Assignment 1, Cloud Computing

24MD0471 - Akhmetkan Temirlan
Exercise 1: Understanding Cloud Computing Models

- **Objective**: Explore different cloud computing models and understand their key differences.
- **Steps**:
    - Research the three primary cloud service models: Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS).
    - Create a table comparing these models in terms of control, flexibility, and use cases.
    - Identify examples of services offered by Google Cloud Platform (GCP) under each model.
- **Questions**:
    - What are the main differences between IaaS, PaaS, and SaaS?

        IaaS or Infrastructure as a Service provides virtualized computing resources over the internet and is particularly useful for custom hardware setups. Meanwhile, **PaaS or Platform as a Service** allows users to manage applications without having to address the underlying infrastructure. PaaS is perfect for quickly building and deploying apps. On the other hand, **SaaS or Software as a Service** requires a subscription to access software features over the internet. SaaS offers the least control, as the provider manages everything, making it ideal for collaboration tools.
   - Which GCP services fall under each of these models?

        In the context of Google Cloud Platform, these models are well represented. For **IaaS**, services such as Google Compute Engine and Google Cloud Storage are prominent. In the **PaaS** category, Google App Engine and Google Cloud Functions cater to application development needs. In the **SaaS** realm, Google Workspace—including tools like Gmail and Google Docs—provides convenient solutions for everyday tasks.
    - Provide a real-world example where each cloud service model might be the most appropriate choice.

        Lets illustrate these concepts with real-world examples. If someone starts a new company and needed a custom server for an application then they would likely choose Google Compute Engine. Google App Engine would be the best option for working on a web app who want to focus on coding instead of server management. And for daily tasks like managing emails and collaborating with a team, subscribing to Google Workspace would be a best choice, as it requires no maintenance from the user.

Exercise 2: Exploring Google Cloud Platform's Core Services

- **Objective**: Get acquainted with the core services provided by Google Cloud Platform.
- **Steps**:
  - Access the Google Cloud Console and navigate to the list of GCP services.
  - Explore and describe the purpose of the following core services:
    - Compute Engine
    - Google Kubernetes Engine (GKE)
    - App Engine
    - Cloud Storage
    - BigQuery
  - For each service, identify a potential use case in a business scenario.
- **Questions**:
  - **What is the primary use case of Compute Engine?**
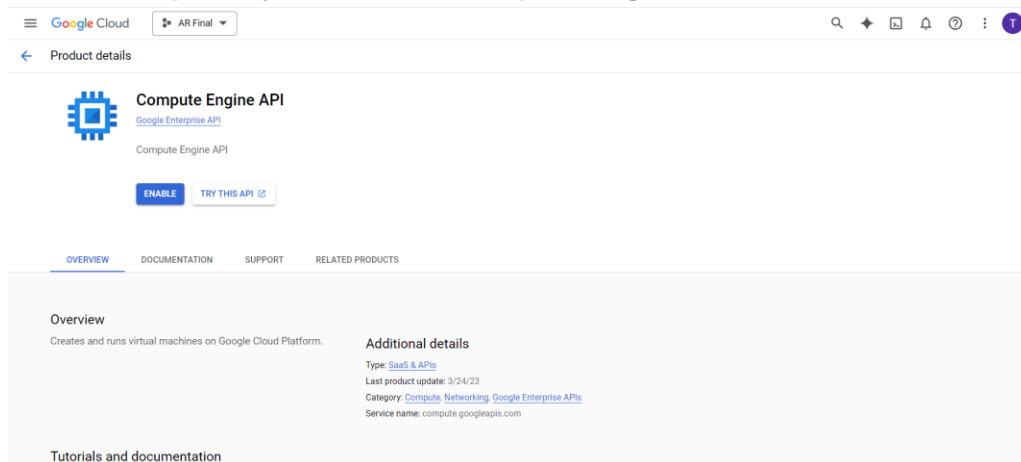
    the primary use case of Compute Engine is to offer scalable and customizable virtual machines for different tasks. Businesses can use it to run applications, host websites, or handle batch processing. It's especially helpful for companies that want the flexibility to adjust their virtual machines according to specific needs, like CPU, memory, and storage options.
 - **How does Google Kubernetes Engine (GKE) simplify the managementof containerized applications?**
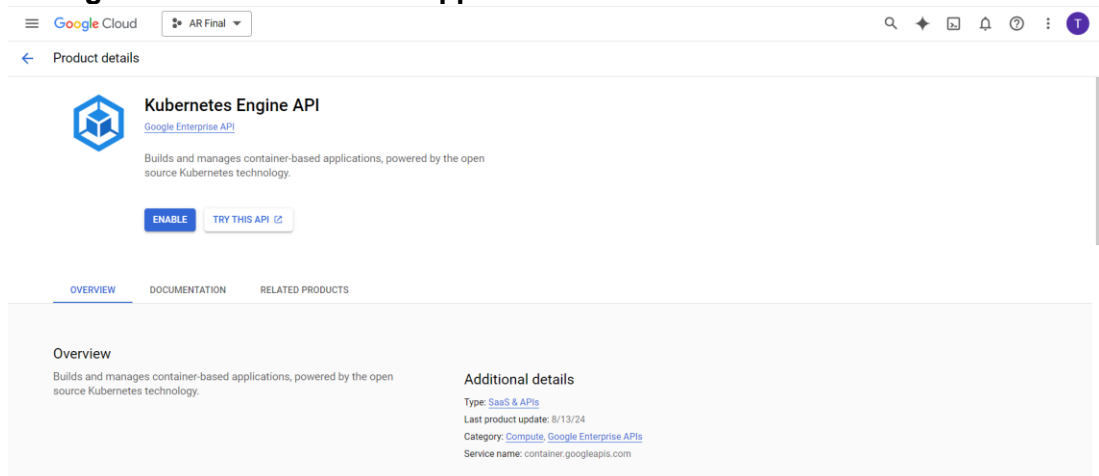
    Google Kubernetes Engine simplifies the management of containerized applications by automating many tasks related to deployment and scaling. It handles load balancing, scaling, and monitoring, allowing

developers to focus more on building their apps instead of worrying about the infrastructure. Also, it automates updates and health checks, making the entire process more efficient.

- **What advantages does Cloud Storage offer for data management?**
Cloud Storage offers several benefits for managing data. It is highly scalable - allowing users to store and access almost unlimited amounts of data without the fear of running out of space. It also provides high durability and availability, which automatically copying data across multiple locations. Another advantage is accessibility - users can reach their data from anywhere on the internet, making sharing and collaboration easy. Also it is cost-effective, with different pricing options that help businesses find what fits their budget and needs best.

