Apache Spark installation

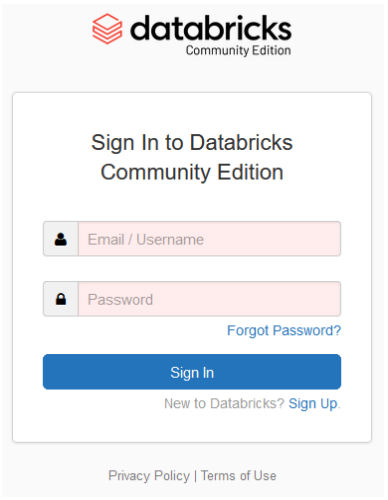
Databricks Community

The easiest way to try Apache Spark is to use Databricks Community which is a 14-day free trial of a cloud hosted instance of Apache Spark.

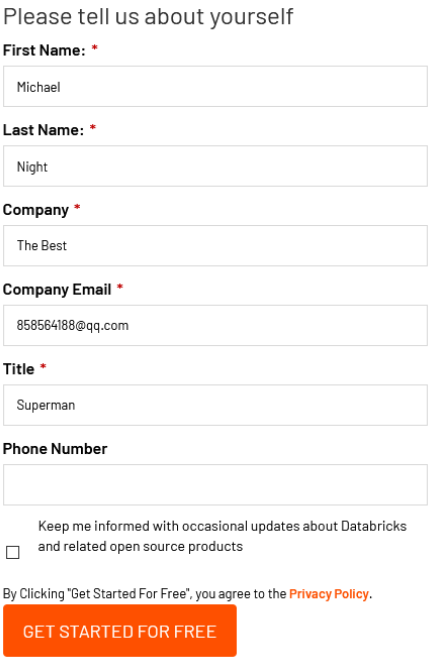
Databricks founders are the original inventor of Apache Spark.

How to sign-up for Databricks Community

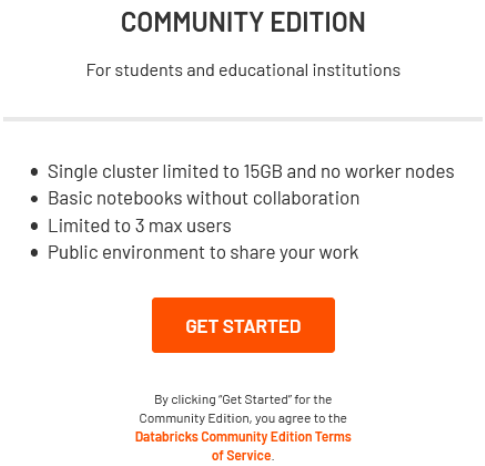
Navigate to https://community.cloud.databricks.com/ and click sign-up.



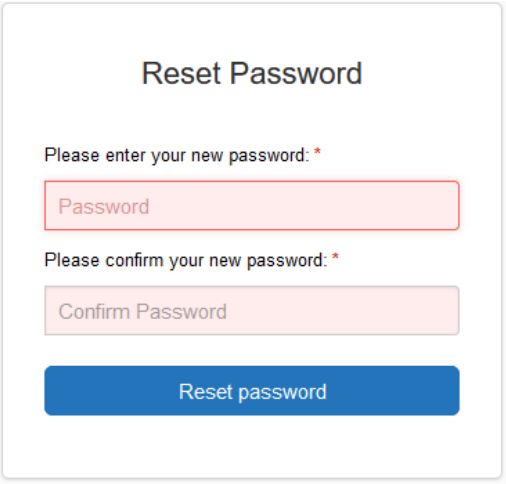
You will get to a registration page https://databricks.com/try-databricks



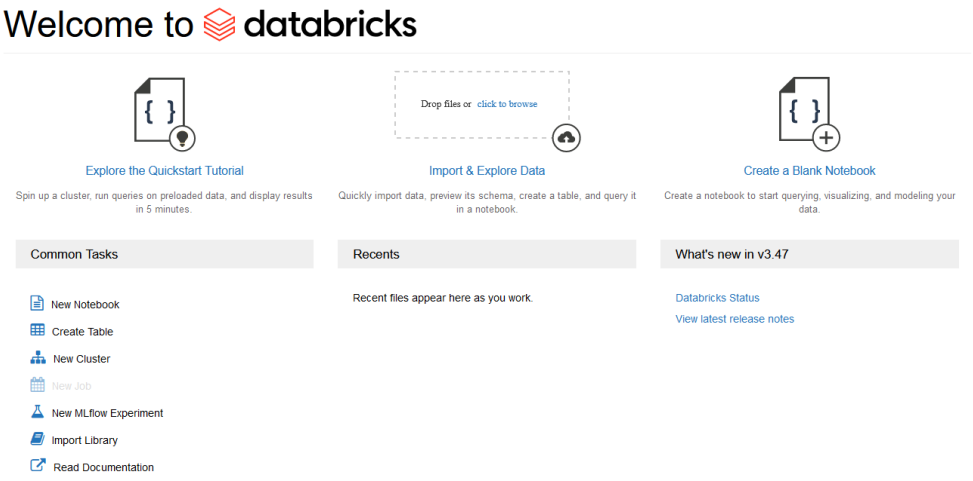
Select the Community edition and after clicking “GET STARTED” button you will receive an email for confirmation.



