

# Apache Spark installation

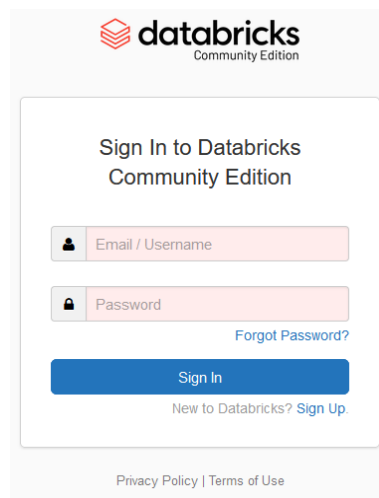
## Databricks Community

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.

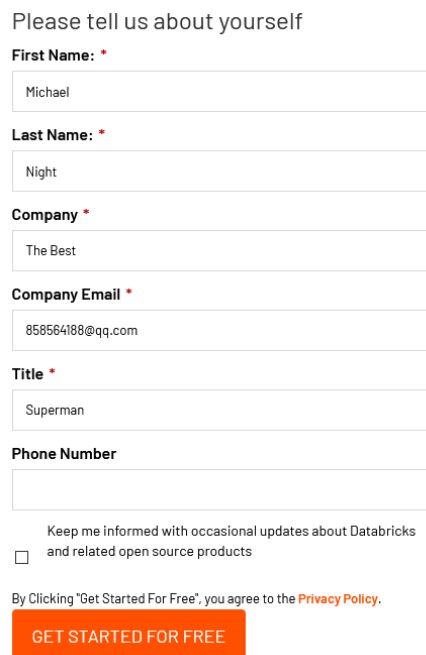
## How to sign-up for Databricks Community

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



The image shows the 'Sign In to Databricks Community Edition' page. At the top is the Databricks logo with 'Community Edition' underneath. The main heading is 'Sign In to Databricks Community Edition'. Below this are two input fields: 'Email / Username' and 'Password'. To the right of the password field is a link that says 'Forgot Password?'. Below the input fields is a blue 'Sign In' button. Underneath the button is a link that says 'New to Databricks? Sign Up.'. At the bottom of the page, there are links for 'Privacy Policy' and 'Terms of Use'.

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



The image shows a registration form titled 'Please tell us about yourself'. It contains several fields with red asterisks indicating required fields: 'First Name' (filled with 'Michael'), 'Last Name' (filled with 'Night'), 'Company' (filled with 'The Best'), 'Company Email' (filled with '858564188@qq.com'), and 'Title' (filled with 'Superman'). There is also a 'Phone Number' field which is empty. Below the fields is a checkbox for 'Keep me informed with occasional updates about Databricks and related open source products', which is currently unchecked. At the bottom, there is a line of text: 'By Clicking "Get Started For Free", you agree to the [Privacy Policy](#).' Below this text is an orange button that says 'GET STARTED FOR FREE'.

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

## COMMUNITY EDITION

For students and educational institutions

- Single cluster limited to 15GB and no worker nodes
- Basic notebooks without collaboration
- Limited to 3 max users
- Public environment to share your work

GET STARTED

By clicking "Get Started" for the Community Edition, you agree to the [Databricks Community Edition Terms of Service](#).

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

### Reset Password


Please enter your new password: \*

Please confirm your new password: \*

Reset password


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

Welcome to  databricks




#### Explore the Quickstart Tutorial

Spin up a cluster, run queries on preloaded data, and display results in 5 minutes.



#### Import & Explore Data








Quickly import data, preview its schema, create a table, and query it in a notebook.



#### Create a Blank Notebook

Create a notebook to start querying, visualizing, and modeling your data.

#### Common Tasks

-  New Notebook
-  Create Table
-  New Cluster
-  New Job
-  New MLflow Experiment
-  Import Library
-  Read Documentation

#### Recents

Recent files appear here as you work.