- **Why would a business choose BigQuery for their data analysis needs?**
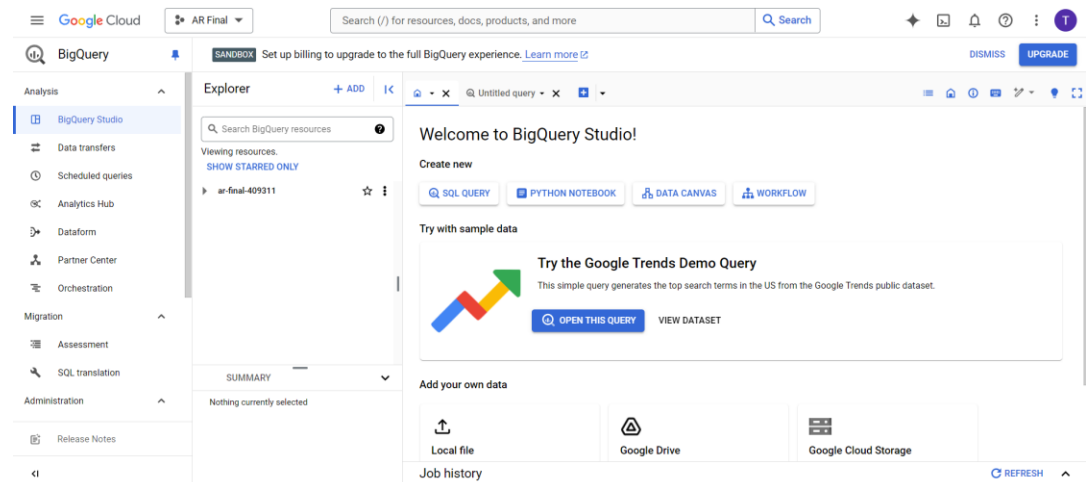
A business choose BigQuery for data analysis because it can quickly and efficiently manage large datasets. As a fully managed data warehouse, it enables real-time analytics and fast SQL queries, allowing companies to gain insights without handling complex infrastructure. Its integration with other Google Cloud services and support for machine learning make it an excellent option for businesses that want to use data for better decision making.

## Exercise 3: Creating and Managing Virtual Machines with Compute Engine

- **Objective**: Learn how to create, manage, and interact with virtual machines (VMs) using Compute Engine.
- **Steps**:
  - In the Google Cloud Console, navigate to Compute Engine and create a new VM instance.
  - Configure the VM with specific parameters, such as the machine type, region, and operating system.
  - Connect to the VM using SSH and install a basic web server (e.g., Apache or Nginx).
  - Stop, start, and delete the VM through the console.
- **Questions**:
  What steps did you follow to create the VM?
  To create a VM instance in Google Cloud Console, I navigated to the Compute Engine

