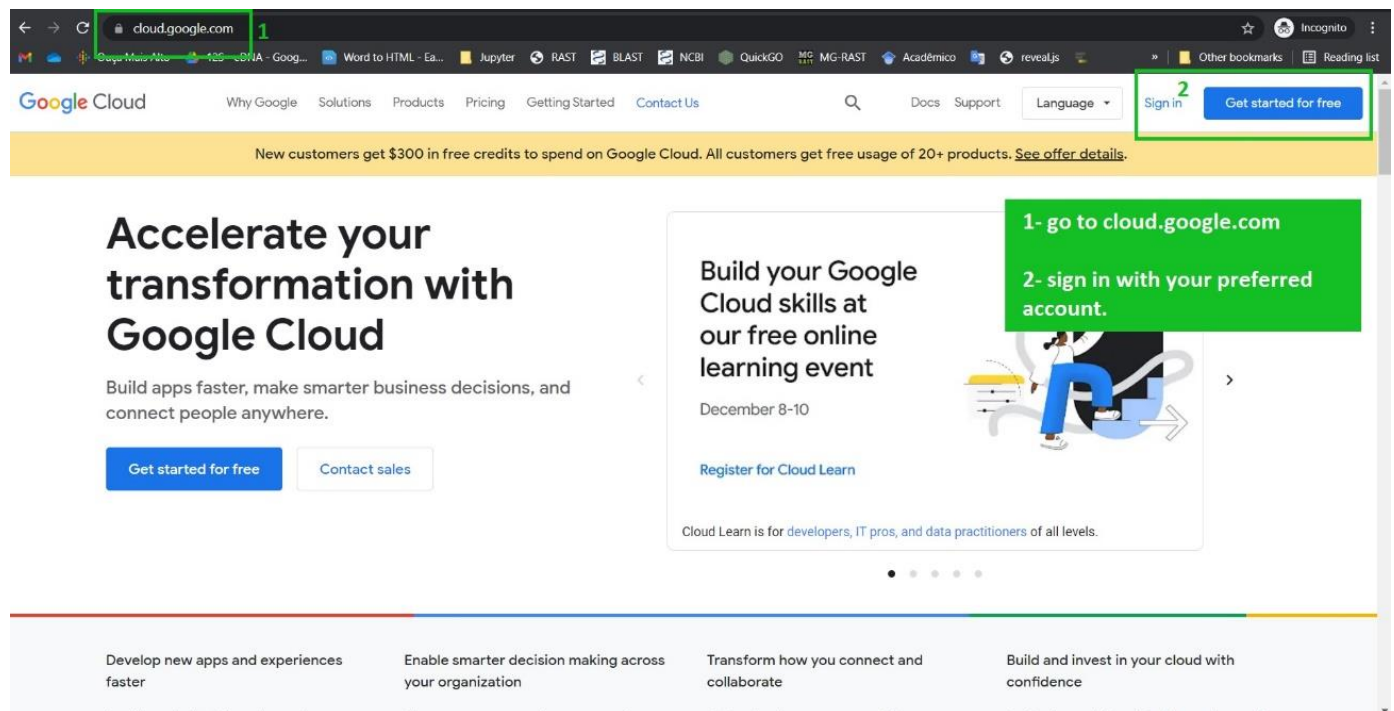


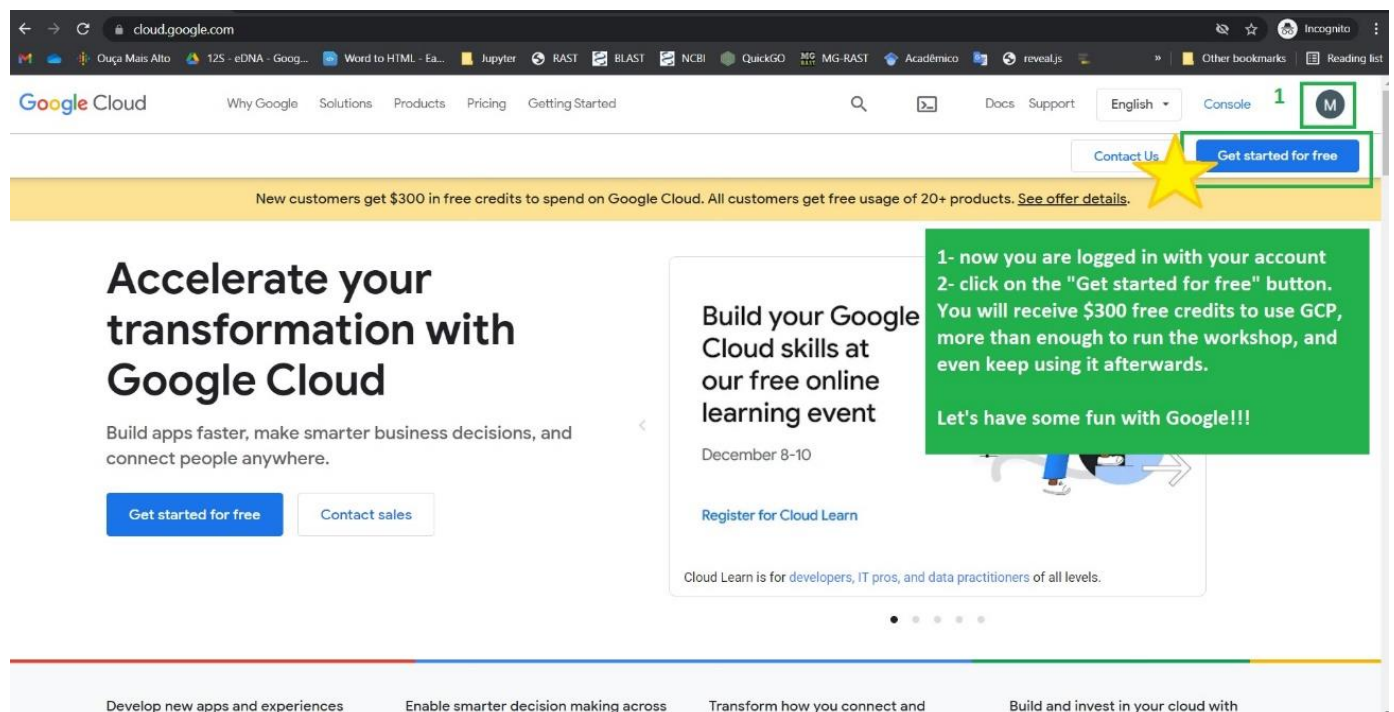
# Tutorial for setting up the METAPIPE Workshop 2021 environment!

1. Go to Google Cloud Platform, GCP, and follow the steps:

<https://cloud.google.com/>



2. After you are logged in, get start with GCP:




### 3. Enter the information requested:

console.cloud.google.com/freetrial/signup/tos?\_ga=2.106239107.1070328130.1639052931-2085689032.1639052931&pli=1

Ouçã Mais Alto 12S - eDNA - Goog... Word to HTML - Ea... Jupyter RAST BLAST NCBI QuickGO MG-RAST Acadêmico reveal.js Other bookmarks Reading list

### Etapa 1 de 3 Informações da conta

 [Redacted]@gmail.com [ALTERAR CONTA](#)

**País** 1- select your country.  
Brasil

**Qual opção descreve melhor sua organização ou as necessidades dela?** 2- select "school project/task"  
Please select  
Projeto escolar / tarefa

**Termos de Serviço**  
☒ Eu concordo com os [Termos de Serviço do Google Cloud Platform](#) e com os [termos de serviço de quaisquer serviços e APIs aplicáveis](#). Também li e concordo com os [Termos de Serviço do período do teste gratuito do Google Cloud Platform](#).  
Necessário para continuar

**Atualizações por e-mail**  
☐ Quero receber e-mails periódicos com notícias, atualizações de produtos e ofertas especiais do Google Cloud e do Google Cloud Partners.

**3- Accept the terms.**

**4- click on the "Continue" button.**

**CONTINUAR**

**Acesso a todos os produtos do Cloud Platform**  
Tenha tudo o que você precisa para criar e executar apps, sites e serviços, incluindo o Firebase e a API Google Maps.

**Crédito de US\$ 300 gratuito**  
Use o Google Cloud com US\$ 300 em crédito para gastar nos próximos 90 dias.

**Nenhuma cobrança automática será feita após o término do período de teste gratuito**  
Solicitamos seu cartão de crédito para ter certeza de que você não é um robô. Você não será cobrado, a menos que atualize manualmente para uma conta paga.

**Unfortunately, I couldn't change Google Chrome idiom to English.**  
1- Select your Country,  
2- Select "School Project/Task" or something similar,  
3- Accept the Terms of Service,  
4- Click on the "Continue" button.

### 4. GCP will send a code to your mobile phone:

console.cloud.google.com/freetrial/signup/verify?\_ga=2.106239107.1070328130.1639052931-2085689032.1639052931&pli=1

Ouçã Mais Alto 12S - eDNA - Goog... Word to HTML - Ea... Jupyter RAST BLAST NCBI QuickGO MG-RAST Acadêmico reveal.js

### Teste o Google Cloud gratuitamente

### Step 2 of 3 Account identity verification and contact information.

### Etapa 2 de 3 Verificação de identidade e informações de contato

Enviaremos uma mensagem de texto com um código de verificação de seis dígitos para confirmar sua identidade e como podemos entrar em contato com você sobre as soluções para ajudar na experiência com o Cloud. São cobradas taxas padrão.

