-244:~$ gcloud alpha genomics variants list --variant-set-id
"10473108253681171589" --reference-name "22" --start 51003835 --end 51003836
VARIANT_SET_ID        REFERENCE_NAME  START     END       REFERENCE_BASES  ALTER
NATE_BASES
10473108253681171589  22              51003835  51003836  A                [u'G'
]
ubuntu@ip-172-31-43-244:~$
```