section and clicked button "enable" and add billing account.
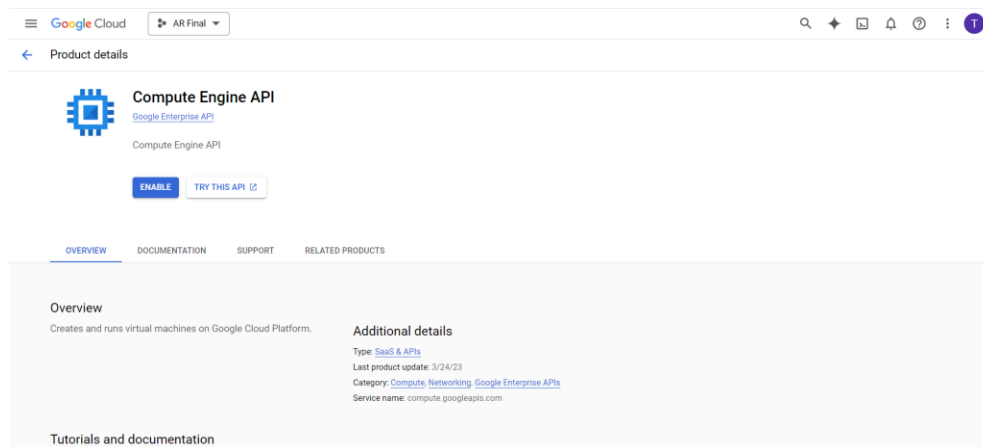
Next, clicked on "Create Instance."
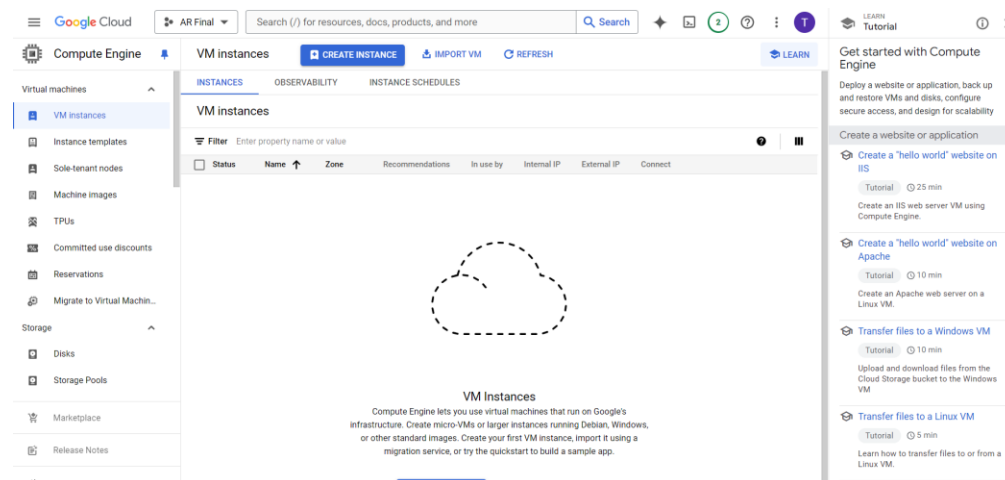
I then configured the VM by selecting the machine type, choosing the region, and picking the operating system I wanted to use. After confirming the settings, I clicked on the "Create" button to launch the VM.
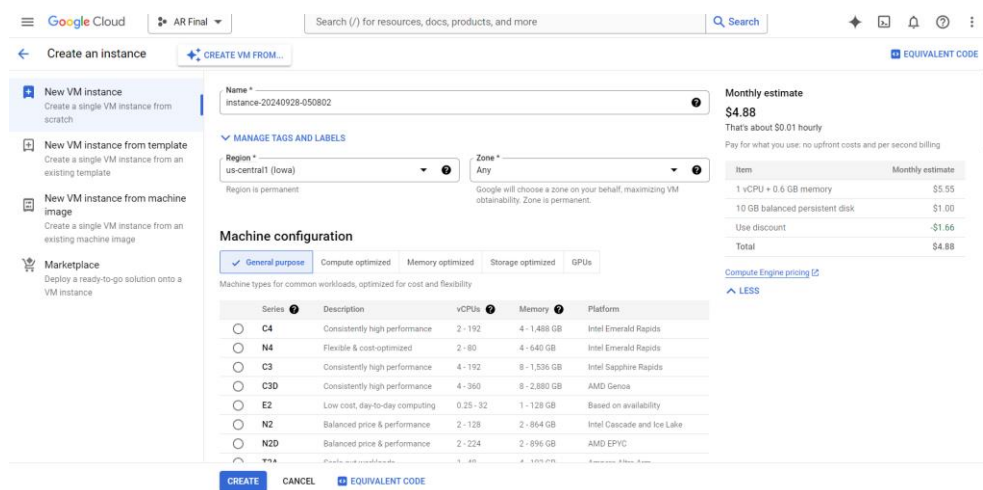
Finally, below is my new VM instance
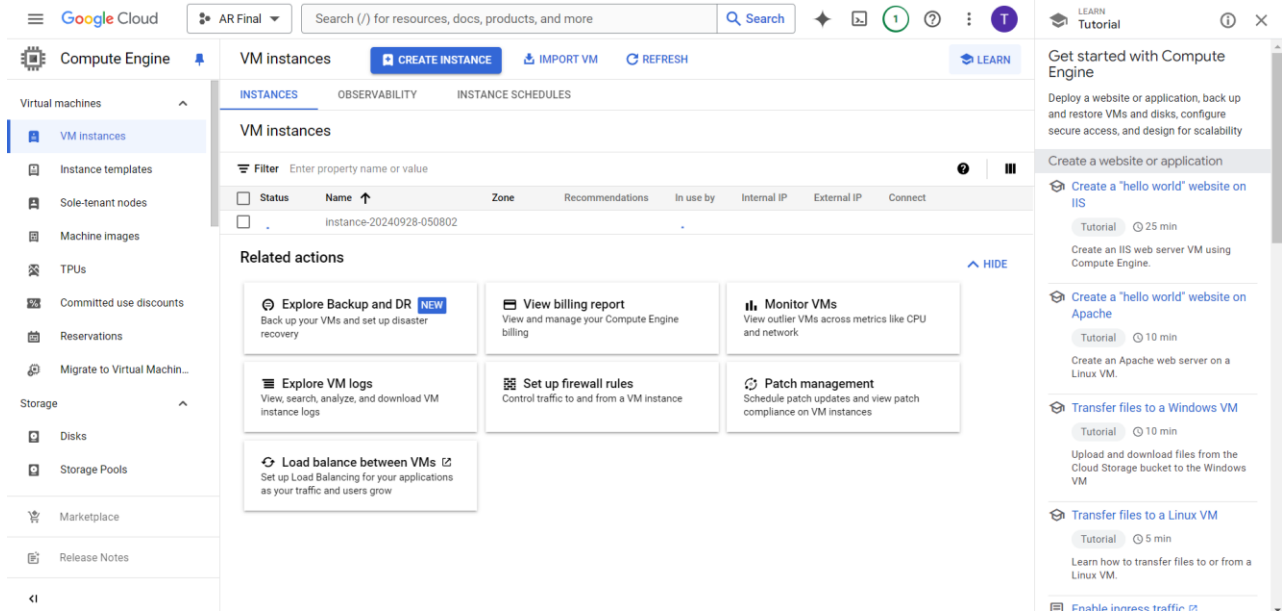
- How did you connect to the VM, and what commands did you use to install the web server?
  To connect to the VM, I used SSH. In the Google Cloud Console, I found the VM instance I created and clicked on the "SSH" button next to it.
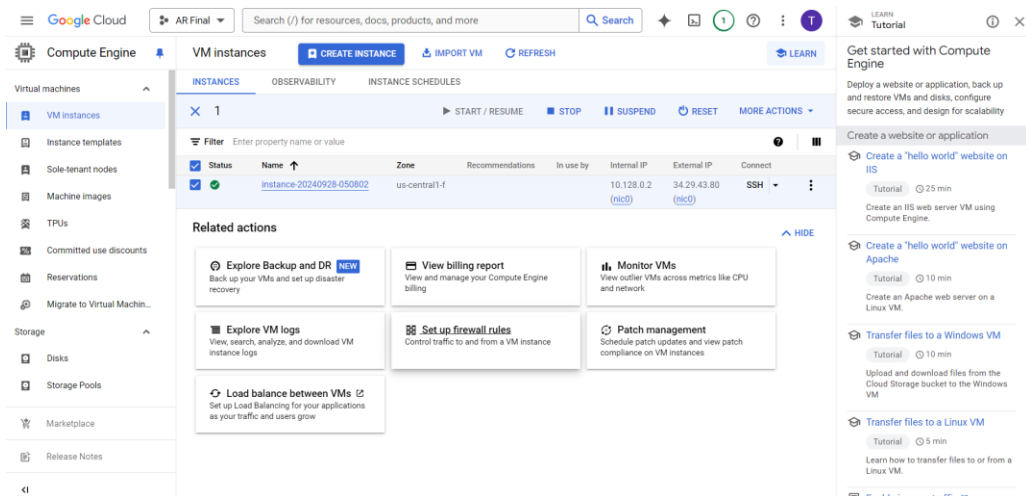
This opened a terminal window where I could run commands. To install a basic web server Apache, i used

