

Tutorial 1: Set up your account for GDSC 2024

Welcome to the first tutorial of the 2024 GDSC: The Grade-AI Generation! The tutorials will teach you all the necessary steps to participate in (and hopefully win) the challenge. This first tutorial covers the (boring) groundwork that we need to cover before we can jump into the AI parts. It explains how to create an account, sign in, create and join a team and how to access AWS.

But before we jump in make sure you join the [GDSC Teams channel](#) for all updates from the organisation team and meeting and connecting with other participants. This is also the place to ask questions if you are stuck.

Here is a quick overview of what we'll cover:

- [1. Signing up and logging to the website](#)
- [2. Creating and joining a team](#)
- [3. Setting up your AWS accounts](#)
- [4. How to use AWS
 - \[4.1. Accessing GenAI Models with AWS Bedrock\]\(#\)
 - \[4.2. Managing source code with AWS CodeCommit\]\(#\)
 - \[4.3. Developing AI solutions with Sagemaker\]\(#\)
 - \[4.4. Keeping track of the costs\]\(#\)](#)
- [5. Running the tutorials](#)
- [6. Conclusion](#)
- [7. Appendix - Where to develop your solution
 - \[7.1. With AWS Sagemaker\]\(#\)
 - \[7.2. On your laptop\]\(#\)](#)

1. Signing up and logging to the website:

The very first thing you need to do to participate in the GDSC is to sign up! Here's how:

1. Navigate to the GDSC Portal's Sign-up page - <https://gdsc.ce.capgemini.com/app/signup/>. Make sure to enter your Capgemini email. The full name is not mandatory to sign up, but it is necessary to receive

a certificate of completion after the challenge.

Sign Up

Please create your account

Full Name e.g.'Micky Mouse'

E-Mail *

Password *

Confirm Password *

Country

None

Skill level

Novice

Affiliation

Other

Certification

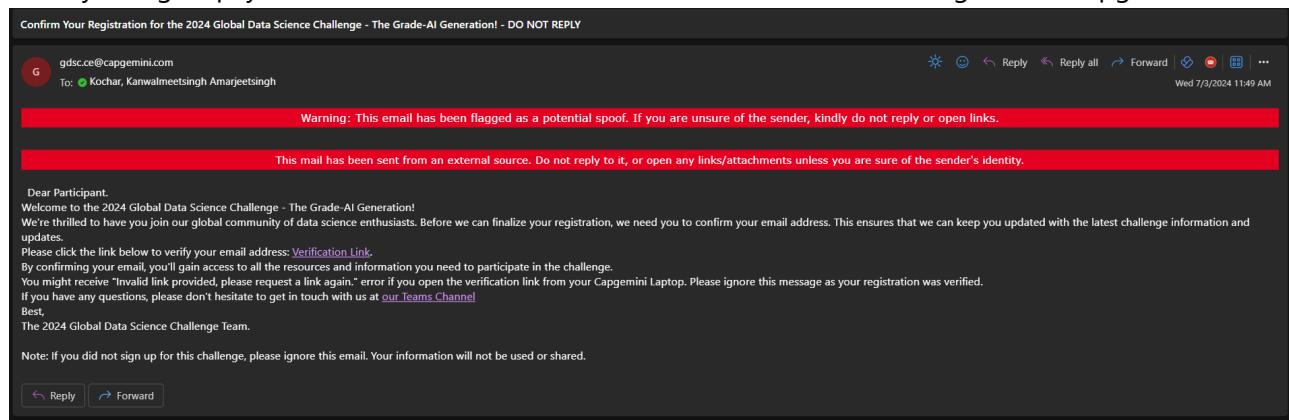
None

I accept the [Website T&C](#) & [Participation T&C](#)

Sign up



2. Once you Sign-up, you will receive an e-mail like the one shown below from gdsc.ce@capgemini.com.



3. Please click on the Verification link provided in the e-mail. You will receive an error such as this. But do not worry, you can now login to the website.



4. To login to the website, navigate to the login page - <https://gdsc.ce.capgemini.com/app/login/> and enter your credentials.