Telefone  
+55 [Redacted]

**ENVIAR CÓDIGO**

**Add your mobile phone number to receive a confirmation code.**

**Then click on "Send Code"**

**Acesso a todos os produtos do Cloud Platform**  
Tenha tudo o que você precisa para criar e executar apps, sites e serviços, incluindo o Firebase e a API Google Maps.

**Crédito de US\$ 300 gratuito**  
Use o Google Cloud com US\$ 300 em crédito para gastar nos próximos 90 dias.

**Nenhuma cobrança automática será feita após o término do período de teste gratuito**  
Solicitamos seu cartão de crédito para ter certeza de que você não é um robô. Você não será cobrado, a menos que atualize manualmente para uma conta paga.

5. Enter the billing information. Probably Google will load it automatically, since you probably pay for some Google service already, like Google Drive. Pay attention to the "Billing address", it must be exactly the same as your credit card. **Don't worry about charges, this whole procedure is 100% free.**

Teste o Google Cloud gratuitamente

## Step 3 of 3 Billing Information

### Etapa 3 de 3 Verificação de informações de pagamento

Suas informações de pagamento nos ajudam a reduzir fraudes e abusos. Só haverá cobrança quando você ativar o faturamento automático.

#### Perfil para pagamentos ⓘ

Escolha o perfil para pagamentos que será associado a essa conta ou transação. As informações de um perfil para pagamentos são compartilhadas e usadas em todos os produtos do Google.

➔ [Redacted] ▼

#### Informações do cliente

##### ➔ Tipo de conta ⓘ

Pessoa física

##### 👤 Informações fiscais ⓘ

Status de informações fiscais: Pessoa física

[Redacted]

##### 📄 Nome e endereço ⓘ

[Redacted]

#### Como você fará o pagamento

##### 📅 Pagamentos automáticos

Você pagará por esse serviço apenas depois de acumular custos. O pagamento será efetuado por meio de uma cobrança automática quando você atingir o limite de faturamento ou 30 dias após o último pagamento automático, o que ocorrer primeiro.

#### Forma de pagamento ⓘ

[Redacted] ▼

As informações pessoais que você fornecer aqui serão adicionadas ao seu perfil para pagamentos. Elas serão armazenadas com segurança e tratadas de acordo com o [Política de Privacidade do Google](#).

In this page Google asks for a billing account. It is highly likely that Google has your billing information already, since it is the same for any **Google services**, like **Google Drive**.

You don't need to worry about it. Besides the \$300 free credits, after 3 months, **Google will not start a subscription automatically**, as usually occurs with other trial subscriptions.

To be **actually billed** for the GCP service, you must **"ACTIVATE"** your GCP account. If you'd like, take a look in the "Terms of service".

For you to have an idea of the cost, I am using GCP for almost 3 months, almost every weekday, and I still have more than half of the \$300 free credits available.

There's no risk\* of you being charged beyond the free credits.

\*I will show you how to **TURN OFF** your session each time you use it. This is the unique way of **wasting** your credits, forgetting to **STOP** your virtual machine each time you finish your work.

## 6. Answer some questions...

The screenshot shows the Google Cloud Platform 'Hello, Ma' onboarding screen. A survey overlay is displayed in the center, asking four questions to personalize the user's experience. The survey questions are:

- Qual opção descreve melhor sua organização ou as necessidades dela?
- Por que você veio para o Google Cloud?  
Seleção: Para saber mais / conhecer em detalhes
- Que serviços você tem interesse em utilizar no Google Cloud?
- Qual opção descreve melhor sua função?

The survey overlay has buttons for 'PRÓXIMA' (Next) and 'CONCLUÍDO' (Concluded). A green callout box on the right provides instructions:

- In this page Google will welcome you, confirming your 3 months \$300 free credits.
- It is asking why are you experiencing the GCP service. Click on the first option, "To know better/more details"
- Then click on "NEXT" button,
- It asks your interests, click on "Data Analysis". You must click on at least one.
- Then click on "NEXT" button,
- Finally it asks what your job/function, click on the "academic" Option.
- Click on the "CONCLUDE" button.

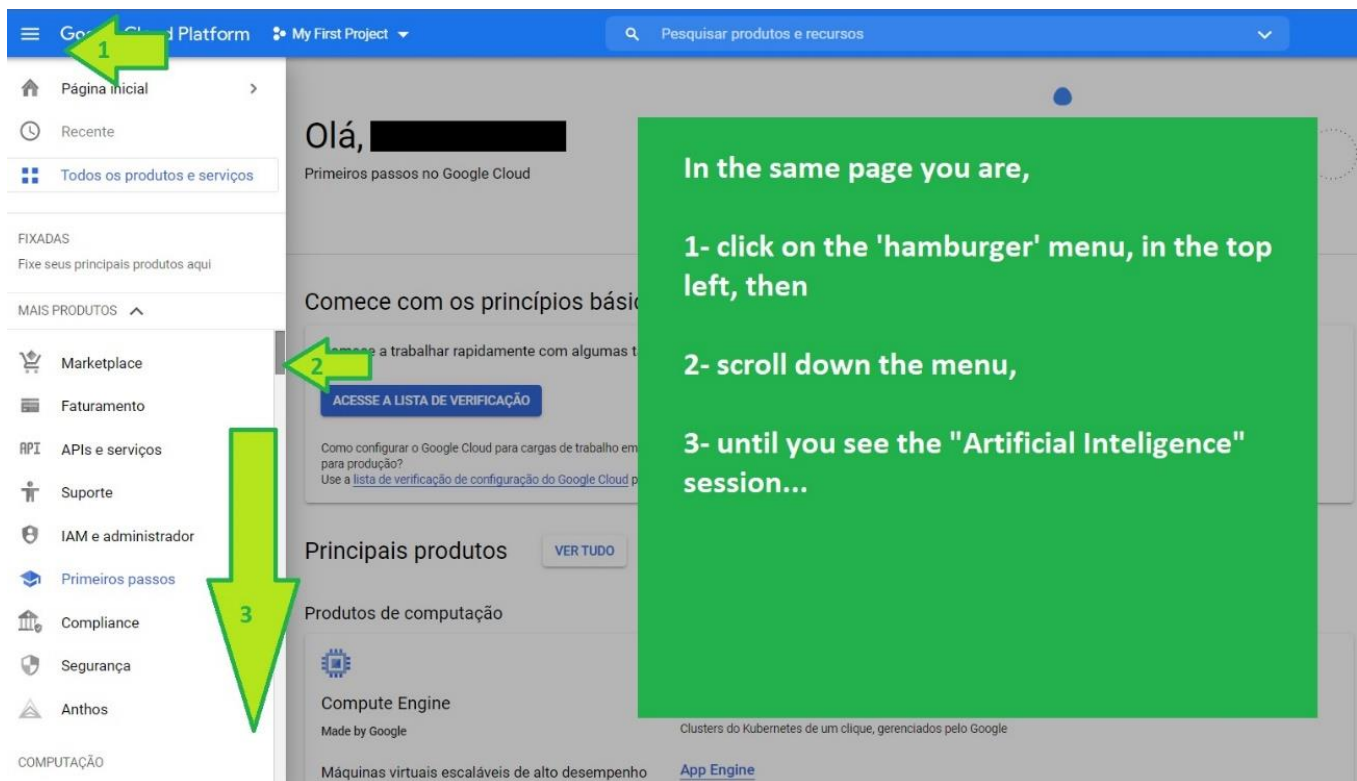
## 7. Your account in GCP was created.

The screenshot shows the Google Cloud Platform 'Hello, Ma' onboarding screen. The 'Comece com os princípios básicos' section is visible, featuring a button for 'ACESSE A LISTA DE VERIFICAÇÃO' (Access the checklist). A green callout box on the right provides instructions:

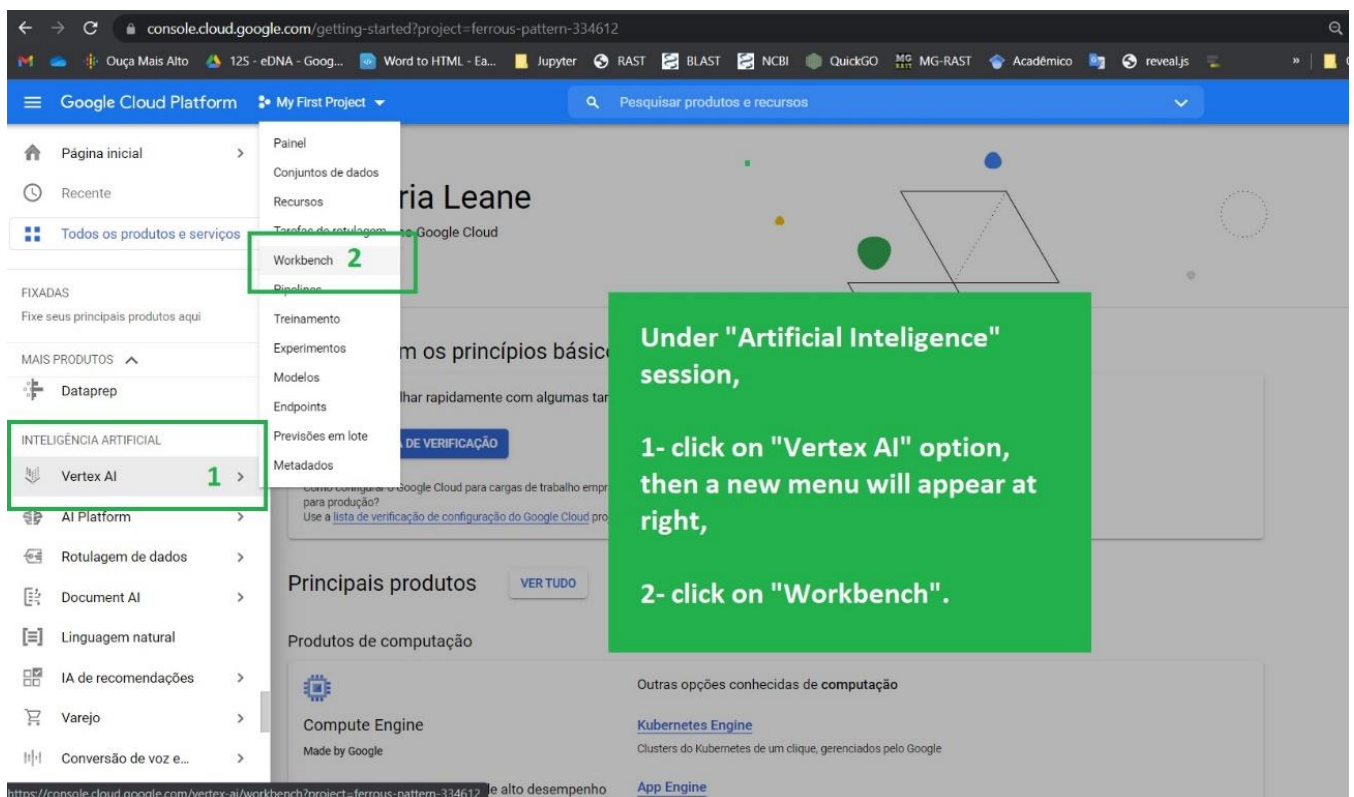
- Now we are ready to start.
- Take some minutes to read and explore this page.
- In the next step we will actually start to prepare our GCP "classroom" for the METAPIPE Workshop 2021.



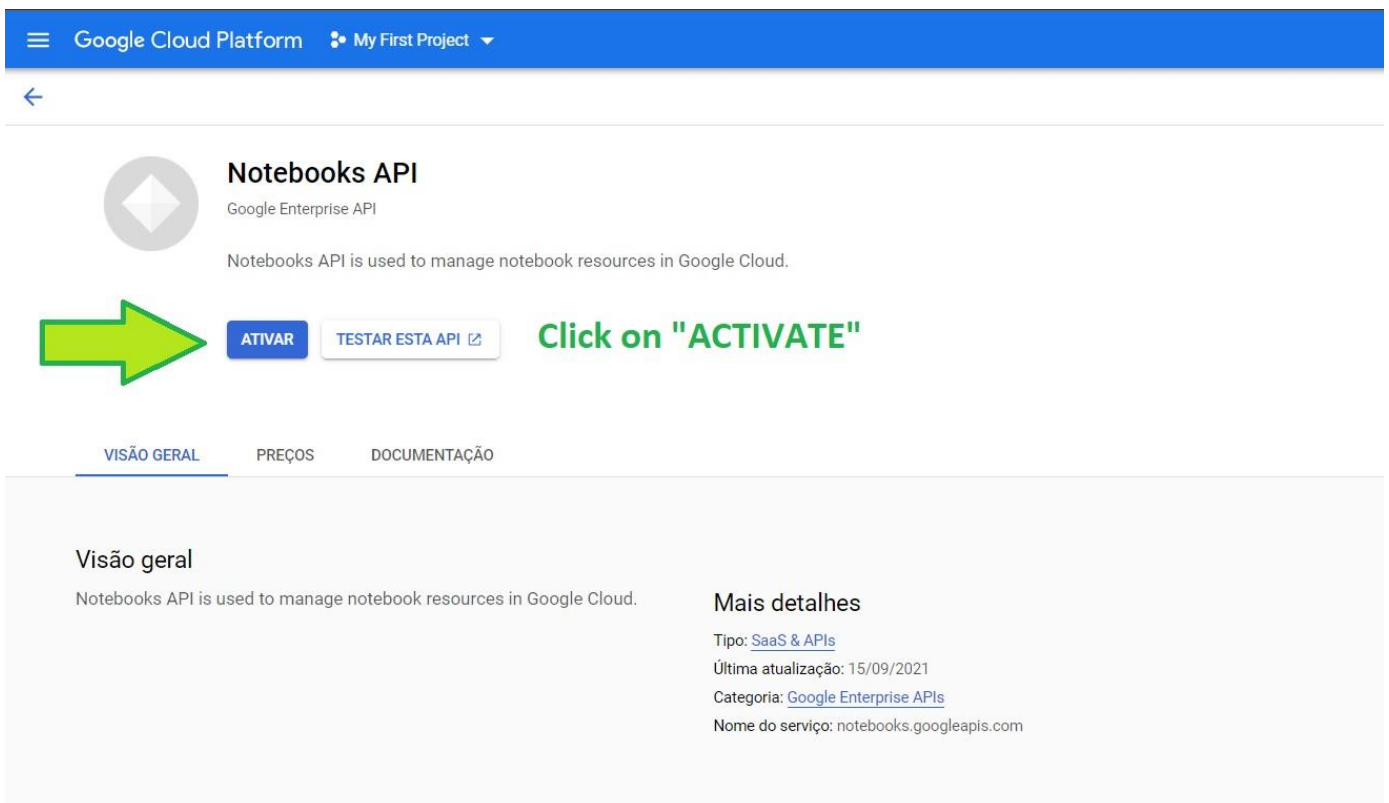
8. Now, using the top left menu, scroll down until you see “Artificial intelligence” session:



9. Click on “Vertex AI” service, then “workbench”, where the Notebooks are automatically created to run in a Default virtual machine (VM).



10. Activate the "Notebooks API". This may take some minutes.



Google Cloud Platform My First Project

## Notebooks API

Google Enterprise API

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

**Click on "ACTIVATE"**

[ATIVAR](#) [TESTAR ESTA API](#)

[VISÃO GERAL](#) [PREÇOS](#) [DOCUMENTAÇÃO](#)

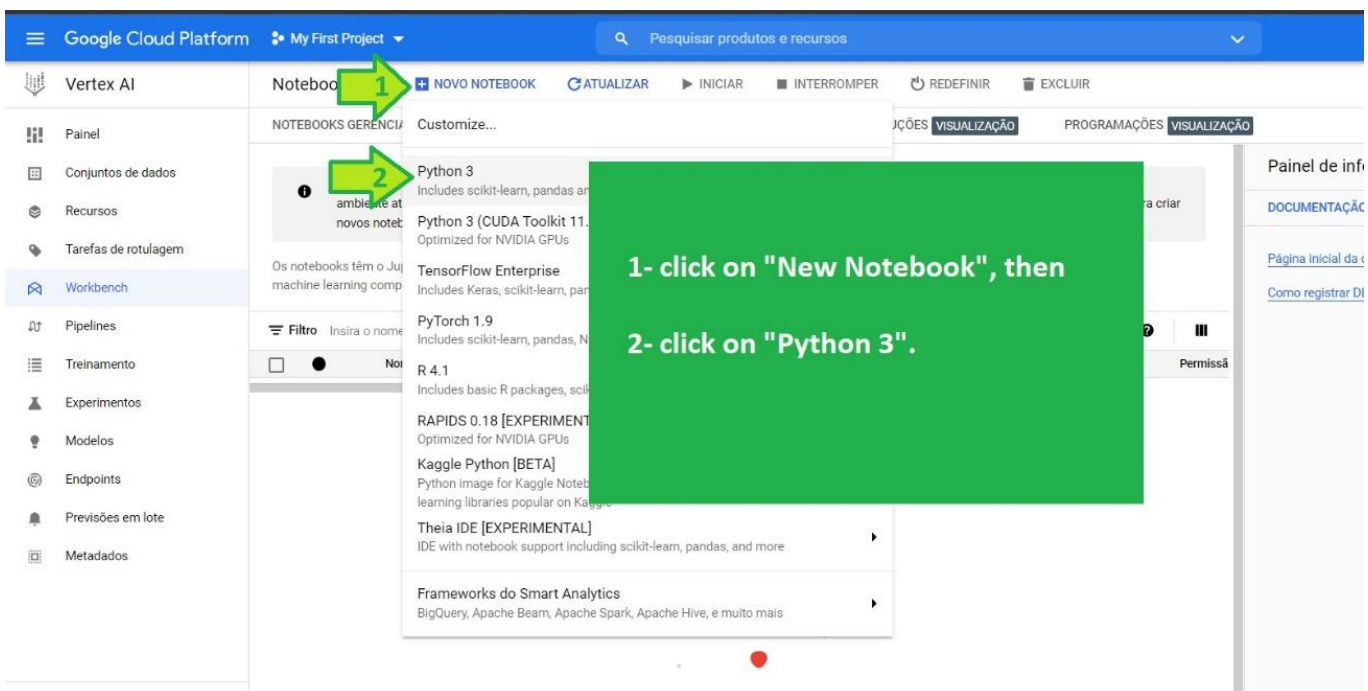
### Visão geral

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

### Mais detalhes

Tipo: [SaaS & APIs](#)  
Última atualização: 15/09/2021  
Categoria: [Google Enterprise APIs](#)  
Nome do serviço: notebooks.googleapis.com