--sudo apt update

--sudo apt install apache2

```
permitted by applicable law.
tema241202@instance-20240928-050802:~$ sudo apt update
Get:1 file:/etc/apt/mirrors/debian.list Mirrorlist [30 B]
Get:5 file:/etc/apt/mirrors/debian-security.list Mirrorlist [39 B]
Get:7 https://packages.cloud.google.com/apt google-compute-engine-bookworm-stable InRelease [1321 B]
Get:8 https://packages.cloud.google.com/apt cloud-sdk-bookworm InRelease [1654 B]
Get:2 https://deb.debian.org/debian bookworm InRelease [151 kB]
Get:9 https://packages.cloud.google.com/apt google-compute-engine-bookworm-stable/main amd64 Packages [3128 B]
Get:3 https://deb.debian.org/debian bookworm-updates InRelease [55.4 kB]
Get:4 https://deb.debian.org/debian bookworm-backports InRelease [59.0 kB]
Get:10 https://packages.cloud.google.com/apt cloud-sdk-bookworm/main all Packages [1555 kB]
Get:6 https://deb.debian.org/debian-security bookworm-security InRelease [48.0 kB]
Get:11 https://packages.cloud.google.com/apt cloud-sdk-bookworm/main amd64 Packages [3337 kB]
Get:12 https://deb.debian.org/debian bookworm-updates/main Sources.diff/Index [11.7 kB]
Get:13 https://deb.debian.org/debian bookworm-updates/main amd64 Packages.diff/Index [11.7 kB]
Get:14 https://deb.debian.org/debian bookworm-updates/main Translation-en.diff/Index [11.7 kB]
Get:15 https://deb.debian.org/debian bookworm-updates/main Sources T-2024-09-10-2011.55-F-2024-09-10-2011.55.pdif
f [562 B]
Get:16 https://deb.debian.org/debian bookworm-updates/main amd64 Packages T-2024-09-10-2011.55-F-2024-09-10-2011.
55.pdiff [1116 B]
Get:15 https://deb.debian.org/debian bookworm-updates/main Sources T-2024-09-10-2011.55-F-2024-09-10-2011.55.pdif
f [562 B]
Get:16 https://deb.debian.org/debian bookworm-updates/main amd64 Packages T-2024-09-10-2011.55-F-2024-09-10-2011.
55.pdiff [1116 B]
Get:20 https://deb.debian.org/debian bookworm-updates/main Translation-en T-2024-09-10-2011.55-F-2024-09-10-2011.
55.pdiff [538 B]
Get:20 https://deb.debian.org/debian bookworm-updates/main Translation-en T-2024-09-10-2011.55-F-2024-09-10-2011.
55.pdiff [538 B]
Get:17 https://deb.debian.org/debian bookworm-backports/main Sources.diff/Index [63.3 kB]
Get:18 https://deb.debian.org/debian bookworm-backports/main amd64 Packages.diff/Index [63.3 kB]
Get:19 https://deb.debian.org/debian bookworm-backports/main Translation-en.diff/Index [63.3 kB]
Get:24 https://deb.debian.org/debian bookworm-backports/main Sources T-2024-09-28-0205.59-F-2024-09-10-2011.55.pd
iff [32.1 kB]
Get:24 https://deb.debian.org/debian bookworm-backports/main Sources T-2024-09-28-0205.59-F-2024-09-10-2011.55.pd
iff [32.1 kB]
Get:25 https://deb.debian.org/debian bookworm-backports/main amd64 Packages T-2024-09-27-2010.31-F-2024-09-11-020
4.35.pdiff [51.7 kB]
Get:25 https://deb.debian.org/debian bookworm-backports/main amd64 Packages T-2024-09-27-2010.31-F-2024-09-11-020
4.35.pdiff [51.7 kB]
```

```
Get:23 https://deb.debian.org/debian-security bookworm-security/main Translation-en [111 kB]
Fetched 5936 kB in 1s (4566 kB/s)
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
4 packages can be upgraded. Run 'apt list --upgradable' to see them.
tema241202@instance-20240928-050802:~$ sudo apt install apache2
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
The following additional packages will be installed:
  apache2-bin apache2-data apache2-utils libapr1 libaprutil1 libaprutil1-dbd-sqlite3 libaprutil1-ldap
  libjansson4 liblua5.3-0 ssl-cert
Suggested packages:
  apache2-doc apache2-suexec-pristine | apache2-suexec-custom www-browser
The following NEW packages will be installed:
  apache2 apache2-bin apache2-data apache2-utils libapr1 libaprutil1 libaprutil1-dbd-sqlite3 libaprutil1-ldap
  libjansson4 liblua5.3-0 ssl-cert
0 upgraded, 11 newly installed, 0 to remove and 4 not upgraded.
Need to get 2378 kB of archives.
After this operation, 8468 kB of additional disk space will be used.
Do you want to continue? [Y/n] Y
Get:1 file:/etc/apt/mirrors/debian.list Mirrorlist [30 B]
Get:2 https://deb.debian.org/debian bookworm/main amd64 libapr1 amd64 1.7.2-3 [102 kB]
Get:3 https://deb.debian.org/debian bookworm/main amd64 libaprutil1 amd64 1.6.3-1 [87.8 kB]
Get:4 https://deb.debian.org/debian bookworm/main amd64 libaprutil1-dbd-sqlite3 amd64 1.6.3-1 [13.6 kB]
Get:5 https://deb.debian.org/debian bookworm/main amd64 libaprutil1-ldap amd64 1.6.3-1 [11.8 kB]
Get:6 https://deb.debian.org/debian bookworm/main amd64 libjansson4 amd64 2.14-2 [40.8 kB]
Get:7 https://deb.debian.org/debian bookworm/main amd64 liblua5.3-0 amd64 5.3.6-2 [123 kB]
Get:8 https://deb.debian.org/debian bookworm/main amd64 apache2-bin amd64 2.4.62-1~deb12u1 [1385 kB]
Get:9 https://deb.debian.org/debian bookworm/main amd64 apache2-data all 2.4.62-1~deb12u1 [160 kB]
Get:10 https://deb.debian.org/debian bookworm/main amd64 apache2-utils amd64 2.4.62-1~deb12u1 [210 kB]
Get:11 https://deb.debian.org/debian bookworm/main amd64 apache2 amd64 2.4.62-1~deb12u1 [223 kB]
Get:12 https://deb.debian.org/debian bookworm/main amd64 ssl-cert all 1.1.2 [21.1 kB]
Fetched 2378 kB in 0s (13.3 MB/s)
Preconfiguring packages ...
Selecting previously unselected package libapr1:amd64.
(Reading database ... 69885 files and directories currently installed.)
Preparing to unpack .../00-libapr1_1.7.2-3_amd64.deb ...
```

To install a basic web server Nginx, i used

```
--sudo apt update
--sudo apt install nginx
```

```
tema241202@instance-20240928-050802:~$ sudo apt update
Get:1 file:/etc/apt/mirrors/debian.list Mirrorlist [30 B]
Get:5 file:/etc/apt/mirrors/debian-security.list Mirrorlist [39 B]
Hit:7 https://packages.cloud.google.com/apt google-compute-engine-bookworm-stable InRelease
Hit:8 https://packages.cloud.google.com/apt cloud-sdk-bookworm InRelease
Hit:2 https://deb.debian.org/debian bookworm InRelease
Hit:3 https://deb.debian.org/debian bookworm-updates InRelease
Hit:4 https://deb.debian.org/debian bookworm-backports InRelease
Get:6 https://deb.debian.org/debian-security bookworm-security InRelease [48.0 kB]
Get:9 https://deb.debian.org/debian-security bookworm-security/main Sources [111 kB]
Fetched 159 kB in 1s (159 kB/s)
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
4 packages can be upgraded. Run 'apt list --upgradable' to see them.
tema241202@instance-20240928-050802:~$
```

```
tema241202@instance-20240928-050802:~$ sudo apt install nginx
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
The following additional packages will be installed:
  nginx-common
Suggested packages:
  fcgiwrap nginx-doc
The following NEW packages will be installed:
  nginx nginx-common
0 upgraded, 2 newly installed, 0 to remove and 4 not upgraded.
Need to get 640 kB of archives.
After this operation, 1696 kB of additional disk space will be used.
Do you want to continue? [Y/n] Y
Get:1 file:/etc/apt/mirrors/debian.list Mirrorlist [30 B]
Get:2 https://deb.debian.org/debian bookworm/main amd64 nginx-common all 1.22.1-9 [112 kB]
Get:3 https://deb.debian.org/debian bookworm/main amd64 nginx amd64 1.22.1-9 [527 kB]
Fetched 640 kB in 0s (4097 kB/s)
Preconfiguring packages ...
Selecting previously unselected package nginx-common.
(Reading database ... 70620 files and directories currently installed.)
Preparing to unpack .../nginx-common_1.22.1-9_all.deb ...
Unpacking nginx-common (1.22.1-9) ...
Selecting previously unselected package nginx.
Preparing to unpack .../nginx_1.22.1-9_amd64.deb ...
Unpacking nginx (1.22.1-9) ...
Setting up nginx-common (1.22.1-9) ...
Created symlink /etc/systemd/system/multi-user.target.wants/nginx.service → /lib/systemd/system/nginx.service.
Could not execute systemctl:  at /usr/bin/deb-systemd-invoke line 145.
Setting up nginx (1.22.1-9) ...
Not attempting to start NGINX, port 80 is already in use.
Processing triggers for man-db (2.11.2-2) ...
tema241202@instance-20240928-050802:~$
```