5. Once you login you will have access to the Resources -

<https://gdsc.ce.capgemini.com/app/portal/resources/>. The page has link to all the information and tutorials that you will be needing for the challenge. You can also access the Usecase page to learn more about the challenge. Feel free to explore the website for information about the current challenge and the past editions.

2. Creating and joining a team

The next step is to form or join a team. Every participant needs to be in a team to participate in a challenge. You have three options:

- Create a team only for yourself
- Create a team, find other people to join, share their team id with them. They will request to join and you can accept
- Request to join an existing team

We highly recommend that you form a team with other participants. This will help you to share ideas, learn from each other and have a better chance of winning the challenge. If you don't have a team, you can find other people in the GDSC Teams channel '[Looking for Team](#)'. It's best to work with people in a similar timezone and with a similar motivation. If you really want to win make sure your teammates have the same goal. If you're here to learn and have fun, make sure your teammates are on the same page. Note that there can be up to 4 people in a team.

To actually create a new team you need to

1. Go to the 'My Team' page - <https://gdsc.ce.capgemini.com/app/portal/>.
2. Enter your team name (keeping in mind that there is a prize for the 'best' name) and press the 'Create Team' button.

My Team

How to get access to AWS

1. Login to AWS using [this URL](#), and the email you used to sign up to GDSC.
2. You will receive an Email From AWS to verify your identity.
3. Once verified you can login to your account.

You are not part of any team yet. Join a team or create your own.

GDSC_ADMINNS	Create new team
Team ID	Join team

3. Once a team is created, you are assigned a Team ID, which you can share with potential team members and ask them to join your team. You are also assigned an AWS account, in this example *AneTestAccount0008*

How to get access to AWS

1. Login to AWS using [this URL](#), and the email you used to sign up to GDSC.
2. You will receive an Email From AWS to verify your identity.
3. Once verified you can login to your account. AneTestAccount0008

Team Information

Team Name: GDSC_ADMINNS
Team ID: 913c0412-b939-4487-86ed-a62e94593e5e
Team Captain: Kanwalmeetsingh-amarjeetsingh.kochar@capgemini.com
Your team has received 0 vote's

Team Members

Full name	E-Mail	Affiliation	Certification	Skill level	Country	Action
Kanwalmeetsingh Kochar	kanwalmeetsingh-amarjeetsingh.kochar@capgemini.com	Other	None	Novice	India	Delete team

Submissions

ID	Date	Autoeval Passed	Autoeval Result Reason	Rating
Your team has no submissions				

4. (Optional) To join a team, simply copy the team ID and paste it in the 'Join Team' section.

How to get access to AWS

1. Login to AWS using [this URL](#) and the email you used to sign up to GDSC.
2. You will receive an Email from AWS to verify your identity.
3. Once verified you can login to your account. AneTestAccount0007

You are not part of any team yet. Join a team or create your own.

Team name	Create new team
913c0412-b939-4487-86ed-a62e94593e5e	
Join team	

If you no longer wish to join a team, you can cancel your request.

Your request to join a team is pending approval.

Team ID: 913c0412-b939-4487-86ed-a62e94593e5e
Team Lead Email (who approves your request): kanwalmeetsingh-amarjeetsingh.kochar@capgemini.com

[Cancel Join Request](#)

5. (Optional) The team lead can approve or decline requests from new members

Team Information

Team Name: GDSC_ADMIN
Team Id: 913c0412-b939-4487-86ed-a62e94593e5e
Team Captain: kanwalmeetsingh-amarjeetsingh.kochar@capgemini.com
Your team has received 0 vote's

Team Members

Full name	E-Mail	Affiliation	Certification	Skill level	Country	Action
Kanwalmeetsingh Kochar	kanwalmeetsingh-amarjeetsingh.kochar@capgemini.com	Other	None	Novice	India	

Awaiting approvals

Email	Actions
mada-ashwini.kumar@capgemini.com	Approve Decline Join Request

X Warning: if you delete the team, all members, permissions and resources in your account will be destroyed

[Delete team](#)

6. (Optional) The team lead (i.e. the team creator) can remove members from the team. Keep in mind that there can be at most 4 members in a team.