11. Create a New Notebook, using Python 3 default environment:



Google Cloud Platform My First Project

Pesquisar produtos e recursos

## Notebooks

[NOVO NOTEBOOK](#) [ATUALIZAR](#) [INICIAR](#) [INTERROMPER](#) [REDEFINIR](#) [EXCLUIR](#)

[NOTESBOOKS GERENCIAR](#) [Customize...](#) [Visualização](#) [Programações](#) [Visualização](#)

**1- click on "New Notebook", then 2- click on "Python 3".**

**Python 3**  
Includes scikit-learn, pandas and more

**Python 3 (CUDA Toolkit 11.0)**  
Optimized for NVIDIA GPUs

**TensorFlow Enterprise**  
Includes Keras, scikit-learn, pandas and more

**PyTorch 1.9**  
Includes scikit-learn, pandas, NumPy and more

**R 4.1**  
Includes basic R packages, scikit-learn, pandas and more

**RAPIDS 0.18 [EXPERIMENTAL]**  
Optimized for NVIDIA GPUs

**Kaggle Python [BETA]**  
Python image for Kaggle Notebooks with popular machine learning libraries popular on Kaggle

**Theia IDE [EXPERIMENTAL]**  
IDE with notebook support including scikit-learn, pandas, and more

**Frameworks do Smart Analytics**  
BigQuery, Apache Beam, Apache Spark, Apache Hive, e muito mais

12. Enter a name for your notebook, select the region (does not need to be exact, I am in South America and I have to select one of the USA servers), and then click on **"ADVANCED OPTIONS"**:

**New notebook**

Notebook name \*  
metapipe-workshop2021

63-char limit with lowercase letters, digits, or '-' only. Must start with a letter. Cannot end with a '-'.

Region \*  
us-east1 (South Carolina)

Zone \*  
us-east1-b

**Notebook properties**

Environment	Python 3 (with Intel® MKL)
Machine type	4 vCPUs, 15 GB RAM
Boot disk	100 GB Standard persistent disk
Data disk	100 GB Standard persistent disk
Subnetwork	default(10.142.0.0/20)
External IP	Ephemeral(Automatic)
Permission	Compute Engine default service account
Estimated cost	\$102.70 monthly, \$0.141 hourly

ADVANCED OPTIONS CANCEL CREATE

1- give a name to your Notebook, using lowercase, '-' and numbers,

2- select your region (para quem está no Brasil, seleciona 'east-1'),

3- click on "ADVANCED OPTIONS"

13. A new page will open, scroll down a little bit

**Create a user-managed notebook**

Notebook name \*  
metapipe-workshop2021

63-char limit with lowercase letters, digits, or '-' only. Must start with a letter. Cannot end with a '-'.

Region \*  
us-east1 (South Carolina)

Zone \*  
us-east1-b

**Environment**

All environment have the latest NVIDIA GPU libraries (CUDA, CuDNN, NCCL) and latest Intel® libraries (Intel® MKL\_DNN/MKL) ready to go, along with the latest supported drivers. Select the specific image based on the primary machine learning framework you will be using. If the library you would like to use is not listed, choose the base image which provides core packages.

Operating System \*  
Debian 10

Environment \*  
Python 3 (with Intel® MKL)

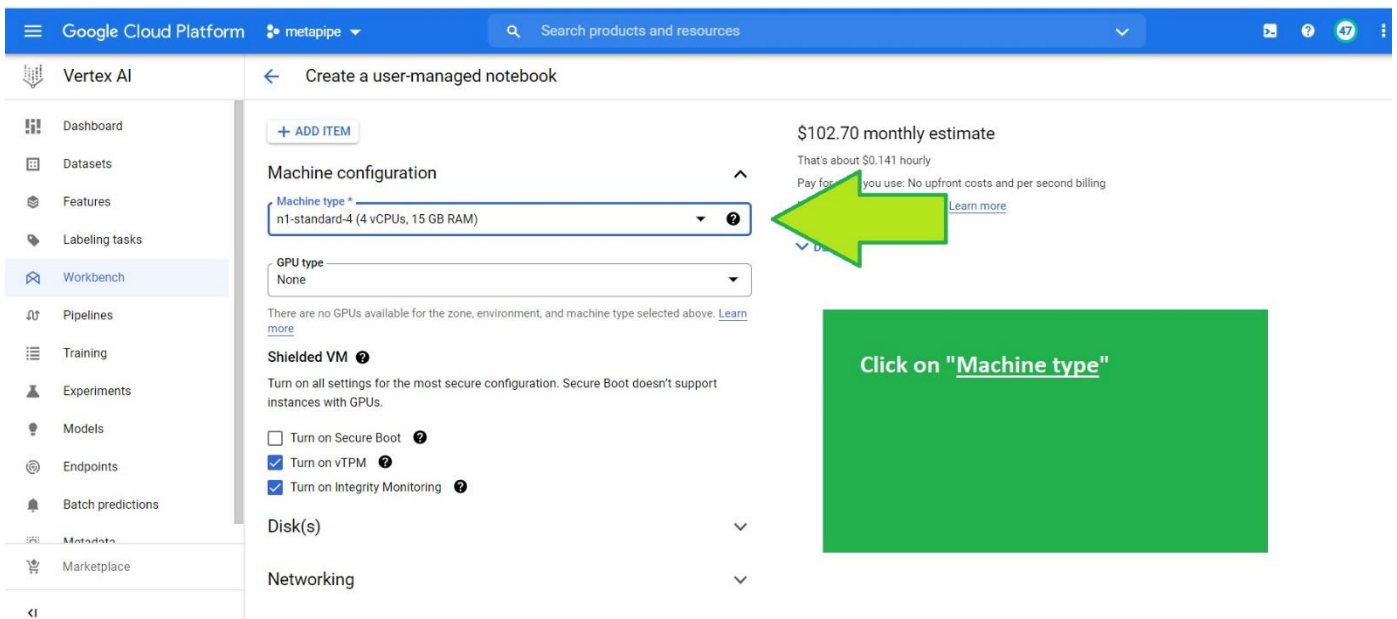
**DETAILS**

**\$102.70 monthly estimate**  
That's about \$0.141 hourly  
Pay for what you use: No upfront costs and per second billing  
Networking cost also applies. [Learn more](#)

Requests to your notebook from the Datalab/Jupyter interface may be routed through a different region than selected above depending on service availability.

This page will open, scroll down until you see  
**"Machine Configuration"**

14. Then you will see the “Machine configuration” session. Click on “Machine type”:



Google Cloud Platform metapipe Search products and resources

Vertex AI Create a user-managed notebook

+ ADD ITEM

Machine configuration

Machine type \*  
n1-standard-4 (4 vCPUs, 15 GB RAM)

GPU type  
None

There are no GPUs available for the zone, environment, and machine type selected above. [Learn more](#)

Shielded VM

Turn on all settings for the most secure configuration. Secure Boot doesn't support instances with GPUs.

☐ Turn on Secure Boot

☒ Turn on vTPM

☒ Turn on Integrity Monitoring

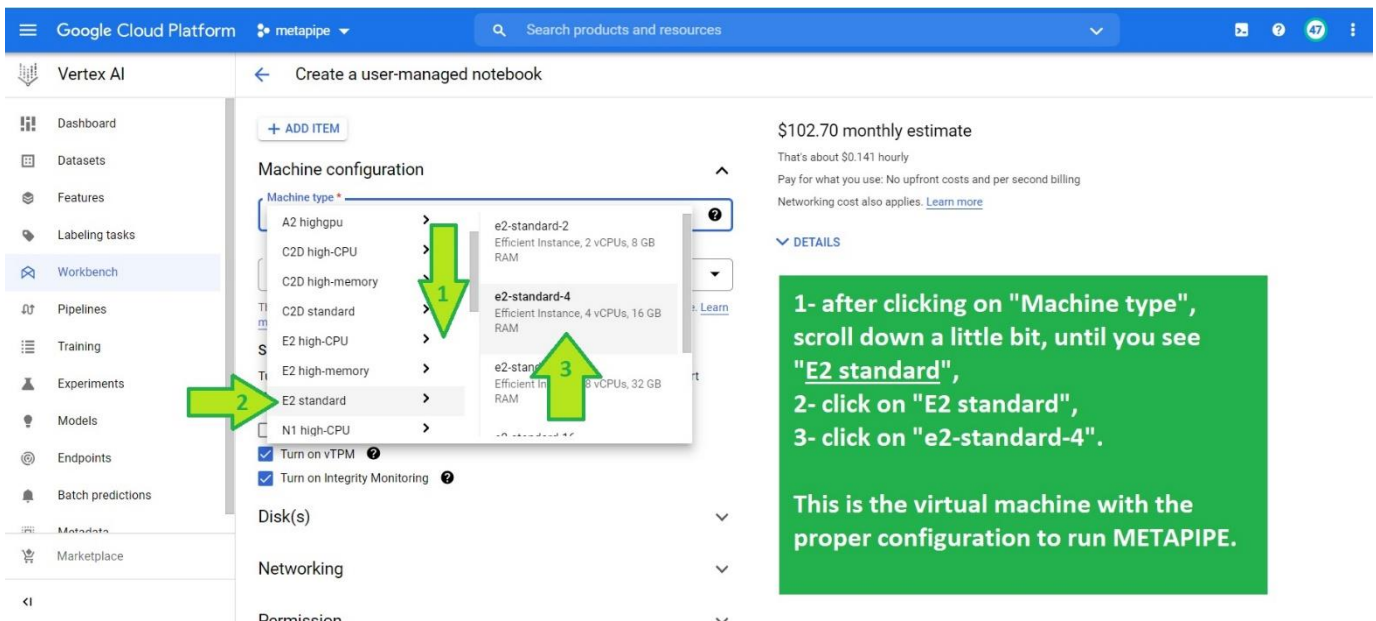
Disk(s)

Networking

\$102.70 monthly estimate  
That's about \$0.141 hourly  
Pay for what you use: No upfront costs and per second billing  
[Learn more](#)

Click on "Machine type"

15. After clicking on “Machine type”, a menu will open. Scroll down a little bit, until you see the “E2 standard” option. Click on it. In the right menu, select the machine “e2-standard-4”.



