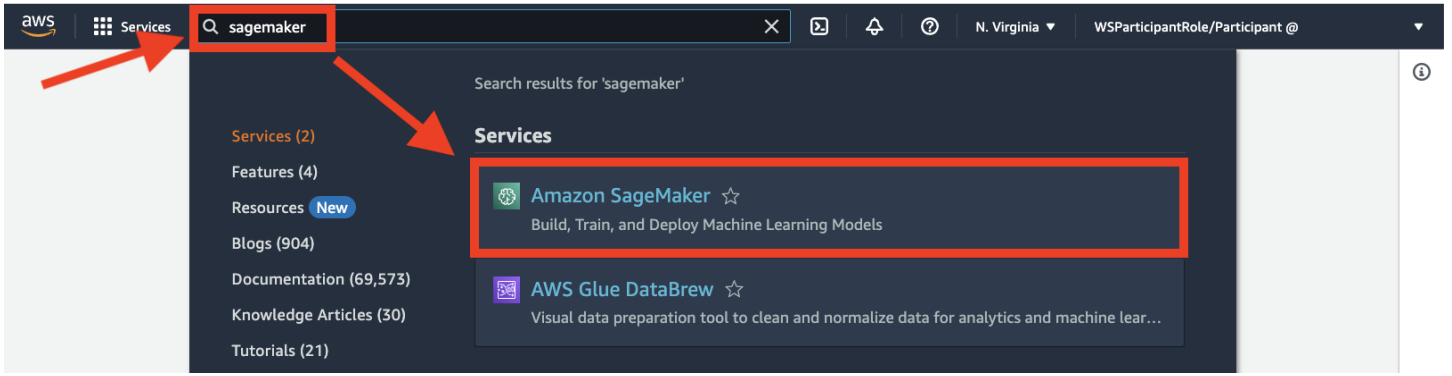
Data Science on AWS

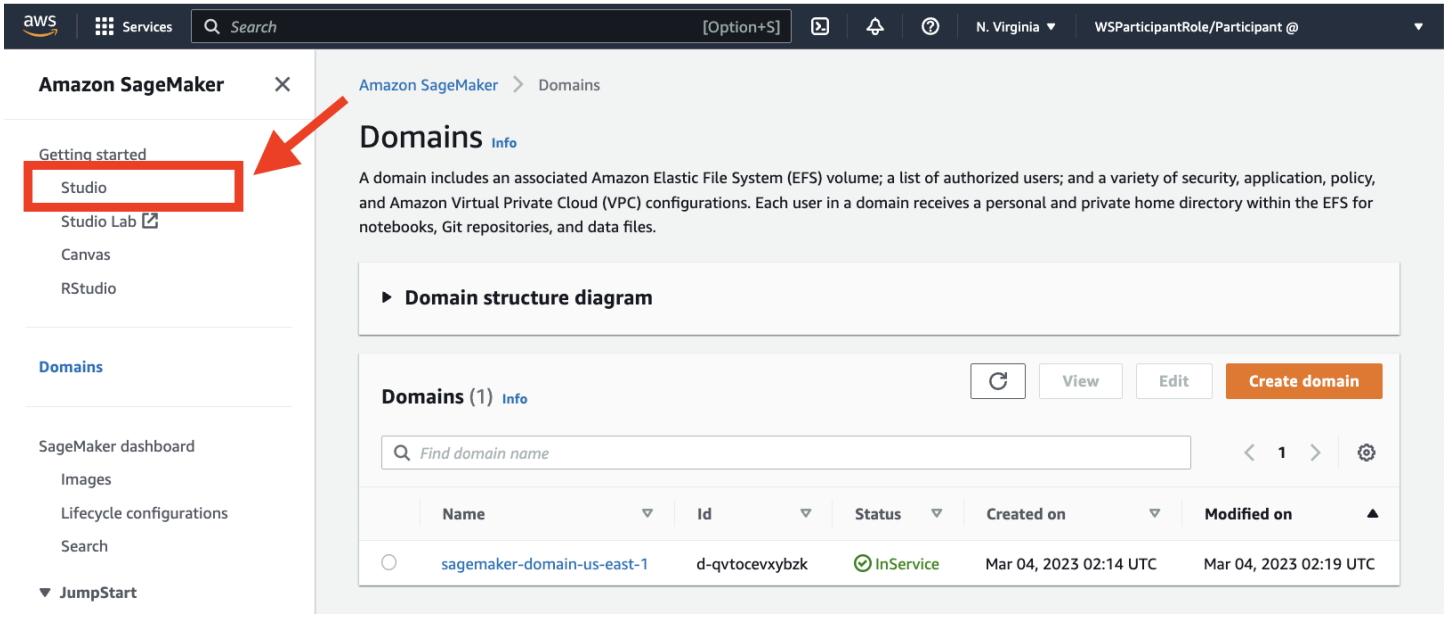
- Introduction
  - Prerequisites
  - AWS Hosted Event
  - Workshop Setup
- Part 1: Distributed Data Preprocessing
- Part 2: Fine-tuning LLMs with Amazon SageMaker
- Part 3: Automating Fine-Tuning workflows with SageMaker Pipelines
- Contributors
- Summary