Team Information

Team Name: GDSC_ADMINs
Team Id: 913c0412-b939-4487-86ed-a62e94593e5e
Team Captain: kanwalmeetsingh-amarjeetsingh.kochar@capgemini.com
Your team has received 0 vote's

Team Members

Full name	E-Mail	Affiliation	Certification	Skill level	Country	Action
fghdfgeg weddfs	mada-ashwini.kumar@capgemini.com	Other	None	Novice	None	REMOVE
Kanwalmeetsingh Kochar	kanwalmeetsingh-amarjeetsingh.kochar@capgemini.com	Other	None	Novice	India	

Warning: If you delete the team, all members, permissions and resources in your account will be destroyed

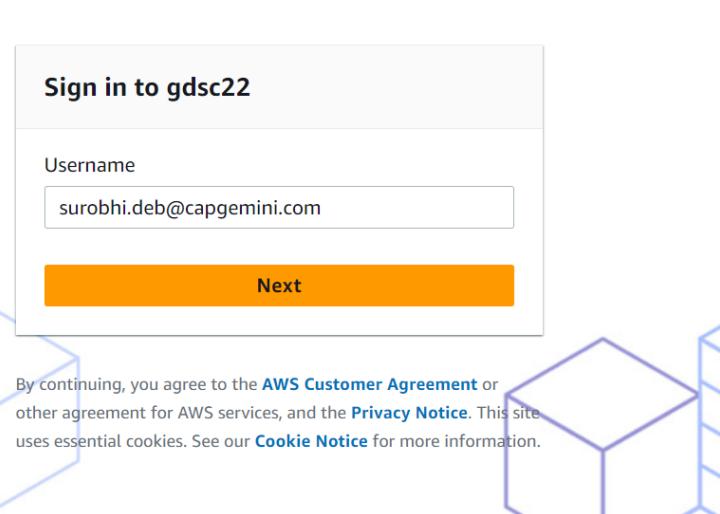
Delete team

- ⚠ Warning for the Team Leads: If you delete the team, all members, permissions and resources in your AWS account will be destroyed. This potentially included the code and models you have been working on. Make sure you have a backup of your code before deleting the team.