- What happens to the VM and its data when it is stopped versus when it
  is deleted?
  When the VM is stopped, it keeps all its data and settings, so it can be
  started again later. While the VM is stopped, You won't be charged for
  the resources while the VM is stopped, but you may still have charges
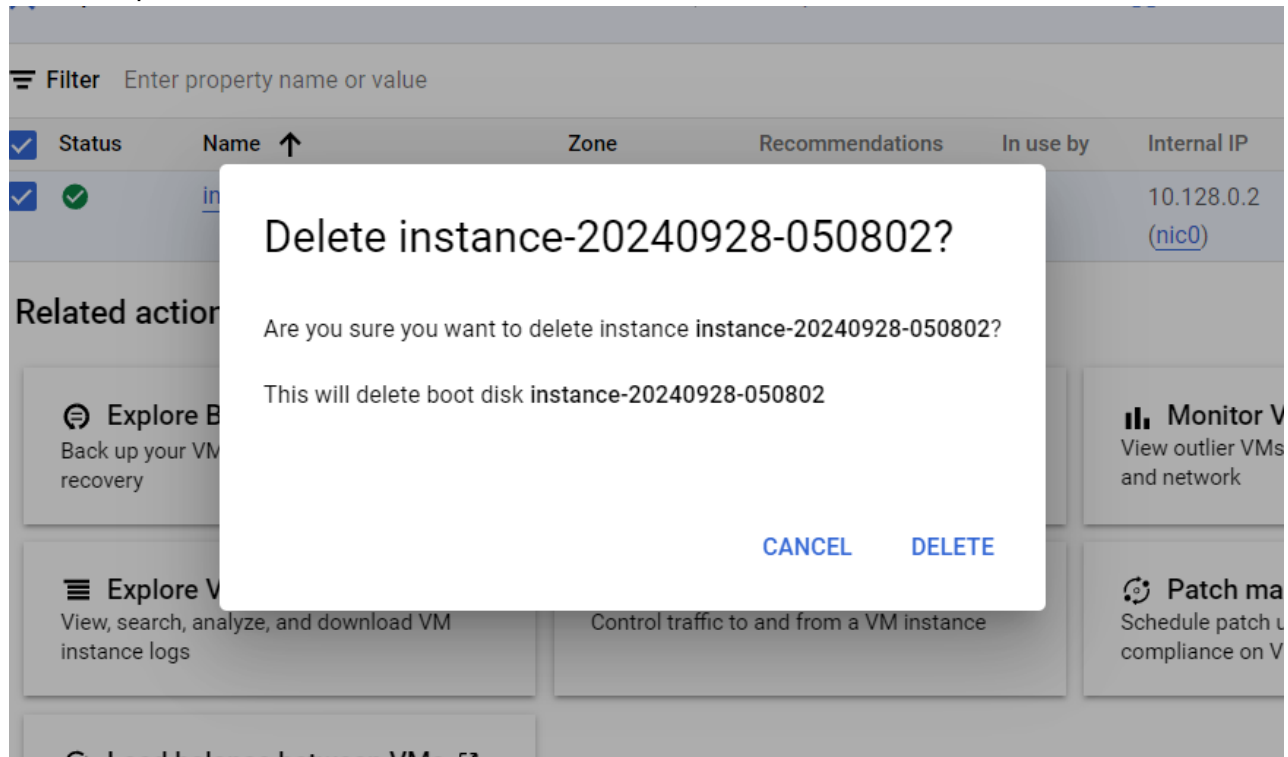 for the persistent disk and any static IP addresses.

☰ **Google** Cloud   • AR Final ▾   Search (/) for resources, docs, products, and more   🔍 Search   ✦ ▣ ① ? ⋮ T

⚙ **Compute Engine** 📌

Virtual machines   ⌃

| | VM instances | | **CREATE INSTANCE** | **⬆ IMPORT VM** | **🔄 REFRESH** | | 🎓 LEARN |

Deploy a website or application, back up and restore VMs and disks, configure secure access, and design for scalability

INSTANCES   OBSERVABILITY   INSTANCE SCHEDULES

Create a website or application

🖥 **VM instances**

VM instances

🎓 Create a "hello world" website on IIS

📄 Instance templates

≡ Filter Enter property name or value   ❷ ▥

Tutorial 🕐 25 min

🖥 Sole-tenant nodes

☐ Status ... Internal IP External IP Connect

Create an IIS web server VM using Compute Engine.

🖥 Machine images

☐ ✅ ... 10.128.0.2 34.29.43.80 SSH ▾ ⋮

🎓 Create a "hello world" website on Apache

⚙ TPUs

Stop instance-20240928-050802?

(nic0) (nic0)

Tutorial 🕐 10 min

🏷 Committed use discounts

Related

You'll be billed only for these preserved resources:

⌃ HIDE

Create an Apache web server on a Linux VM.

📅 Reservations

• Persistent disks

🎓 Transfer files to a Windows VM

⚙ Migrate to Virtual Machin...

• Static IP addresses

E... Monitor VMs

Tutorial 🕐 10 min

Storage   ⌃

The VM will gracefully shut down in 90 seconds. If processes are still running, the VM will be forced to stop and files may get corrupted.

Back u... recove... ew outlier VMs across metrics like CPU d network

Upload and download files from the Cloud Storage bucket to the Windows VM

💾 Disks

E...

🎓 Transfer files to a Linux VM

💾 Storage Pools

CANCEL **STOP**

Patch management

Tutorial 🕐 5 min

View, ... instan...

hedule patch updates and view patch mpliance on VM instances

Learn how to transfer files to or from a Linux VM.

🛒 Marketplace

🔄 **Load balance between VMs** ↗

📄 Release Notes

Set up Load Balancing for your applications as your traffic and users grow

⊲

Enable ingress traffic ↗

VM instance stopped   ✕

If the VM is deleted, all data on it is permanently lost, including any installed software and settings. However, if the data was saved on a separate persistent disk, that disk can be kept for future use.

Status	Name ↑	Zone	Recommendations	In use by	Internal IP
✓	in				10.128.0.2 (nic0)

Delete instance-20240928-050802?

Are you sure you want to delete instance **instance-20240928-050802**?