Go to the email and click the verification link and you will be redirected to enter the password for your account.

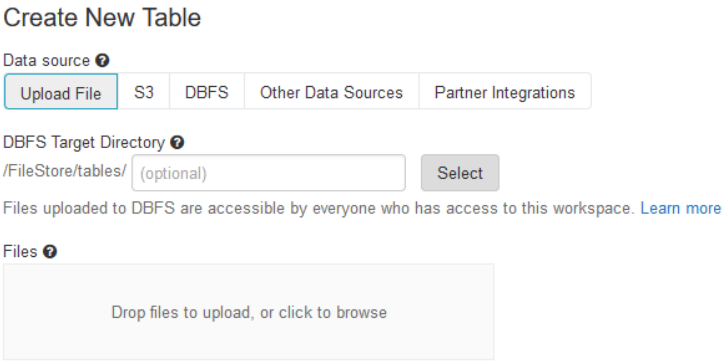


After resetting your password you will be redirected to Databricks workspace.



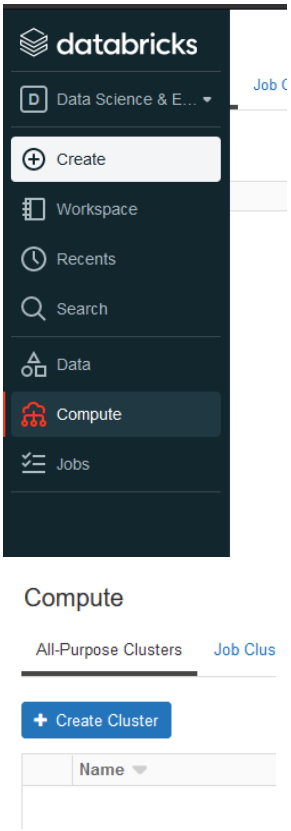
How to import data to Databricks

The easiest way to import the data is via workspace - on the homepage of your workspace in the middle click **Import & Explore Data**.

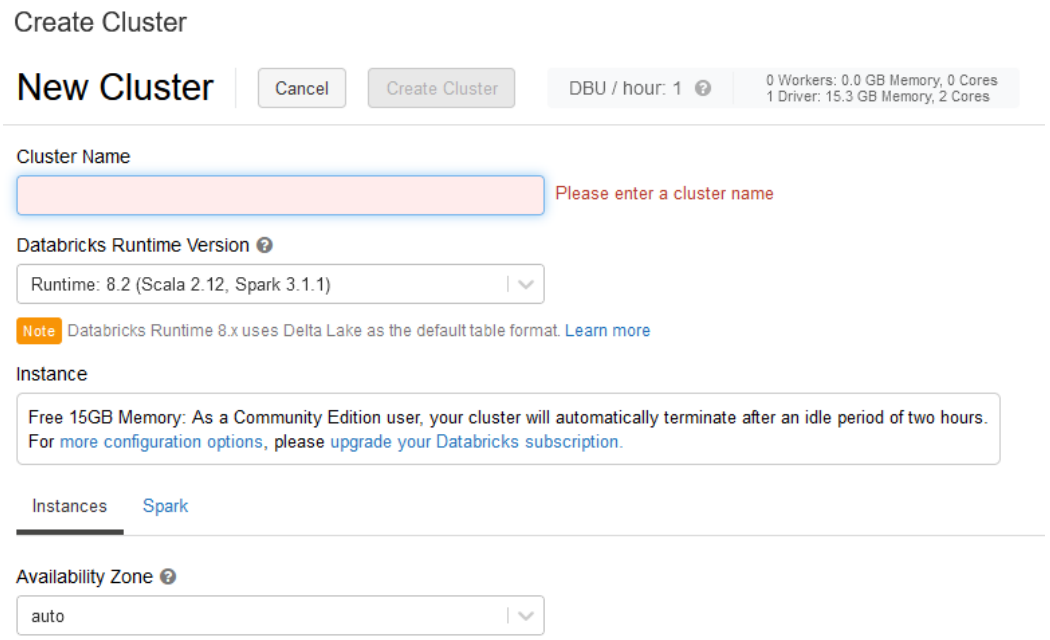


How to create a cluster

Go to the Compute menu in the left navigation pane and click the “Create cluster” button.