Google Cloud Platform metapipe Search products and resources

Vertex AI Create a user-managed notebook

+ ADD ITEM

Machine configuration

Machine type \*

A2 highgpu

C2D high-CPU

C2D high-memory

C2D standard

E2 high-CPU

E2 high-memory

E2 standard

N1 high-CPU

e2-standard-2  
Efficient Instance, 2 vCPUs, 8 GB RAM

e2-standard-4  
Efficient Instance, 4 vCPUs, 16 GB RAM

e2-standard-8  
Efficient Instance, 8 vCPUs, 32 GB RAM

☒ Turn on vTPM

☒ Turn on Integrity Monitoring

Disk(s)

Networking

Permissions

\$102.70 monthly estimate  
That's about \$0.141 hourly  
Pay for what you use: No upfront costs and per second billing  
Networking cost also applies. [Learn more](#)

1- after clicking on "Machine type", scroll down a little bit, until you see "E2 standard", 2- click on "E2 standard", 3- click on "e2-standard-4".

This is the virtual machine with the proper configuration to run METAPIPE.



16. Then you will return to the “advanced options” page. Scroll down and click on “CREATE”. It will take about 2 minutes to open your new virtual machine and load the JupyterLab service, which will run our Notebooks.

Google Cloud Platform metapipe Search products and resources

Vertex AI Create a user-managed notebook

**Shielded VM**   
 Turn on all settings for the most secure configuration. Secure Boot doesn't support instances with GPUs.   
 ☐ Turn on Secure Boot   
 ☒ Turn on vTPM   
 ☒ Turn on Integrity Monitoring

**Disk(s)**   
 **Networking**   
 **Permission**   
 **Security**   
 **Environment**   
 **Enter environment upgrade and system health**

**\$105.85 monthly estimate**   
 That's about \$0.145 hourly   
 Pay for what you use: No upfront costs and per second billing   
 Networking cost also applies. [Learn more](#)

**DETAILS**

**CREATE** **CANCEL**

After selection the "e2-standard-4" machine,   
 1- scroll down, and then   
 2- clicl on "CREATE".

17. When ready, the “OPEN JUPYTERLAB” button will appear. Click on it to start.

Google Cloud Platform My First Project Pesquisar produtos e recursos

Vertex AI Notebooks **NOVO NOTEBOOK** **ATUALIZAR** **INICIAR** **INTERROMPER** **REDEFINIR** **EXCLUIR** **OCULTAR PAINEL DE INFORMAÇÕES**

NOTEBOOKS GERENCIADOS **VISUALIZAÇÃO** NOTEBOOKS GERENCIADOS PELO USUÁRIO EXECUÇÕES **VISUALIZAÇÃO** PROGRAMAÇÕES **VISUALIZAÇÃO**

**1** A partir da versão DLVM M80, todos os ambientes incluirão o JupyterLab 3.x por padrão. Para continuar usando a versão do JupyterLab 1.x de um ambiente atual, desative o upgrade automático (se ativado) e não faça upgrade do ambiente manualmente para uma nova versão do ambiente. Para criar novos notebooks com o JupyterLab 1.x instalado, consulte como criar versões específicas de notebooks.

Os notebooks têm o JupyterLab pré-instalado e são configurados com frameworks de machine learning compatíveis com GPU. [Learn more](#)

**Filtro** Insira o nome ou o valor da propriedade

	Nome do notebook	Zona	Upgrade automático	Ambiente	Tipo de máquina
<input type="checkbox"/>	metapipe-workshop2021	us-east1-b	—	NumPy/SciPy/scikit-learn	4 vCPUs, 15 GB RAM

**ABRIR JUPYTERLAB**

Click on "Open JupyterLab"

**Painel de informações**   
 DOCUMENTAÇÃO RÓTULOS   
 [Página inicial da documentação](#)   
 [Como registrar DLVMs legadas](#)

18. Select the default Python 3 environment.

File Edit View Run Kernel Git Tabs Settings Help

Filter files by name: /

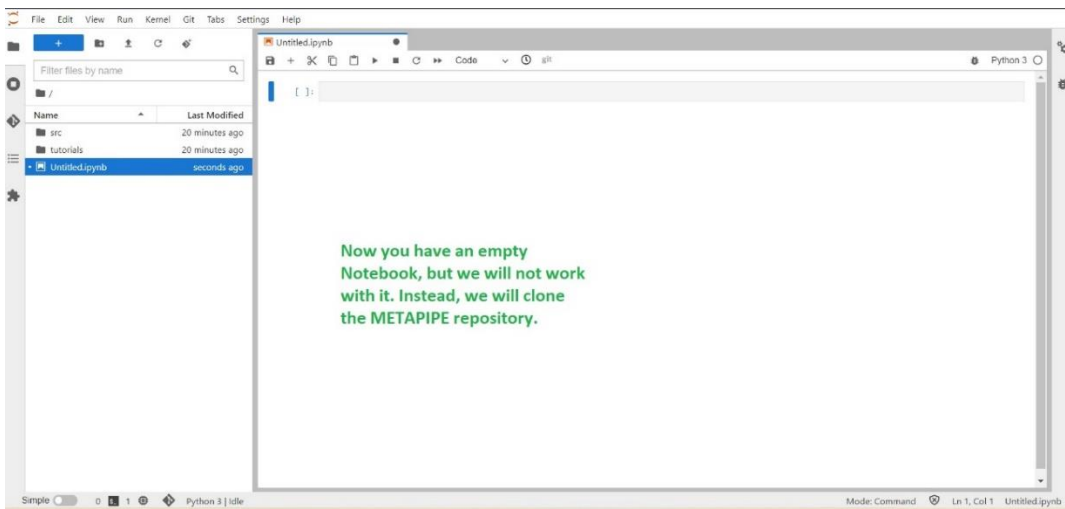
**Launcher**

**Notebook**   
 Python 3 Python (conda envroot) \*

**Console**   
 Python 3 Python (conda envroot) \*

**Other**   
 Terminal Text File Markdown File Python File Show Contextual Help

19. Now you have an empty Notebook. We need to call the METAPIPE scripts and files from Github. Follow the steps:



+++++

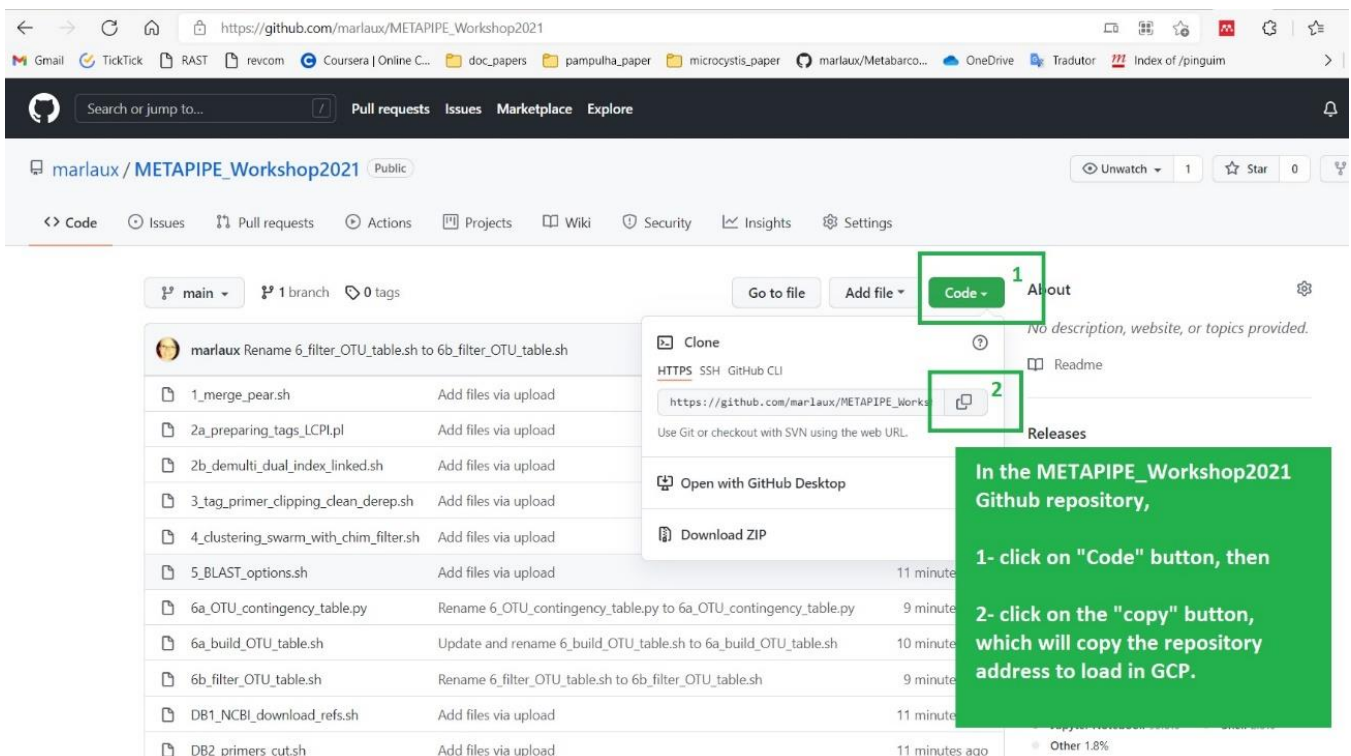
**NOW WE ARE GOING TO CLONE THE METAPIPE WORKSHOP REPOSITORY, CONTAINING ALL THE SCRIPTS, NOTEBOOKS AND FILES REQUIRED.**