This will delete boot disk **instance-20240928-050802**

CANCEL   DELETE

Related action

Explore B
Back up your VM
recovery

Explore V
View, search, analyze, and download VM
instance logs

Control traffic to and from a VM instance

Monitor V
View outlier VMs
and network

Patch ma
Schedule patch
compliance on V

Load balance between VMs

or other standard images. Create your first VM insta
migration service, or try the quickstart to build
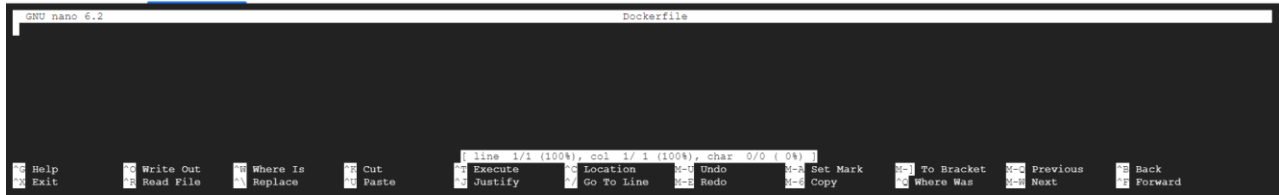
Instance deleted   ✕

CREATE INSTANCE   TAKE THE QU

Exercise 4: Deploying a Containerized Application on Google Kubernetes Engine (GKE)

- **Objective**: Understand how to deploy and manage containerized applications using Google Kubernetes Engine.
- **Steps**:
  - Create a simple Docker container for a web application.

- Push the container image to Google Container Registry (GCR).
- Create a GKE cluster in Google Cloud Console.
- Deploy the containerized application to the GKE cluster.
- Expose the application to the internet and verify its accessibility.

- **Questions**:
  - How did you create and push the Docker container to GCR?
    I opened the Terminal, using Google Cloud Shell, open it from the Google Cloud Console and use a text editor nano in the terminal to create a Dockerfile.

```
GNU nano 6.2                                        Dockerfile

                              [ line  1/1 (100%), col  1/ 1 (100%), char  0/0 ( 0%) ]
^G Help      ^O Write Out  ^W Where Is  ^K Cut       ^T Execute   ^C Location   M-U Undo    M-A Set Mark   M-] To Bracket   M-Q Previous   ^B Back
^X Exit      ^R Read File  ^\ Replace   ^U Paste     ^J Justify   ^/ Go To Line M-E Redo    M-6 Copy       ^Q Where Was     M-W Next       ^F Forward
```

Write the necessary instructions simple web application using Nginx in the Dockerfile and save it with pressing Ctrl + O

```
GNU nano 6.2                                        Dockerfile *
from nginx:latest
COPY ./usr/share/nginx/html


^G Help      ^O Write Out  ^W Where Is  ^K Cut       ^T Execute   ^C Location   M-U Undo    M-A Set Mark   M-] To Bracket   M-Q Previous   ^B Back
^X Exit      ^R Read File  ^\ Replace   ^U Paste     ^J Justify   ^/ Go To Line M-E Redo    M-6 Copy       ^Q Where Was     M-W Next       ^F Forward
```

After saving, pressed Ctrl + X to exit the Nano editor and proceeded to build Docker image "tempo" by running the following command in the terminal:

From that moment my google cloud account started to charge, so next steps I will tell without screenshots

--docker build -t tempo .

Next step was authenticate with Google Cloud, still in the terminal, I logged in to Google Cloud account and get window with password:

--gcloud auth login

Using the terminal I tagged image for Google Container Registry and finally, pushed tagged image to Google Container Registry by commands below:

--docker tag tempo gcr.io/340511/tempo

--docker push gcr.io/340511/tempo

- What steps were involved in setting up the GKE cluster?
  To set up the GKE cluster, I navigated to the Google Cloud Console and selected Kubernetes Engine. I clicked on "Create Cluster"

configured the settings, such as choosing the cluster type, specifying the number of nodes, and selecting the region.

After reviewing the settings, I clicked "Create" to launch the cluster.

- How did you verify that your application was successfully deployed and accessible?
  Open terminal Google Cloud Shell and I connected to GKE cluster

gcloud container clusters get-credentials autopilot-cluster-1 --region us-central1 --project project1-340511

After running the command, I checked the current context

kubectl config current-context:

 Used the command

kubectl create deployment tempodeployment --image=gcr.io/340511/tempo

to create a new deployment

To see the status of your deployments,I runned:

kubectl get deployments

Next, I checked the service associated with application to see how it's exposed to the internet.

I identified that it is ClusterIP: This type is only accessible within the cluster and won't work for external access. So i decided create new service with LoadBalancer.

kubectl expose deployment tempodeployment --type=LoadBalancer --name=kubernetes2 --port=80 --target-port=8080

Check again services and get kubernetes2

So my external IP is 34.41.112.201 and can use to access application.

Exercise 5: Storing and Accessing Data in Google Cloud Storage

- **Objective**: Learn how to store, manage, and access data using Google Cloud Storage.
- **Steps**:
    - Create a new Cloud Storage bucket in the Google Cloud Console.
    - Upload various types of files (e.g., text, images, videos) to the bucket.

