**Community Databricks Cluster Guide**

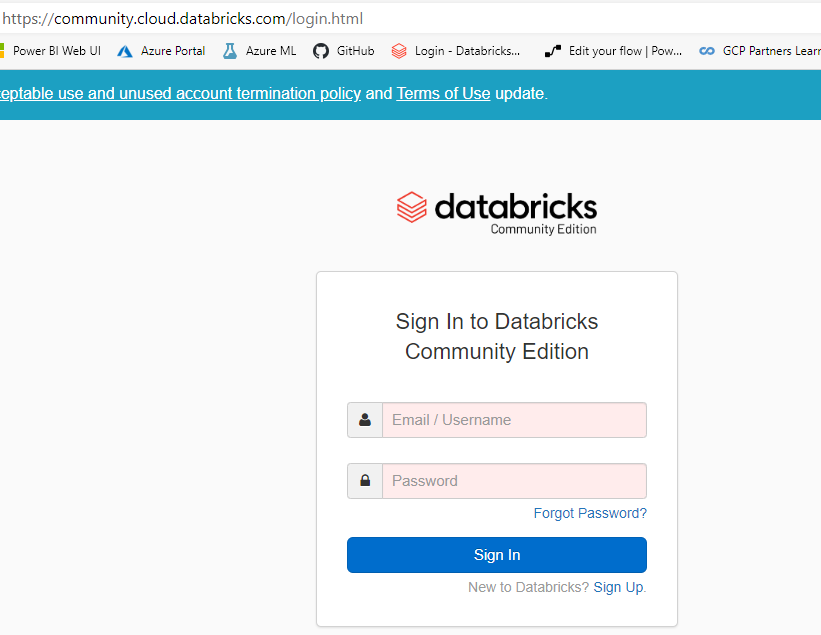
The Databricks UI provides an easy-to-use graphical interface to workspace folders and their contained objects, data objects, and computational resources.

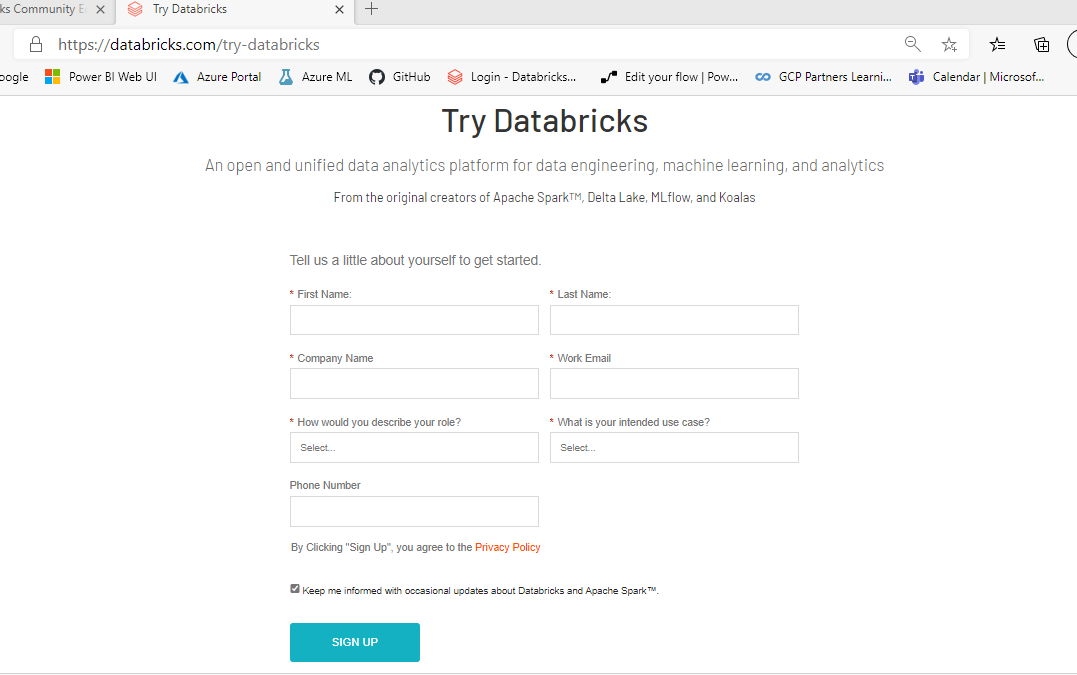
You run these workloads as a set of commands in a notebook.