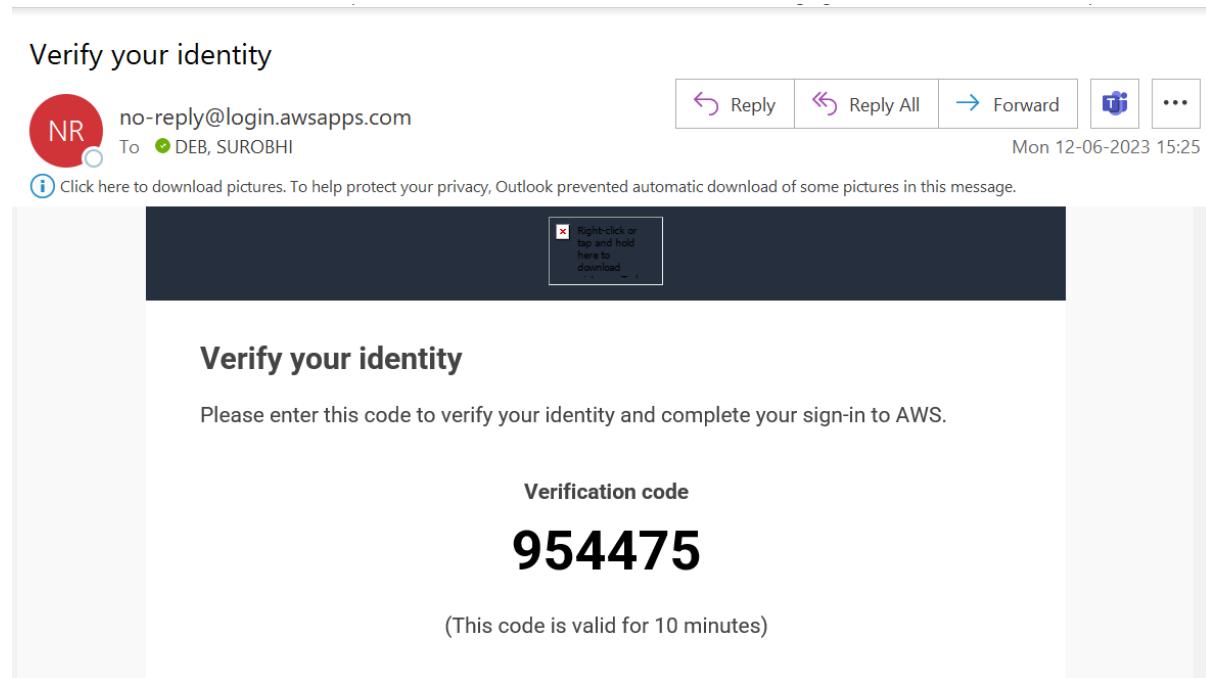
3. Setting up your AWS accounts

AWS is an integral part of the GDSC. It is where you will be developing your solution and accessing the LLM models. You will be assigned an AWS account when you create a team. Here is how you can access your AWS account:

1. Go to the [My Team](#). In the *How to get access to AWS* you can see the access link and your account name.
2. Login to your AWS account with the Capgemini email address that you used to sign up for the GDSC as the username <https://gdsc22.awsapps.com/start#/>



3. You will receive an email from *no-reply@login.awsapps.com* with the verification code. Enter it and set your password.



4. You should now be logged into AWS and see the AWS Portal with your account.

The screenshot shows the AWS Access Portal interface. At the top, it says "AWS access portal". Below that, there are tabs for "Accounts" and "Applications", with "Accounts" being the active tab. Under the "Accounts" tab, there is a section titled "AWS accounts (1441)". A search bar is present with the text "AneTestAccount0008". An orange arrow icon next to the search bar points to the account name "AneTestAccount0008" listed below it. This account entry includes the ID "058264371772" and the email "gdsc-account.de+test0008@capgemini.com". At the bottom of this section, there are links for "TeamAccess" and "Access keys".

Click on *TeamAccess* to access your account!