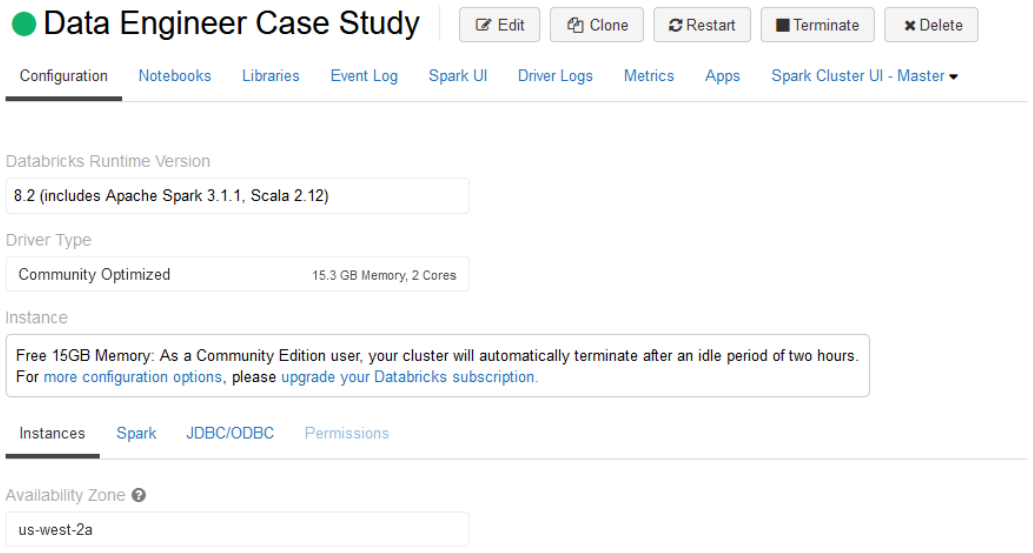


Enter the name of your cluster (any name you like) and click the “create cluster” button.



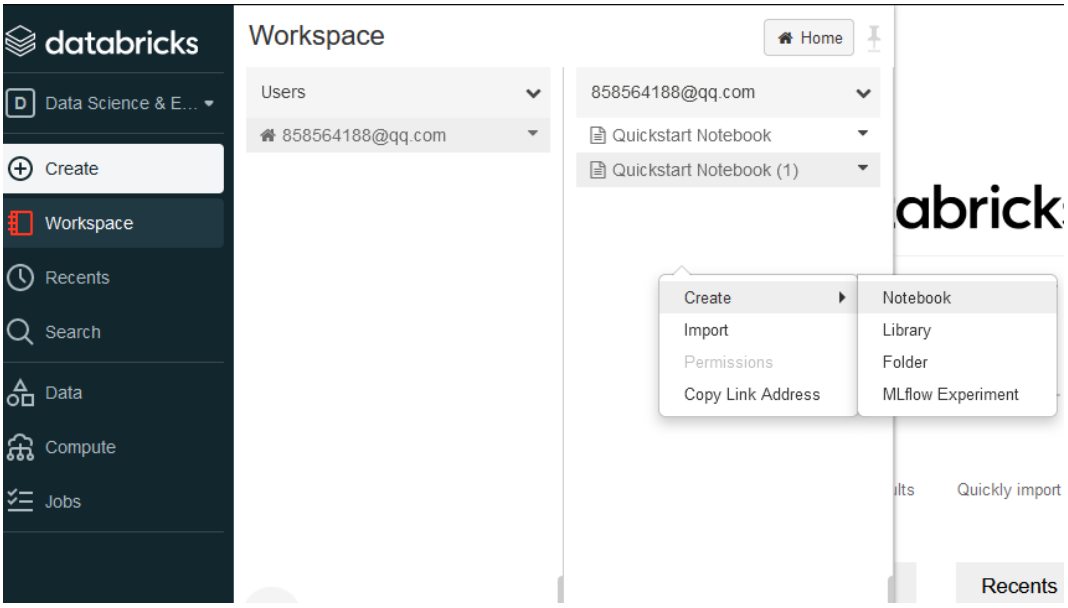
It will take some time to initiate the cluster. Cluster by default will be terminated after 2 hours of being idle.

You can also start, restart or terminate or edit the cluster from within the Compute menu.

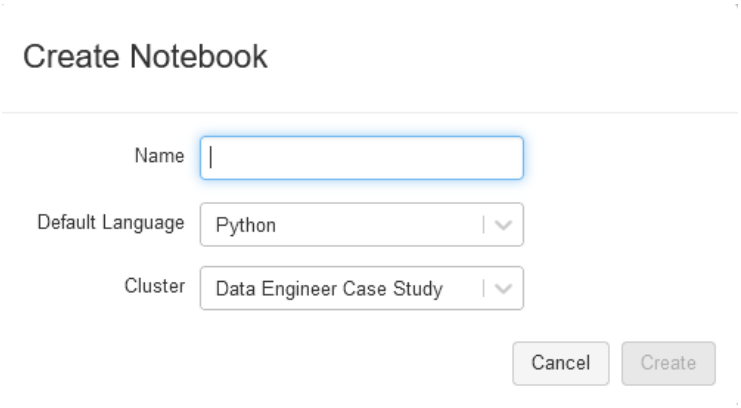


How to create a notebook to run code

Go to workspace in the left navigation pane, expand blades and under your user account right click and choose create notebook.



Name your notebook and select the main language and cluster for it. Click create and you can start coding :)



Additional resources

Community edition FAQ - https://databricks.com/product/faq/community-edition

Databricks documentation - https://docs.databricks.com/