The workspace is an environment for accessing all of your Databricks assets.

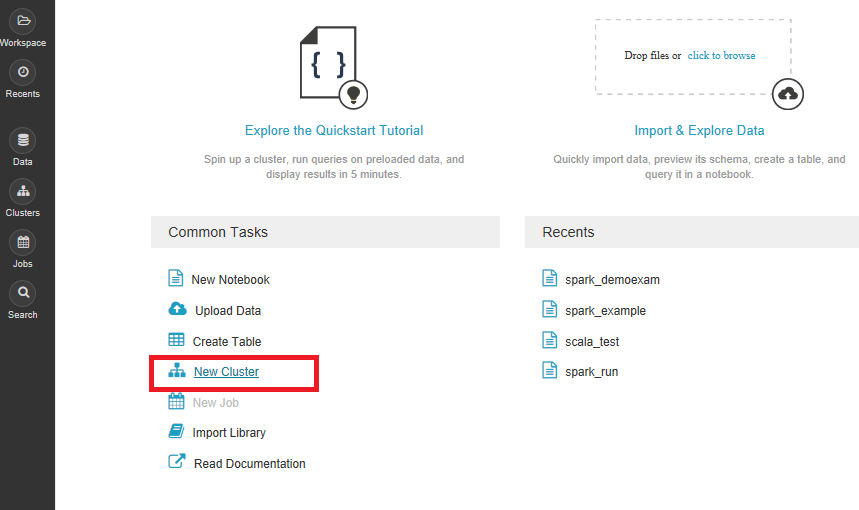
A Databricks cluster is a set of computation resources and configurations on which you run data engineering, data science, and data analytics workloads, such as production ETL pipelines, streaming analytics, ad-hoc analytics, and machine learning.

1. Login to <https://community.cloud.databricks.com/login.html>
2. Preferred Browser: Chrome or Edge
3. Create User ID using work email in **Databricks Community Edition** and login





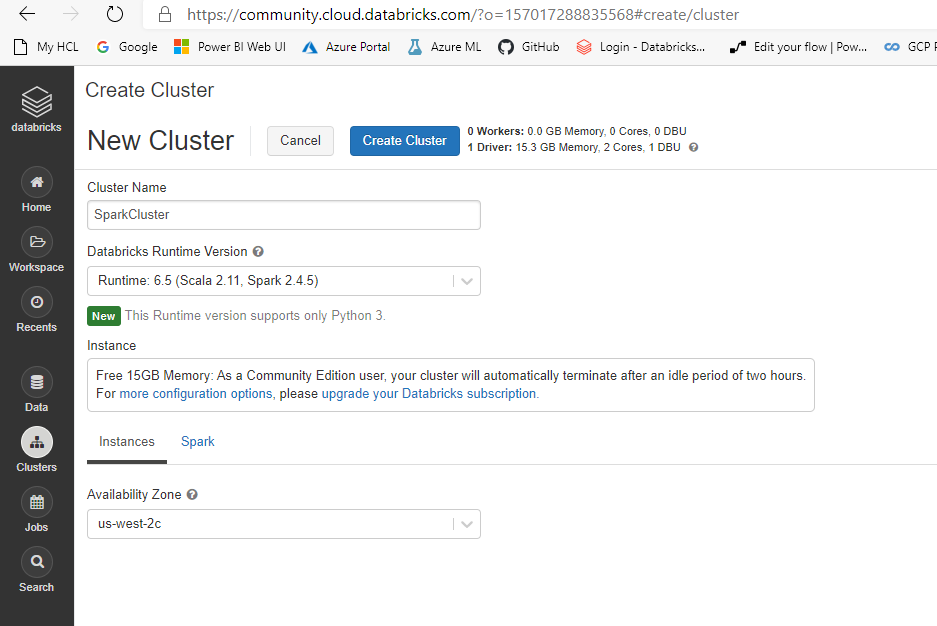
1. Click New cluster to create the cluster