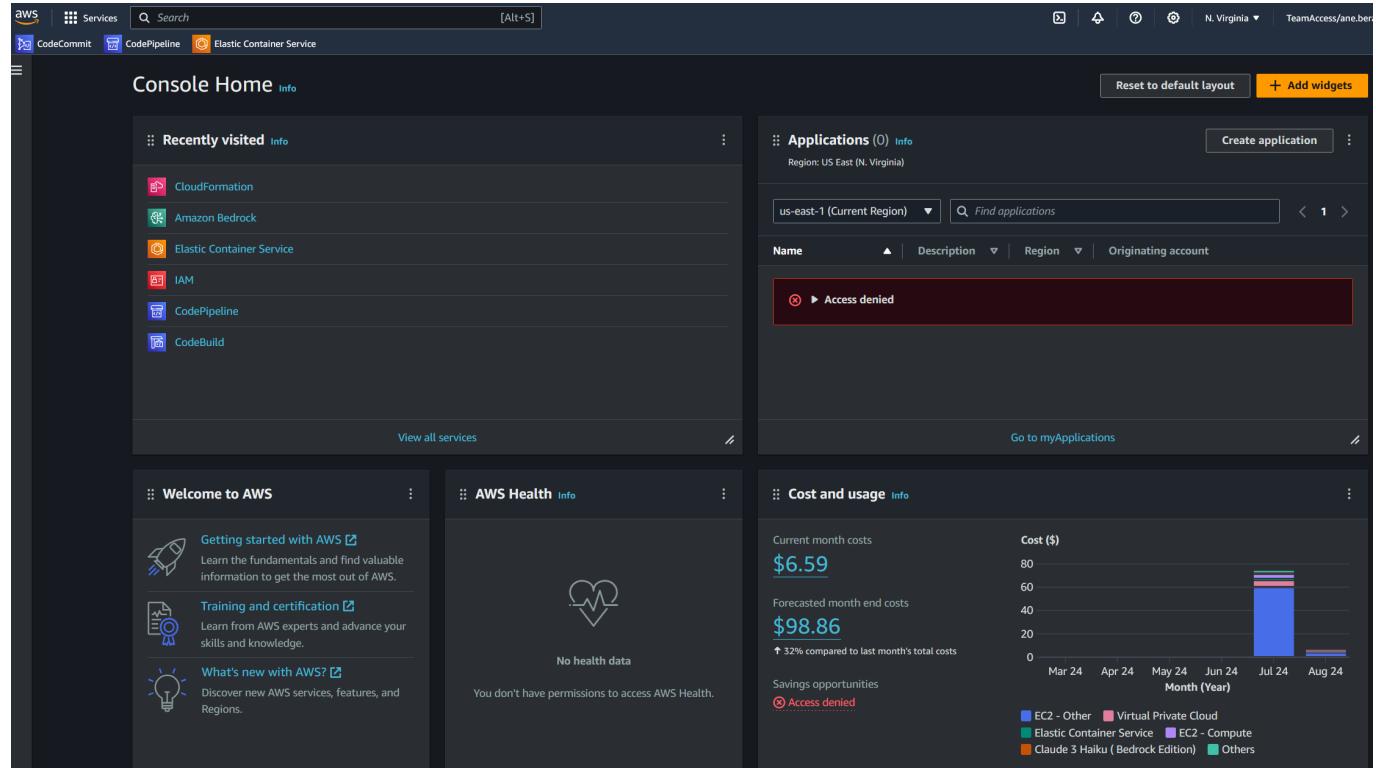
5. (Optional) If you would like to work on your laptop you need to export the credentials. Click on *Access keys* and follow the instructions. We recommend that you go with *Option 1: Set AWS environment variables*. You need to paste the instruction in your terminal and then run the command `aws sts get-caller-identity`. If you see the response shown in the screenshot, you are all set. If it says the token is expired, refresh the AWS page and copy the credentials again

```
16:16:30 ~$ aws sts get-caller-identity
{
    "UserId": "AROA2UC3F3GURJ73GZSPI:ane.berasategi@capgemini.com",
    "Account": "730335664553",
    "Arn": "arn:aws:sts::730335664553:assumed-role/AWSReservedSSO_TeamAccess_9d4878c752
om"
}
```

4. How to use AWS

Almost done! Now that you have access to AWS, let's see how you can use it for the GDSC. In this section, we will cover the main services you will be using in the GDSC: Bedrock, CodeCommit, Sagemaker and Cost Explorer.

Your screen should look like this when you log in.



But before we start, here are some general tips for using AWS:

- Make sure you are in N. Virginia region (look at the top right)
- You can change language and the theme in the settings button to the left of the region
- On the top left, you can search for services

4.1. Accessing GenAI Models with AWS Bedrock

Large Language Models (LLM) are the core of the GDSC. We'll access them via AWS Bedrock. Search for Bedrock in the search bar and click on the service. Click on the button on the top left to open the panel. Your screen should look like this:

Amazon Bedrock

The easiest way to build and scale generative AI applications with foundation models (FMs)

Try Bedrock

Get started

Base models lists all models that are available on AWS. For the GDSC, we gave you access to Claude 3 Haiku and Claude 3.5 Sonnet. You may use other models but will have to organize access and payment to them yourself. The Playgrounds are perfect for testing out the different models. The Chat interface allows you to chat with the model, similar that the openAI chatGPT website.

4.2. Managing source code with AWS CodeCommit

Source code is usually managed via a version control system like Git. AWS CodeCommit is a managed service that hosts secure Git repositories.