Workshop Setup

1. In the AWS Management Console search bar, type "sagemaker". Then, select **Amazon SageMaker**.



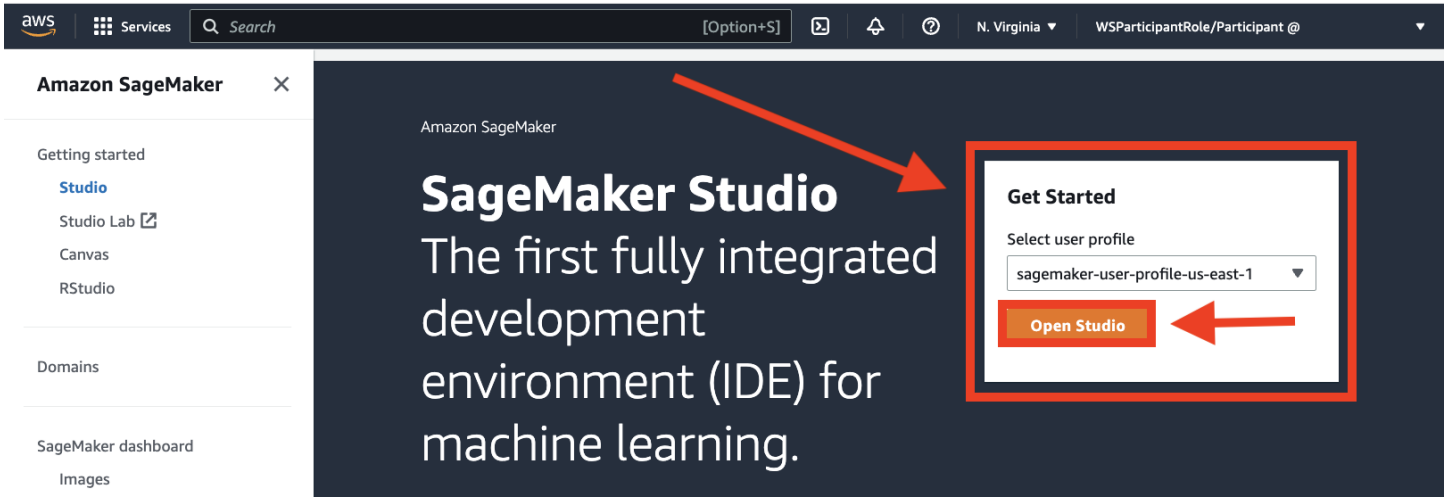
2. In the Amazon SageMaker Console, click **Studio** in the left-side navigation menu.