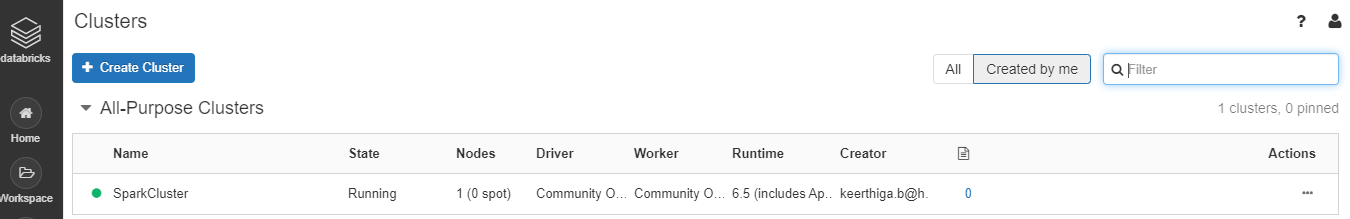
1. Give name to the cluster. The cluster has 15GB RAM, 2 cores

The cluster will automatically terminate after an idle period of two hours. Please recreate it.

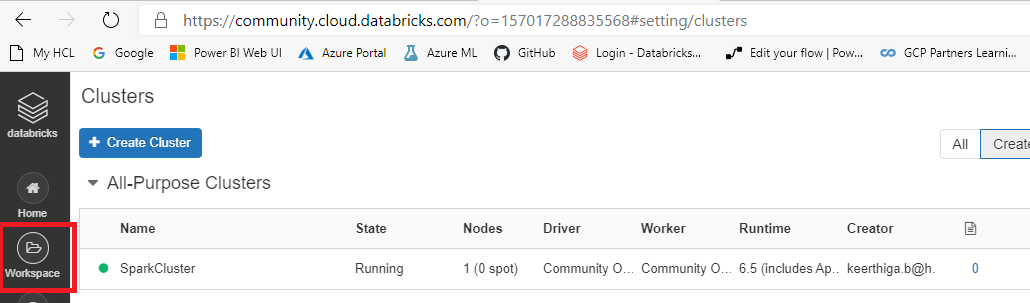
1. Click Create Cluster



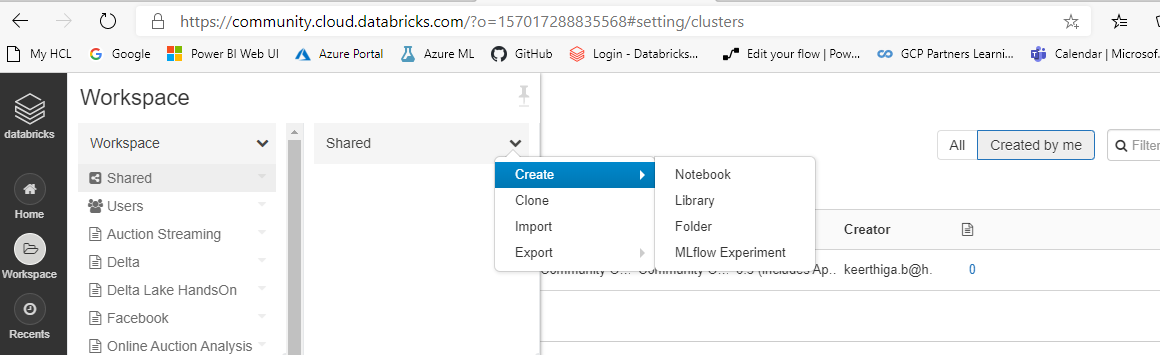
1. Your Spark Cluster is ready now to use