    - Set access permissions for the bucket and test public and private access to the files.
    - Use the Cloud Console to download, move, and delete files in the bucket.
- **Questions**:
    - How do you create a Cloud Storage bucket, and what options are available during setup?
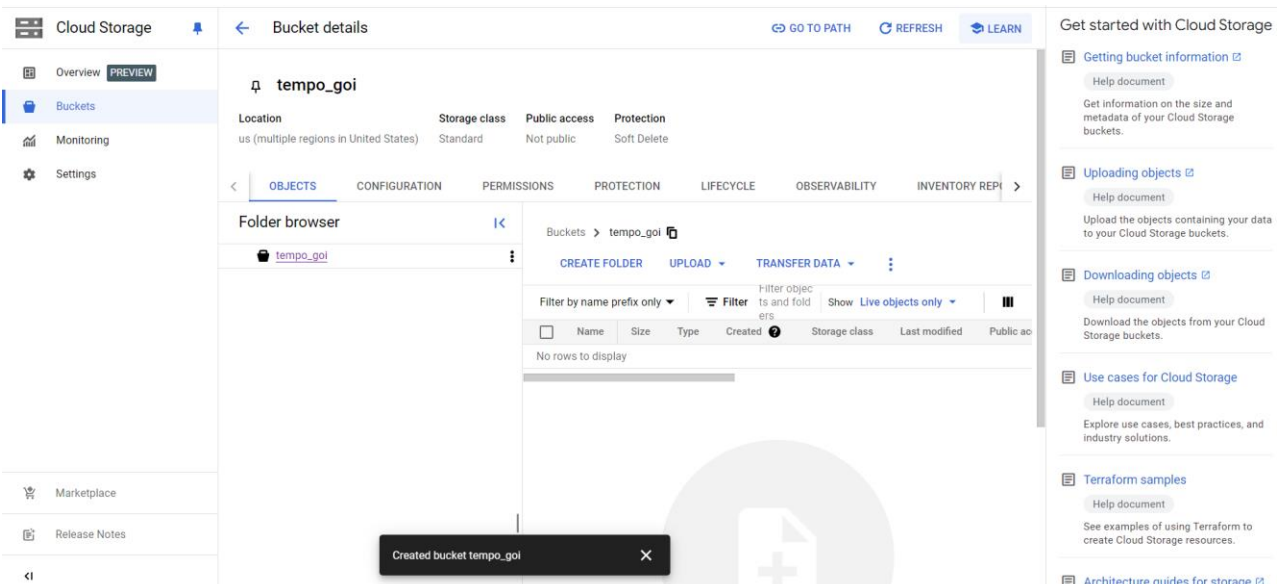
To create a Cloud Storage bucket, I go to the Google Cloud Console and navigate to the Cloud Storage section. Then, I click on "Create" near buckets.
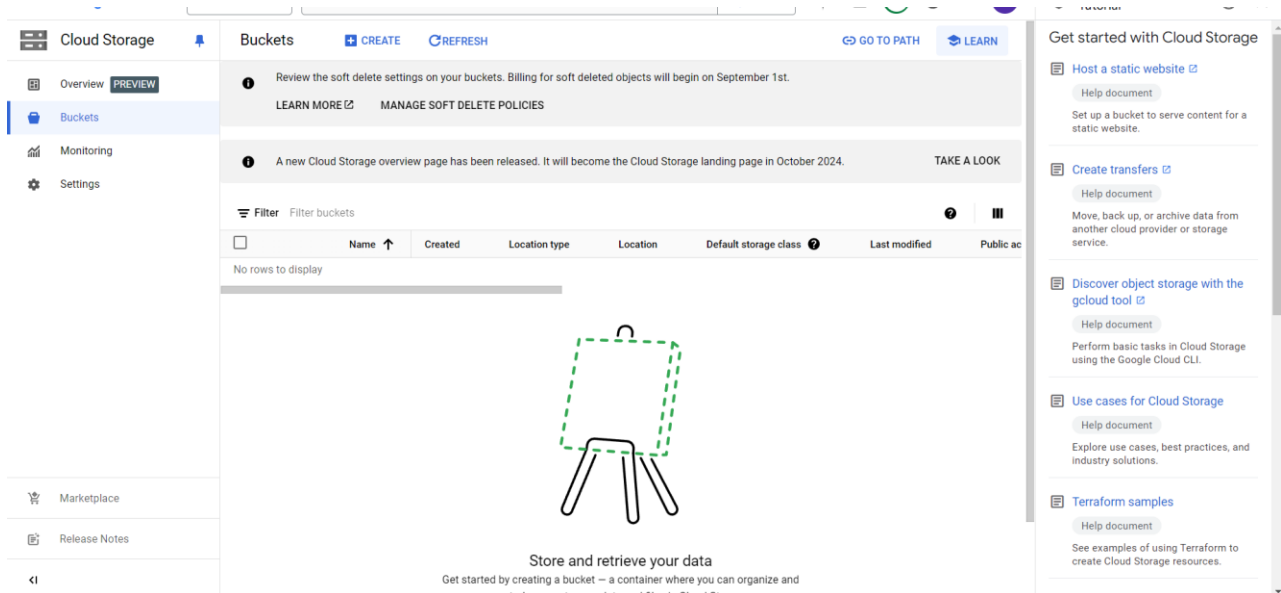
Set upped my bucket, choosing a unique name for my bucket, select a location and pick a storage class and decide on the access control method. Clicked create.
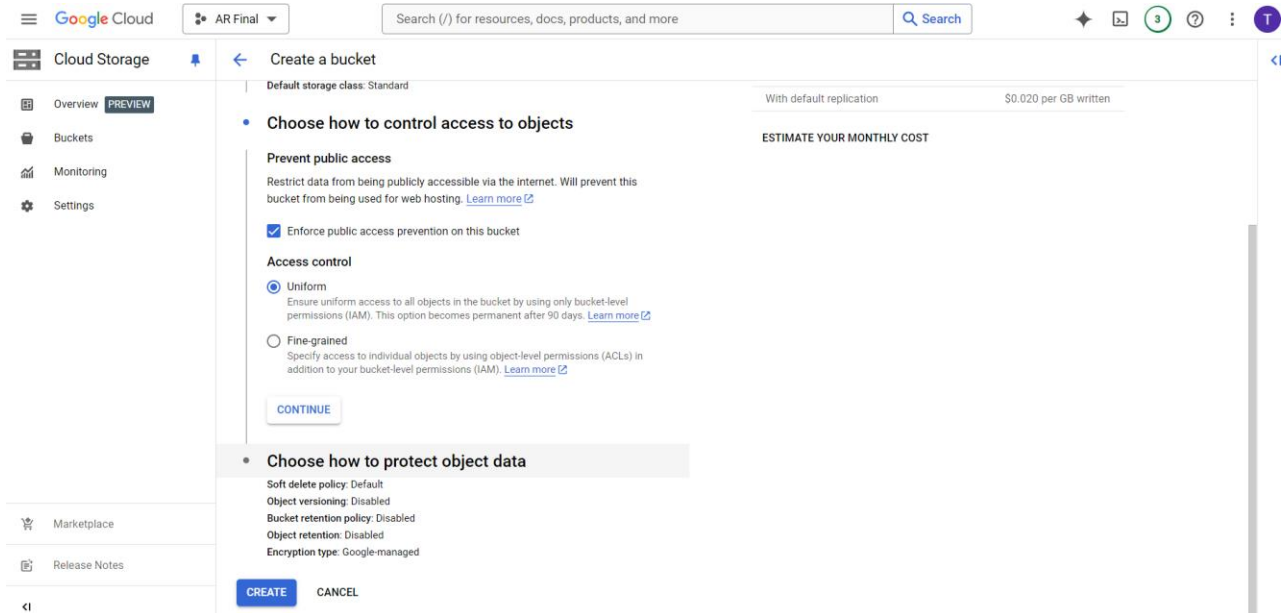
And got tempo_goi bucket:

- What are the differences between setting a bucket to public versus private?
  Setting a bucket to public means anyone on the internet can access its files without needing authentication, which is great for sharing things widely. But a private bucket restricts access, so only people with the right permissions can view or download the files. This is important if I have sensitive data that shouldn't be accessible to everyone.
- How can you manage access permissions for individual files in a bucket?
  To manage access for individual files, I can select the file in the Cloud Console and go to the Permissions tab. From there, I can add or remove users, groups, or service accounts, and specify what roles they have, like Viewer or Editor. This way, I can control who can see or modify each file separately.

Exercise 6: Analyzing Data with BigQuery

- **Objective**: Perform data analysis tasks using BigQuery.
- **Steps**:
    - Access BigQuery in the Google Cloud Console.
    - Create a dataset and table by importing a sample dataset provided by Google.
    - Write and execute SQL queries to perform basic data analysis, such as filtering, aggregation, and sorting.
    - Visualize the results using Google Data Studio or another visualization tool.
- **Questions**:
    - What steps did you take to create a dataset and table in BigQuery?