+++++

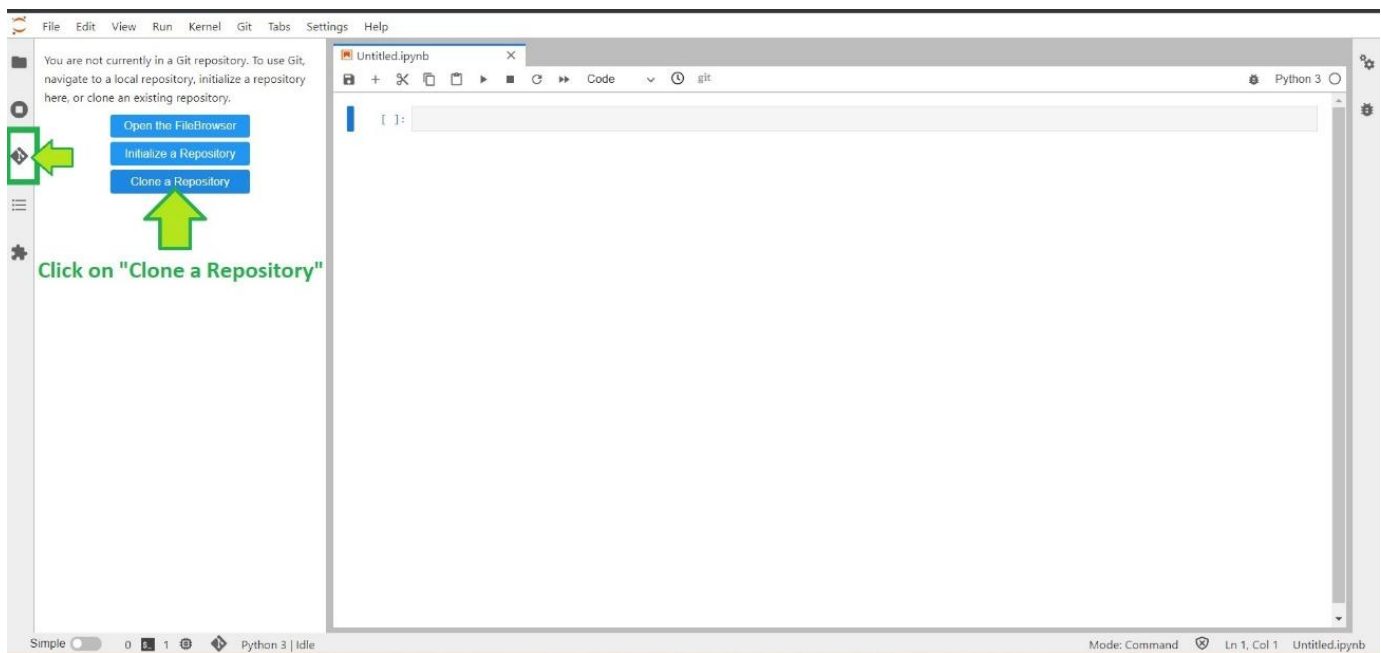
20. Go to METAPIPE Github repository:

[marlaux/METAPIPE\\_Workshop2021 \(github.com\)](https://github.com/marlaux/METAPIPE_Workshop2021)

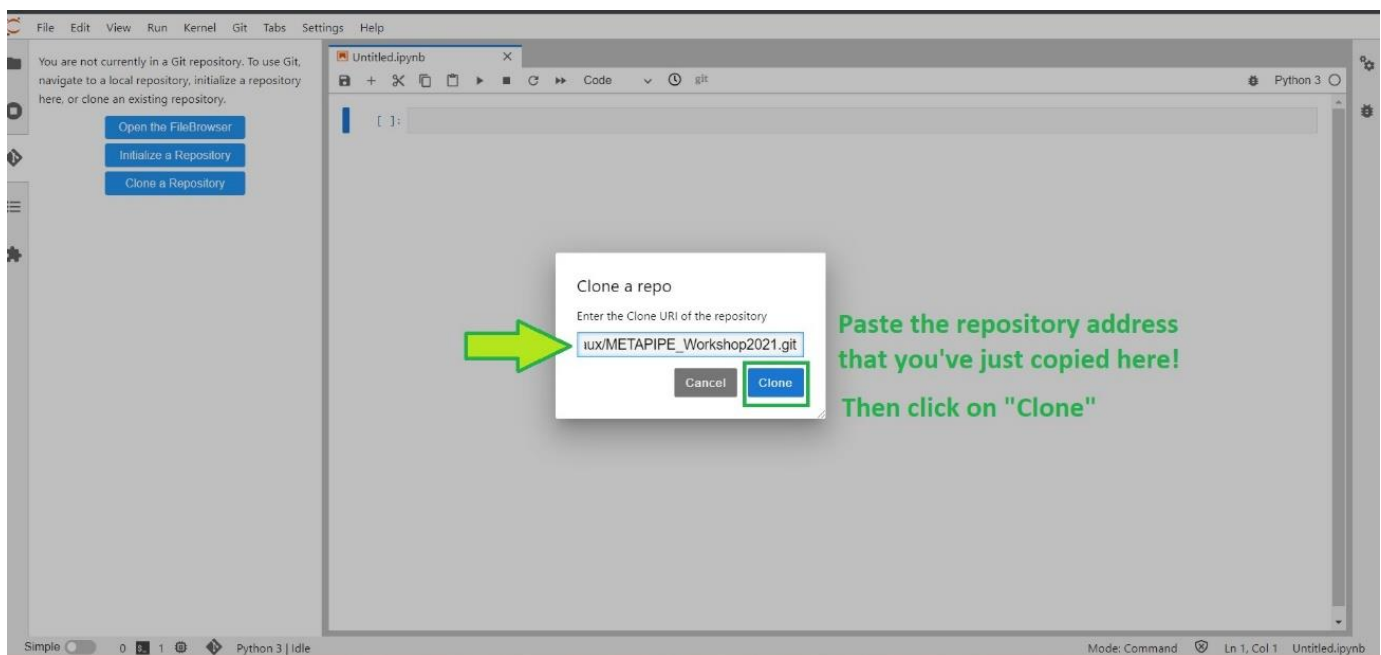
21. and copy the repository URL.



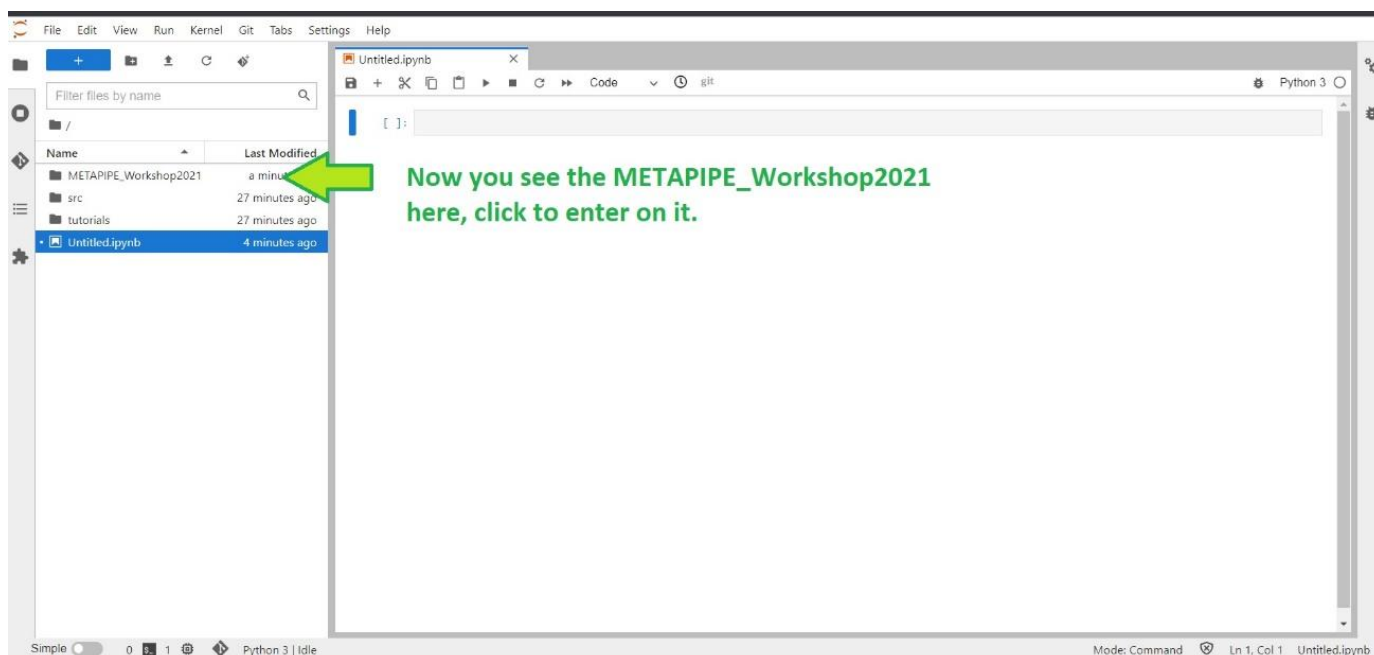
22. Go back to your Notebook. You will see a Github icon on the left, as shown in the following picture. Click on it, and then click on “Clone repository”.