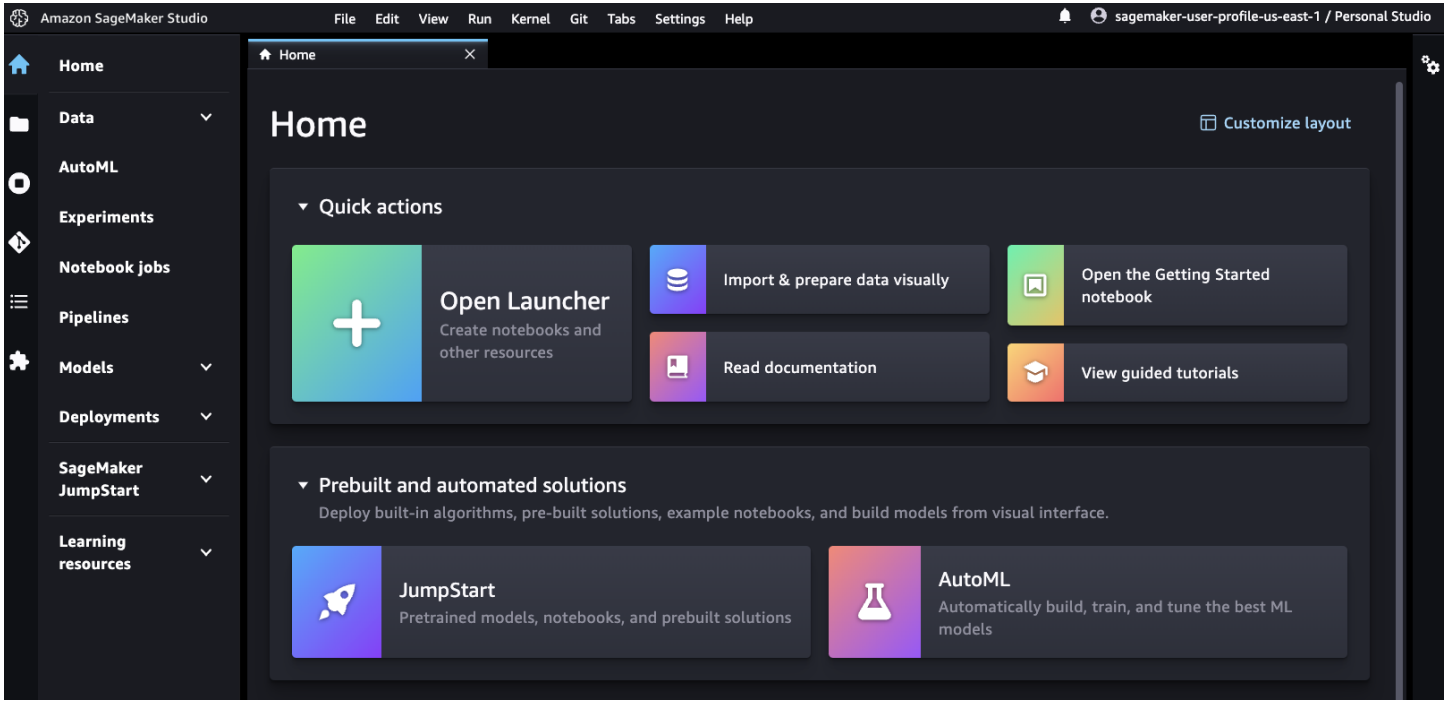
Data Science on AWS

- Introduction
  - Prerequisites
  - AWS Hosted Event
  - **Workshop Setup**
- Part 1: Distributed Data Preprocessing
- Part 2: Fine-tuning LLMs with Amazon SageMaker
- Part 3: Automating Fine-Tuning workflows with SageMaker Pipelines
- Contributors
- Summary