#### What's new in v3.47

[Databricks Status](#)  
[View latest release notes](#)

## 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**.

Create New Table

Data source ⓘ

[Upload File](#) [S3](#) [DBFS](#) [Other Data Sources](#) [Partner Integrations](#)

DBFS Target Directory ⓘ

/FileStore/tables/ (optional) [Select](#)

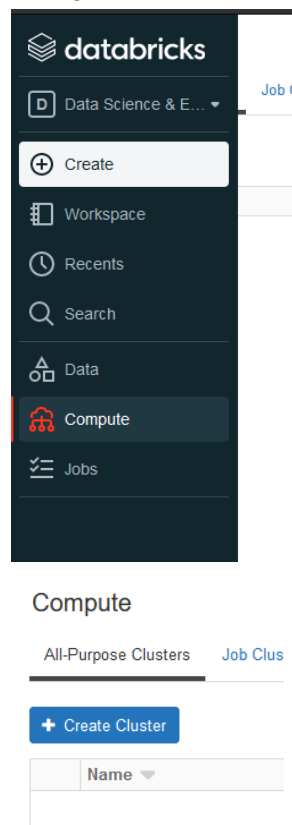
Files uploaded to DBFS are accessible by everyone who has access to this workspace. [Learn more](#)

Files ⓘ

Drop files to upload, or click to browse

## 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.

## Create Cluster

### New Cluster

Cancel

Create Cluster

DBU / hour: 1 ?

0 Workers: 0.0 GB Memory, 0 Cores  
1 Driver: 15.3 GB Memory, 2 Cores

Cluster Name

Please enter a cluster name

Databricks Runtime Version ?

Runtime: 8.2 (Scala 2.12, Spark 3.1.1)

**Note** Databricks Runtime 8.x uses Delta Lake as the default table format. [Learn more](#)

Instance

Free 15GB Memory: As a Community Edition user, your cluster will automatically terminate after an idle period of two hours. For [more configuration options](#), please [upgrade your Databricks subscription](#).

Instances

Spark

Availability Zone ?

auto

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.

### Data Engineer Case Study

Edit

Clone

Restart

Terminate

Delete

Configuration

Notebooks

Libraries

Event Log

Spark UI

Driver Logs

Metrics

Apps

Spark Cluster UI - Master ▼

Databricks Runtime Version

8.2 (includes Apache Spark 3.1.1, Scala 2.12)

Driver Type

Community Optimized

15.3 GB Memory, 2 Cores

Instance

Free 15GB Memory: As a Community Edition user, your cluster will automatically terminate after an idle period of two hours. For [more configuration options](#), please [upgrade your Databricks subscription](#).

Instances

Spark

JDBC/ODBC

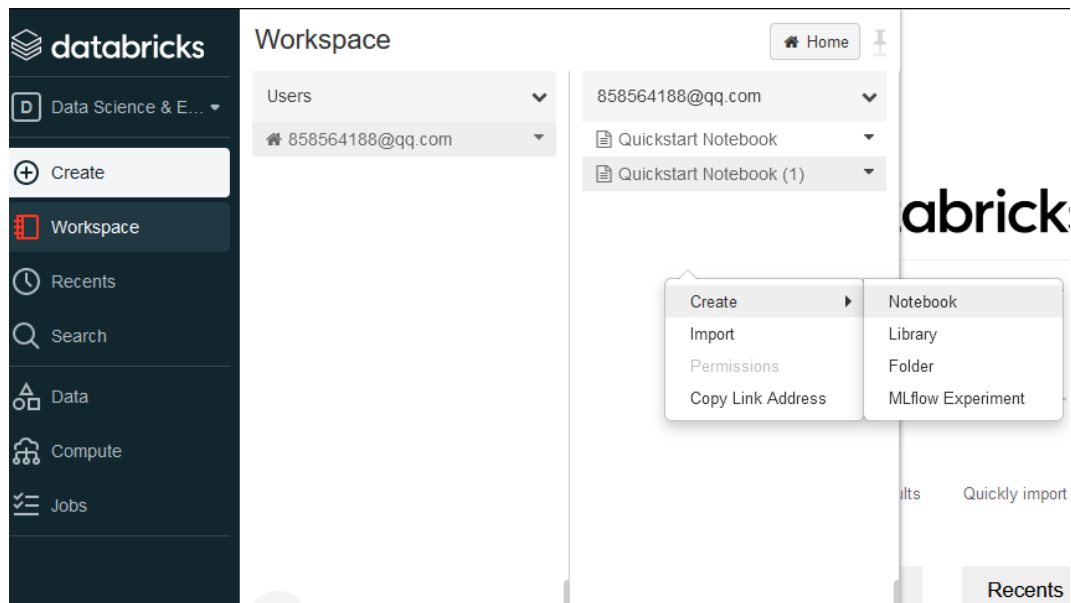
Permissions

Availability Zone ?

us-west-2a

## 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 :)

### Create Notebook

Name

Default Language

Cluster

## Additional resources

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

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