To create a dataset in BigQuery, I first logged into the Google Cloud Console and went to the BigQuery section. I clicked on my project, then hit "Create Dataset."
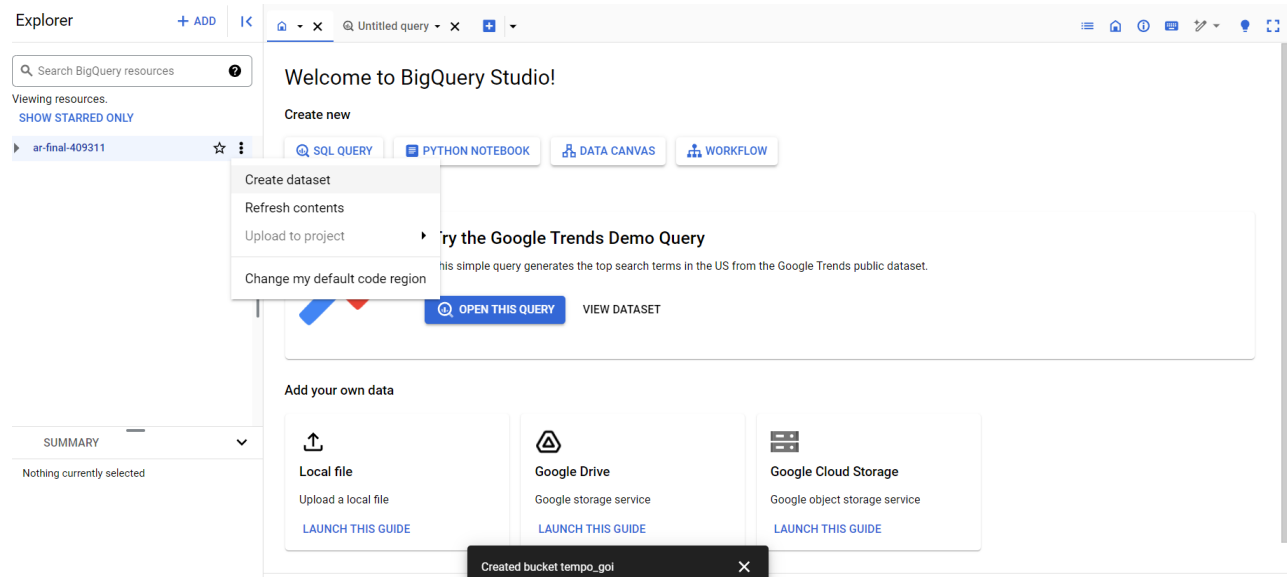
I picked a name " test_dataset" for the dataset and set up the necessary options.

Create dataset

Project ID *
ar-final-409311   CHANGE

Dataset ID *
test_dataset

Letters, numbers, and underscores allowed

Location type ❓

○ Region
   Specify a region to colocate your datasets with other Google Cloud services.

◉ Multi-region
   Allow BigQuery to select a region within a group to achieve higher quota limits.

Multi-region *
US (multiple regions in United States)   ▾

External Dataset

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

☐ Link to an external dataset ❓

Default table expiration

☐ Enable table expiration ❓

Default maximum table age   Days

Tags   ⌄

CREATE DATASET   CANCEL

After that, I created a table "test_table" by clicking "Create Table,"

Search BigQuery resources

Viewing resources.
SHOW STARRED ONLY

ar-final-409311

- Workflows
- External connections
- test_dataset

SUMMARY

Nothing currently selected

Welcome to BigQuery Studio!

Create new

SQL QUERY   PYTHON NOTEBOOK   DATA CANVAS   WORKFLOW

Recently opened

...taset
...nal-409311   Just now

Open
Open in

Create table   ...ple data

Share   **Try the Google Trends Demo Query**
Copy ID   This simple query generates the top search terms in the US from the Google Trends public dataset.

Refresh contents
Delete   OPEN THIS QUERY   VIEW DATASET

Add your own data

Job history   "test_dataset" created.   GO TO DATASET   ✕   ↻ REFRESH

SCHEMA	DETAILS	PREVIEW	TABLE EXPLORER PREVIEW	INSIGHTS	LINEAGE	DATA PROFILE

Filter   Enter property name or value

	Field name	Type	Mode	Key	Collation	Default Value	Policy Tags ❓	Description
☐	**id**	INTEGER	NULLABLE	-	-	-	-	-
☐	**name**	STRING(100)	NULLABLE	-	-	-	-	-
☐	**surname**	STRING(100)	NULLABLE	-	-	-	-	-
☐	**course**	STRING(100)	NULLABLE	-	-	-	-	-

EDIT SCHEMA   VIEW ROW ACCESS POLICIES

- How did you write and execute SQL queries in BigQuery?

I wrote SQL queries in the BigQuery editor. After selecting my dataset and table, I typed out the queries I wanted to run and clicked the "Run" button. I used commands to filter the data, do some aggregations, and sort the results based on what I needed to analyze.

average attendance for season:

🔍 **Untitled query**   ▶ RUN   💾 SAVE ▾   ⬇ DOWNLOAD   👤 SHARE ▾   🕐 SCHEDULE   OPEN IN ▾   ⚙ MORE ▾

```
1  select year, AVG(attendance) as average_attendance
2  from ar-final-409311.test_dataset.exercise_6
3  group by year
4  order by average_attendance
```

Query results

⬇ SAVE RESULTS ▾

JOB INFORMATION	RESULTS	CHART	JSON	EXECUTION DETAILS	EXECUTION GRAPH

Row	year ▾	average_attendance
1	2016	30154.39397703...

The pitch types and their average speed:

🔍 **Untitled query**   ▶ RUN   💾 SAVE ▾   ⬇ DOWNLOAD   👤 SHARE ▾   🕐 SCHEDULE   OPEN IN ▾   ⚙ MORE ▾

```
1  select pitchTypeDescription, AVG(pitchSpeed) as average_pitchSpeed
2  from ar-final-409311.test_dataset.exercise_6
3  group by pitchTypeDescription
4  order by average_pitchSpeed DESC
```

Press Alt+F1

Query results

⬇ SAVE RESULTS ▾   📈 EXPLO

JOB INFORMATION	RESULTS	CHART	JSON	EXECUTION DETAILS	EXECUTION GRAPH

Row	pitchTypeDescription ▾	average_pitchSpeed
1	Fastball	92.36582038947...
2	Sinker	91.22538459864...
3	Cutter	88.08570851603...
4	Forkball	86.26754385964...
5	Pitchout	85.81976744186...
6	Splitter	84.45959934822...

Results per page: 50 ▾   1 - 14 of 14

- What insights were you able to derive from the data analysis?

From my analysis, I found a few key insights. First, the average performance score was highest in the IT department, suggesting they might have better support or engagement. Second, employees with more training hours had higher satisfaction scores, showing that training is important for keeping employees happy. Finally, many employees working significant overtime were in high-demand roles, which might indicate they are overloaded. This suggests that the company should consider balancing workloads to prevent burnout.