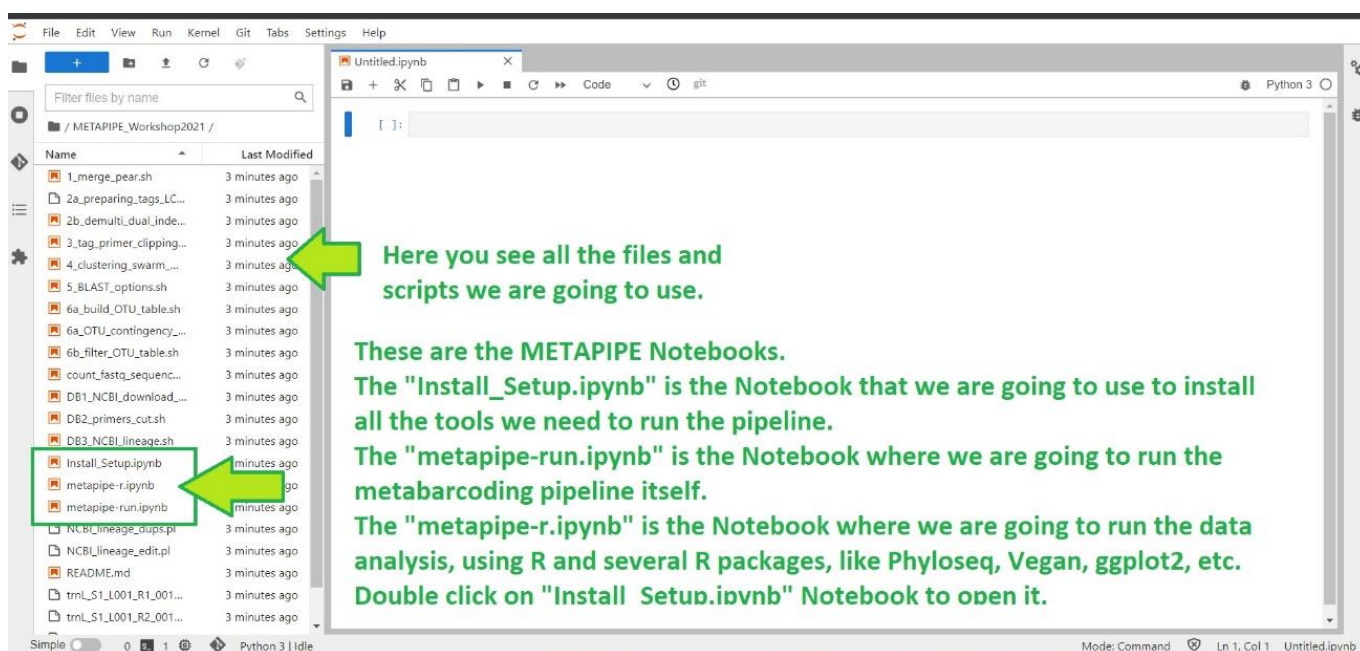
23. Paste the address you’ve just copied from Github and click on “Clone”.



24. Now the METAPIPE repository is in your VM, ready to run. Double click to open it.

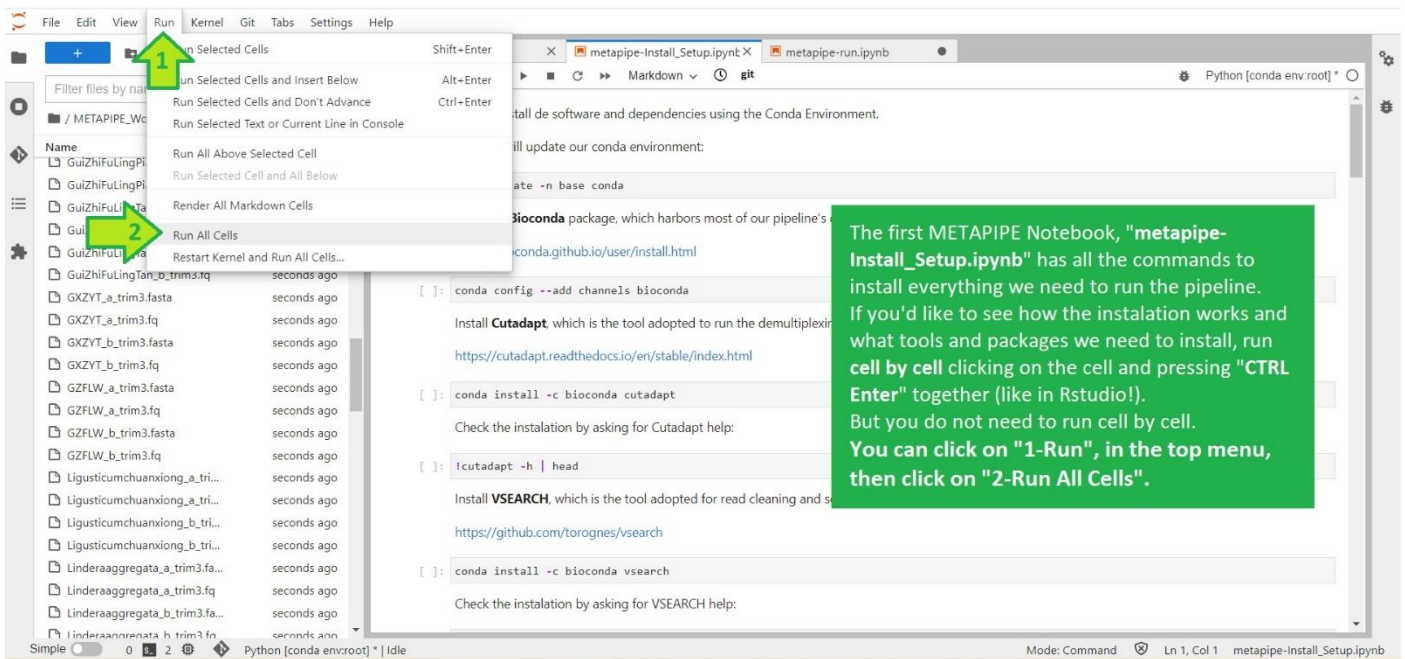


25. These are all our scripts and files. Pay attention in the description found in the following picture (Attention, the 'Install\_Setup.ipynb' Notebook was renamed to 'metapipe-Install\_Setup.ipynb'):

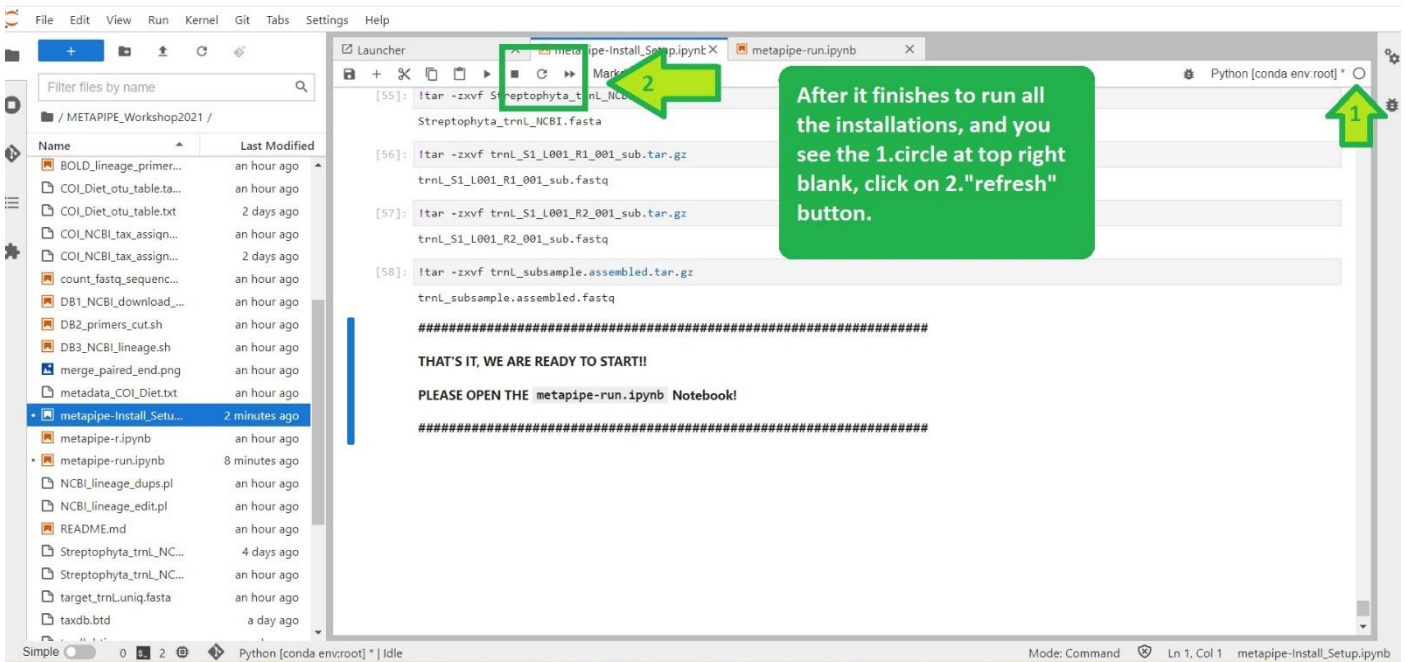




26. Double click on the “metapipe-Install\_Setup.ipynb” Notebook, the first one, which has all the commands to install everything we need to run the pipeline.