Search and navigate to the CodeCommit service to access the tutorials and code for the challenge. **As a change to the previous years you will also need the repository to submit a solution!**

GDSC

GDSC Info

Name

src

requirements.txt

Notify

main

Create pull request

To test your code before submitting a solution, you'll need to push it to the branch `test_submission`. To actually submit a solution you need to push the code to the branch `submission`. We'll show an example of how to do this in [tutorial 4](#).

4.3. Developing AI solutions with Sagemaker

AWS Sagemaker is a fully managed service that provides every developer and data scientist with the ability to build, train, and deploy machine learning models quickly. We'll use it to develop our AI solutions.

Like before, search for Sagemaker in the search bar and click on the service and click on **Notebooks** under **Applications and IDEs**. Your screen should look like this:

Name	Instance	Creation time	Status	Actions
Notebook-1	m1t2.medium	7/16/2024, 5:18:00 PM	Stopped	Start
Notebook-2	m1t2.medium	7/16/2024, 5:18:00 PM	Stopped	Start
Notebook-4	m1t2.medium	7/16/2024, 5:18:00 PM	Stopped	Start
Notebook-3	m1t2.medium	7/16/2024, 5:17:59 PM	Stopped	Start

You will find 4 stopped notebooks created for your team, one for each team member. Align with your team member who will work on which notebook, then press the "Start" button to start your notebook and for around 5 min until the Status is a green "InService". Under the actions column, click on "Open JupyterLab" and you should see the following screen:

JupyterLab is already connected to the GDSC CodeCommit repository, the code can be found on the left side. You are able to pull and push the code from JupyterLab.

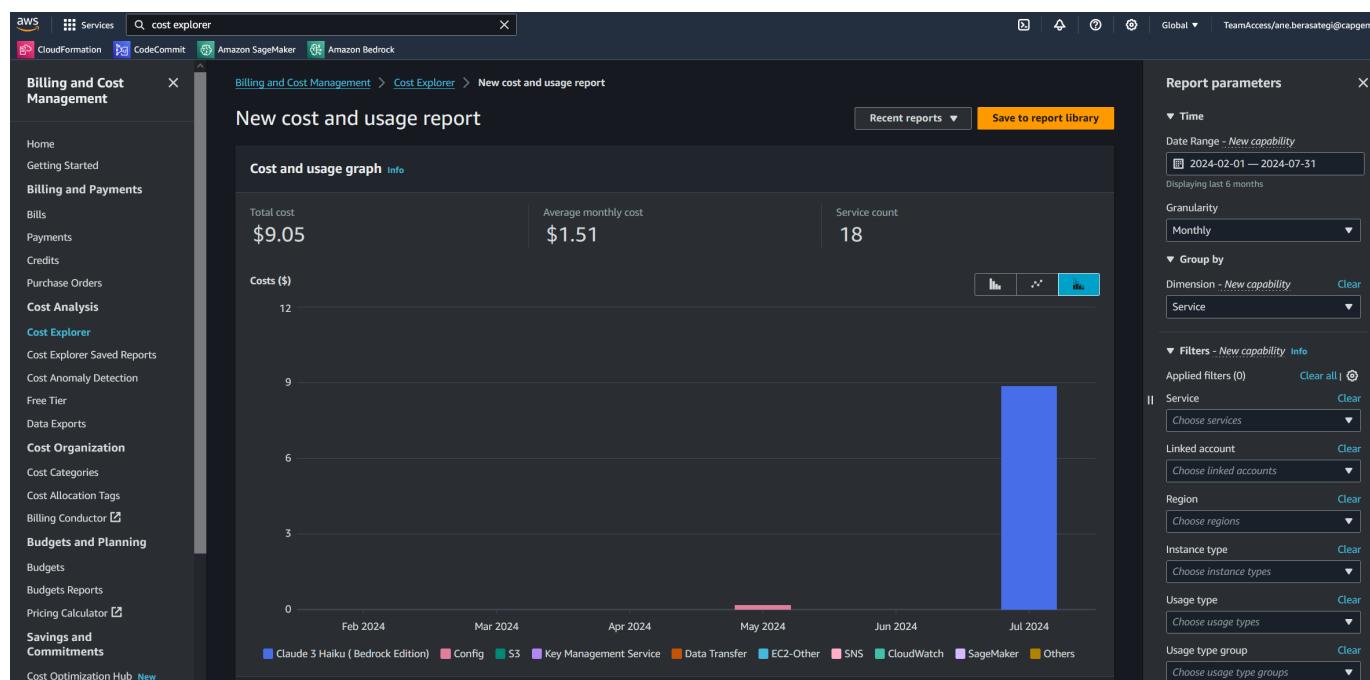
Note that you can edit the appearance of the environment in the Settings, for instance changing to dark mode or changing the language.