3. Next, click **Open Studio** with the pre-selected user profile.



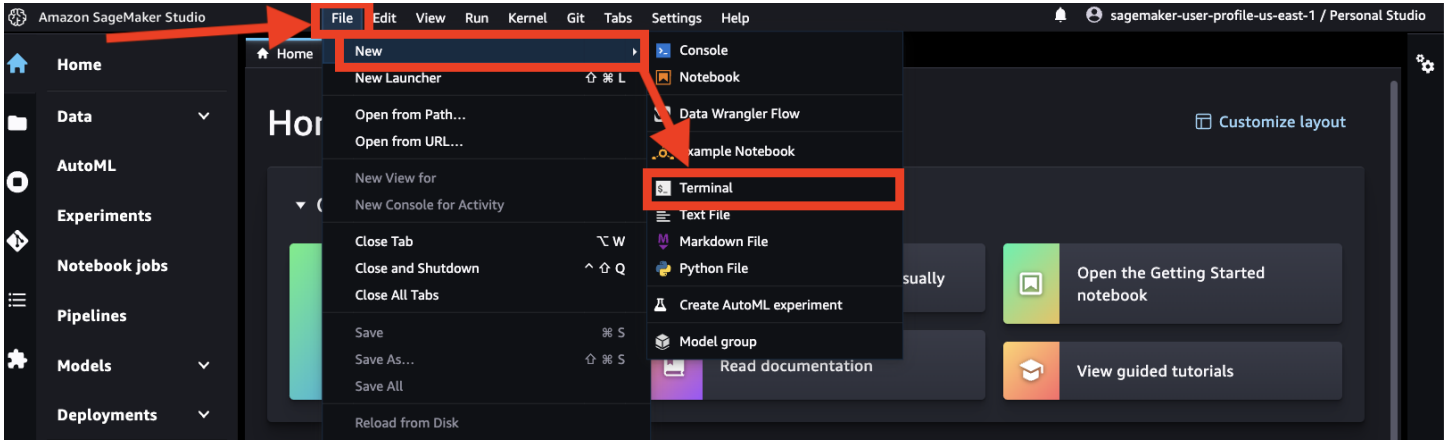
This takes you to Amazon SageMaker Studio.



4. In the top navigation bar, select **File > New > Terminal**.

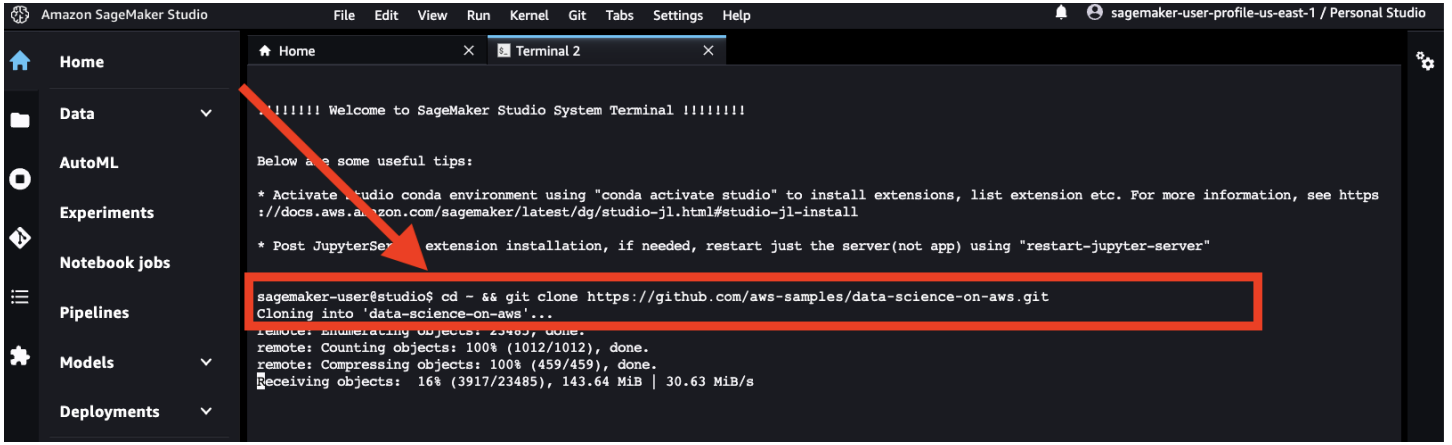
Data Science on AWS

- Introduction
  - Prerequisites
  - AWS Hosted Event
  - **Workshop Setup**
  - Part 1: Distributed Data Preprocessing
  - Part 2: Fine-tuning LLMs with Amazon SageMaker
  - Part 3: Automating Fine-Tuning workflows with SageMaker Pipelines
  - Contributors
  - Summary



5. Within the Terminal, run the following command to clone the workshop's GitHub repo:

```
cd ~ && git clone https://github.com/aws-samples/data-science-on-aws.git
```



If you see an error like the following:

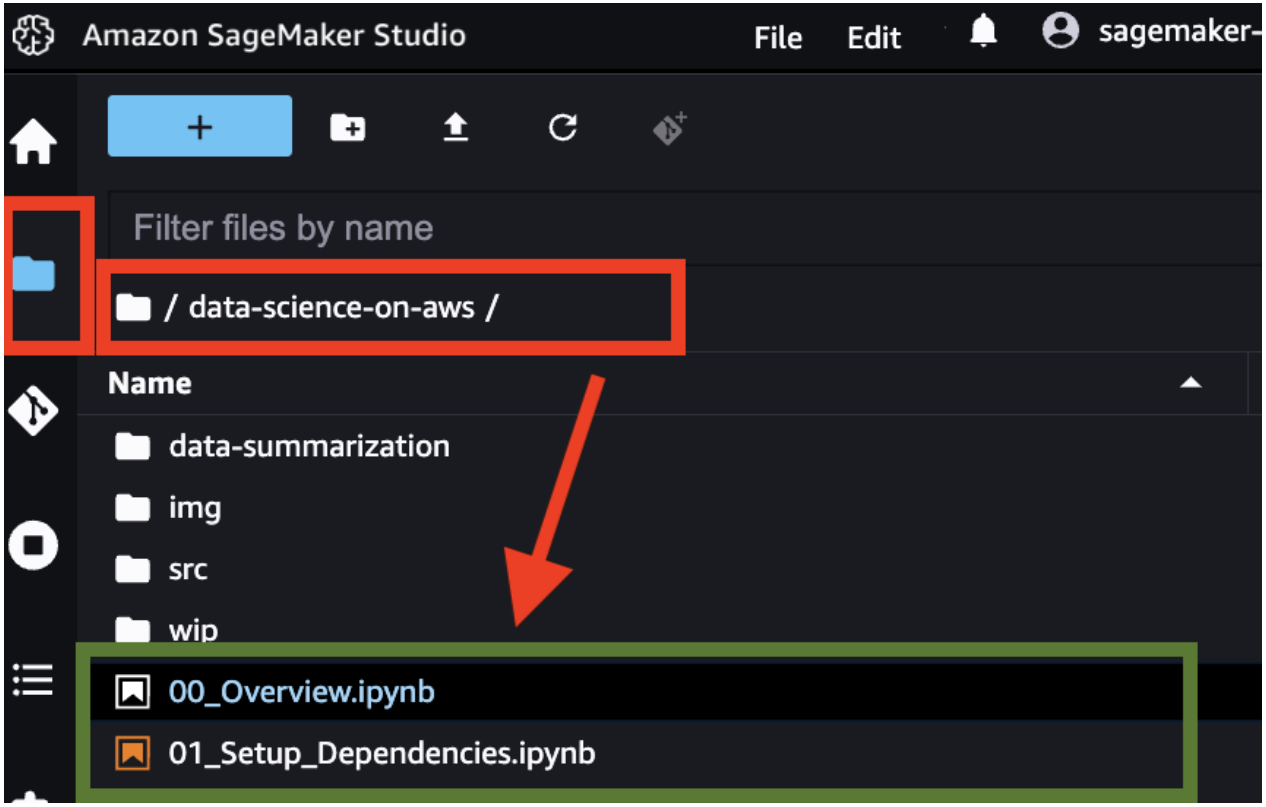
```
fatal: Unable to create '.git/index.lock': File exists.

Another git process seems to be running in this repository, e.g. an editor opened by 'git commit'.
Please make sure all processes are terminated then try again.
If it still fails, a git process may have crashed in this repository earlier:
remove the file manually to continue.
```

Just re-run the command again until it works. ***Note: This is not a fatal error ^^ above ^^. Just re-run the command again until it works.***

Data Science on AWS

- Introduction
  - Prerequisites
  - AWS Hosted Event
  - **Workshop Setup**
- Part 1: Distributed Data Preprocessing
- Part 2: Fine-tuning LLMs with Amazon SageMaker
- Part 3: Automating Fine-Tuning workflows with SageMaker Pipelines
- Contributors
- Summary

6. Start the workshop! In SageMaker Studio, click the folder icon in the left-side navigation menu, and navigate to the `data-science-on-aws` directory.



Congratulations
You have successfully completed the workshop setup. You can now start with the hands-on labs.

Please proceed to the next steps for instructions on each lab.