27. After it finishes to run all the installation (takes around 1 hour), you will see the little circle in the top right blank, instead of grey. This indicated that the machine is not working anymore. At this point, click on “refresh” button, indicated in the following picture:



28. Now let's turn off your virtual machine. Follow the steps: Go back to "Vertex AI Notebooks" tab in your browser:

You always **MUST** turn off your virtual machine when you finish the work. Go back to "Vertex AI Notebooks" tab.

```
[ ]: conda update -n base conda
```

Install de **Bioconda** package, which harbors most of our pipeline's

<https://bioconda.github.io/user/install.html>

```
[ ]: conda config --add channels bioconda
```

Install **Cutadapt**, which is the tool adopted to run the demultiplexir

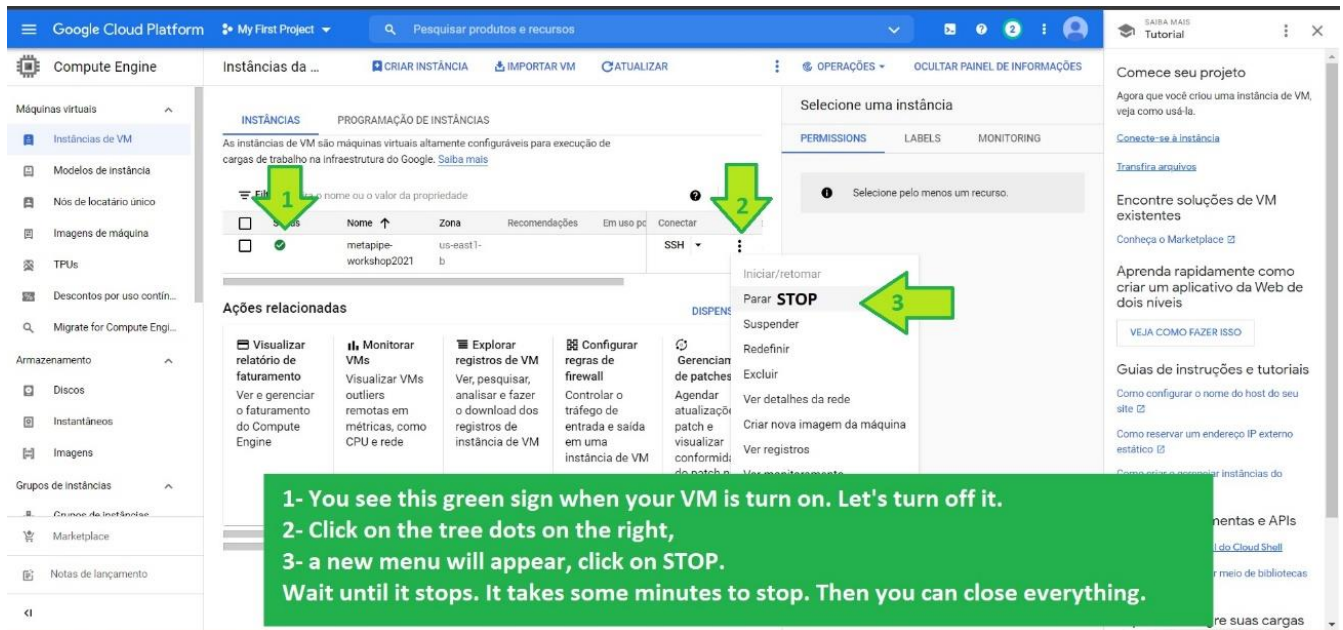
29. Click on the top left menu, scroll down until the "compute" session, then click on "Compute Engine" and "VM Instances":

1- Click on the "hamburger" menu in the top left,

2- scroll down until you see the "Compute" session, then click on "Compute Engine",

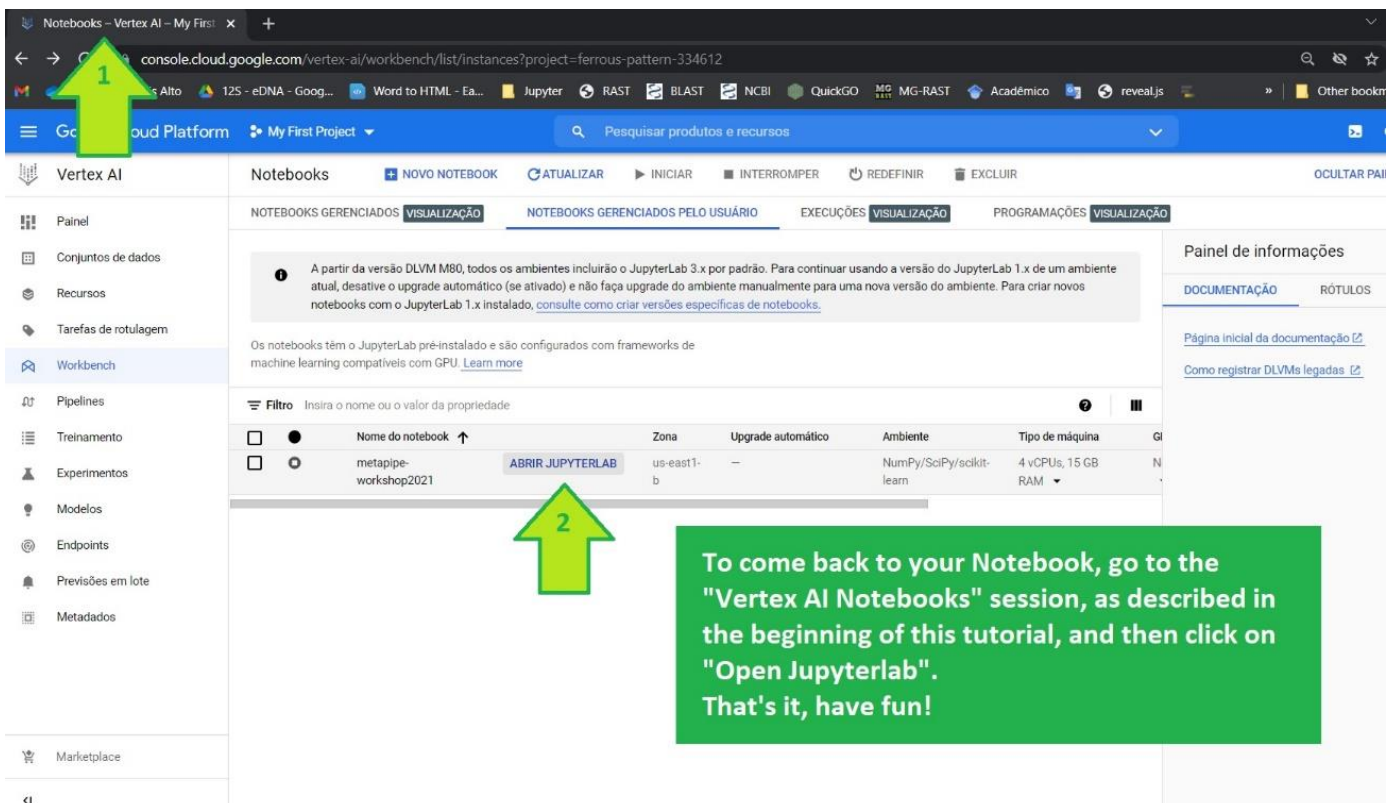
3- A new menu will open on the right. Click on "VM Instances".

30. Click on the three dots, on the right, and then click on “Stop”, and “Stop” again. That’s it, your VM is off. Wait until it finishes the stopping and close the window.



1- You see this green sign when your VM is turn on. Let's turn off it.  
2- Click on the tree dots on the right,  
3- a new menu will appear, click on STOP.  
Wait until it stops. It takes some minutes to stop. Then you can close everything.

31. To go back to your Notebook, you just need to go to the “Vertex AI Notebooks”, as described before, and then click on “Open JupyterLab”.



To come back to your Notebook, go to the "Vertex AI Notebooks" session, as described in the beginning of this tutorial, and then click on "Open Jupyterlab". That's it, have fun!

That's it 😊