4.4. Keeping track of the costs

Finally, it is important to keep track of the costs of the AWS services you are using. Each team gets a fixed budget every week. Developing **and** submitting a solution costs money. **If you spent more than your budget, you will not be able to access the services anymore for the rest of the week!** Hence, budget control is essential!

You can see a summary of your current monthly costs in the AWS overview in the main page when you log in. For a more detailed view you can look at the **Billing and cost management** service. On the main page, you can see a summary of the existing costs, and forecasted costs for the current month.

The **Cost explorer** tab on the left, is the best way to get down to the details. It looks like this:

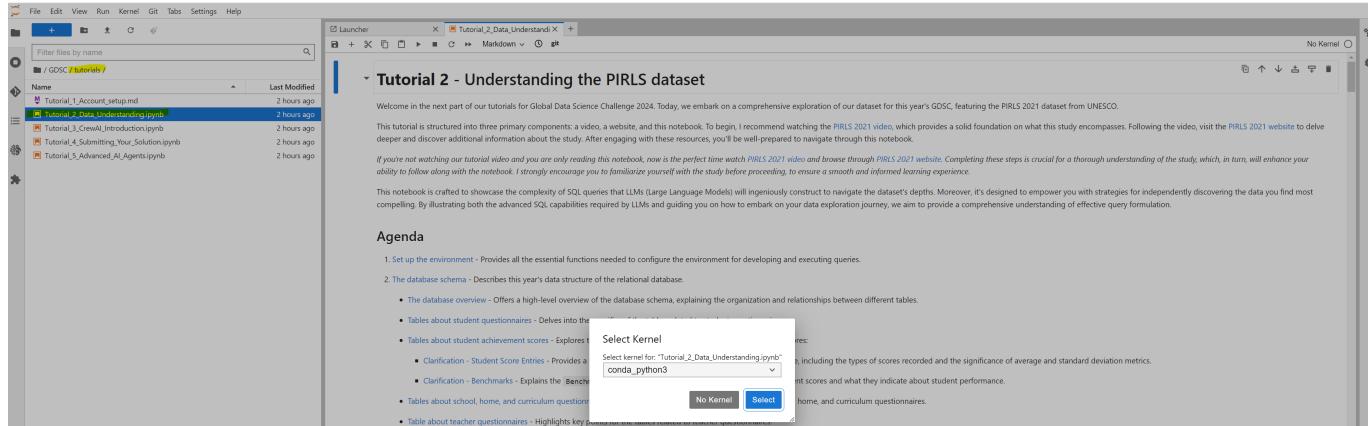


On the right side, you can filter by time range, granularity and dimension. You can see the costs for the different services you are using. You can also see the forecasted costs for the month. Keep in mind that it takes around 1 day for the costs to be updated.

5. Running the tutorials

All that's left now is to do one final check that everything is working. We'll do this by opening the second tutorials. Go back to Sagemaker, open your notebook and start JupyterLab as described in section [4.3](#).

In the left sidebar, you will see the **tutorials** folder. The first tutorial is this file. Open the second tutorial by clicking on [Tutorial_2_Data_Understanding.ipynb](#). Your screen should look like this:



In the **Select Kernel** box click on **Select**. Scroll down to the first Python cell in the **Set up the environment** section. Click in the cell and press **Shift + Enter** to run the cell. You should see the output of the cell below the cell.

Set up the environment

Firstly we install and import necessary modules.

```
[39]: # Installing necessary modules
!pip install psycopg2-binary pandas seaborn matplotlib

Requirement already satisfied: psycopg2-binary in /home/ec2-user/anaconda3/envs/python3/lib/python3.10/site-packages (2.9.9)
Requirement already satisfied: pandas in /home/ec2-user/anaconda3/envs/python3/lib/python3.10/site-packages (2.2.2)
Requirement already satisfied: seaborn in /home/ec2-user/anaconda3/envs/python3/lib/python3.10/site-packages (0.13.2)
Requirement already satisfied: matplotlib in /home/ec2-user/anaconda3/envs/python3/lib/python3.10/site-packages (3.8.4)
Requirement already satisfied: psycopg2==2.9.9 in /home/ec2-user/anaconda3/envs/python3/lib/python3.10/site-packages (from psycopg2-binary) (2.9.9)
Requirement already satisfied: numpy>=1.22.4 in /home/ec2-user/anaconda3/envs/python3/lib/python3.10/site-packages (from pandas) (1.22.4)
Requirement already satisfied: python-dateutil>=2.8.2 in /home/ec2-user/anaconda3/envs/python3/lib/python3.10/site-packages (from pandas) (2.9.0)
Requirement already satisfied: pytz>=2020.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.10/site-packages (from pandas) (2024.1)
Requirement already satisfied: tzdata>=2022.7 in /home/ec2-user/anaconda3/envs/python3/lib/python3.10/site-packages (from pandas) (2024.1)
Requirement already satisfied: contourpy>=1.0.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.10/site-packages (from matplotlib) (1.2.1)
Requirement already satisfied: cycler>=0.10 in /home/ec2-user/anaconda3/envs/python3/lib/python3.10/site-packages (from matplotlib) (0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.10/site-packages (from matplotlib) (4.53.1)
Requirement already satisfied: kiwisolver>=1.3.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.10/site-packages (from matplotlib) (1.4.5)
Requirement already satisfied: packaging>=20.0 in /home/ec2-user/anaconda3/envs/python3/lib/python3.10/site-packages (from matplotlib) (21.3)
Requirement already satisfied: pillow>=8 in /home/ec2-user/anaconda3/envs/python3/lib/python3.10/site-packages (from matplotlib) (10.4.0)
Requirement already satisfied: pyparsing>=2.3.1 in /home/ec2-user/anaconda3/envs/python3/lib/python3.10/site-packages (from matplotlib) (3.1.2)
Requirement already satisfied: six>=1.5 in /home/ec2-user/anaconda3/envs/python3/lib/python3.10/site-packages (from python-dateutil>=2.8.2->pandas) (1.16.0)
```

And that's it! You are now ready to start the second tutorial.

6. Conclusion

In this tutorial you have learnt how to sign up for the challenge and form your team. You have been assigned an AWS account, where you have explored the code, coding environment, LLM models and costs.

Remember that you can always ask for help in the [GDSC Teams channel](#) and that there is a community voted **special prize** for the team that helps other the most!

You are now ready for tutorial 2, where you will learn about the dataset for the challenge.

7 Appendix - Where to develop your solution

We already touched on this previously, but let's go into more detail. You can develop your solution in two ways: either locally on your own laptop or with AWS using AWS Sagemaker.

7.1. With AWS Sagemaker

To use AWS Sagemaker you will need to follow these steps:

- Start your notebook instance when you are ready to develop in Sagemaker as described in the [previous section](#)
- Use CodeCommit to manage your source code and submit solutions. We'll show how in [Tutorial 4](#)
- Remember to stop your instance when you are not using it to save costs. If not, it will be automatically stopped if it's not in use for one hour

7.2. On your laptop

To work on your laptop you'll need to:

- Set up your development tools: code editor, terminal, git, aws cli library
- Export your AWS credentials as explained in the [section 3.5](#) so you are authenticated to AWS
- Pull the code from the account's CodeCommit to your computer and push it when you are ready to test or submit