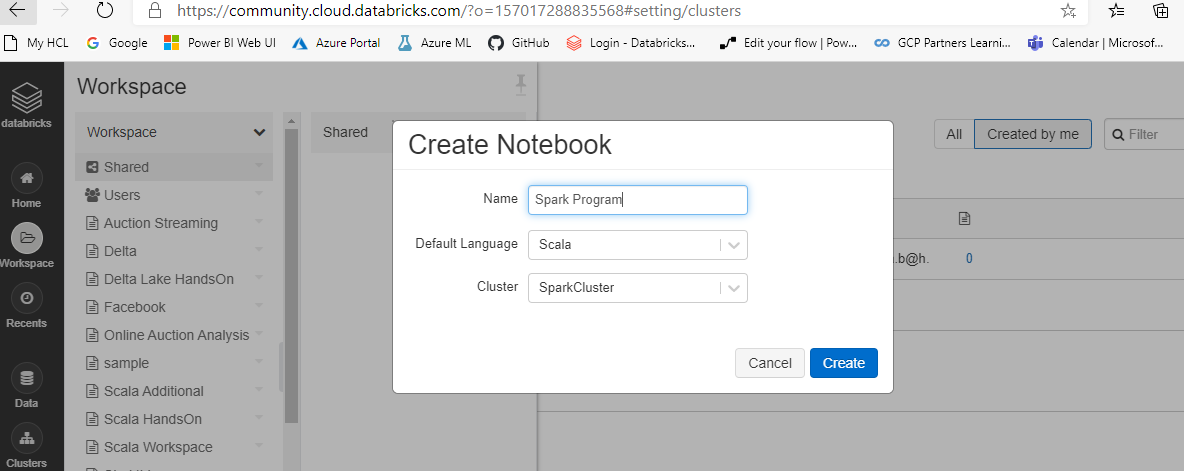
1. To execute spark jobs, click Workspace



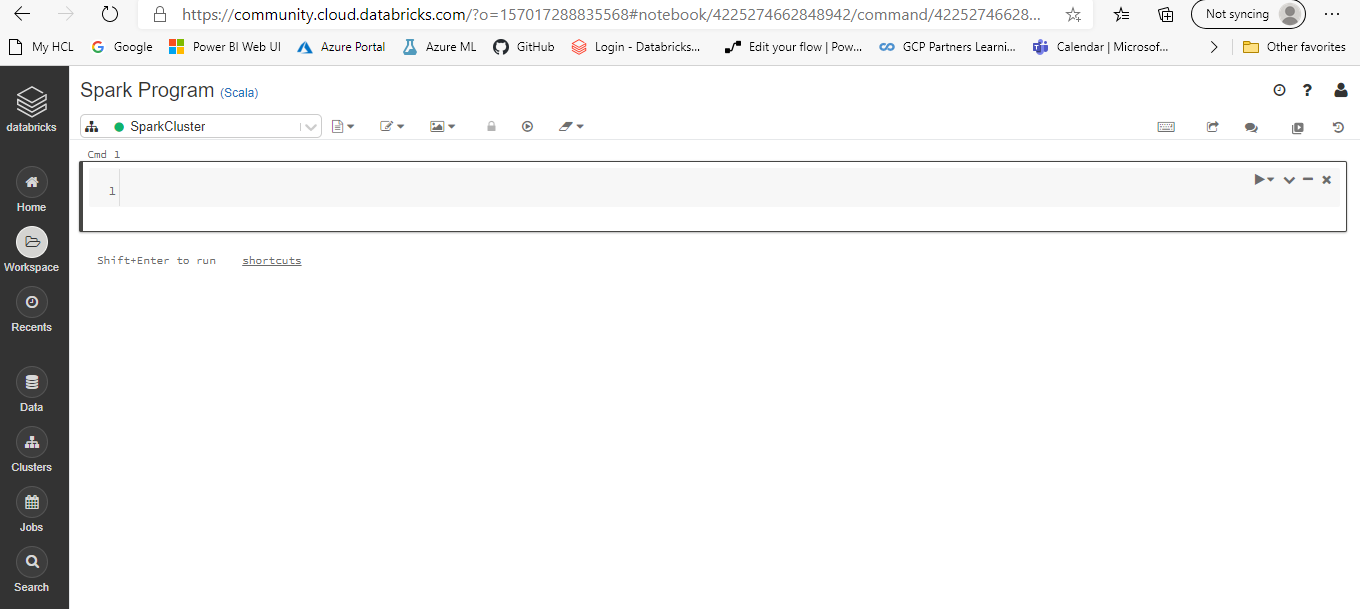
1. Click Workspace->Shared->Create->Notebook



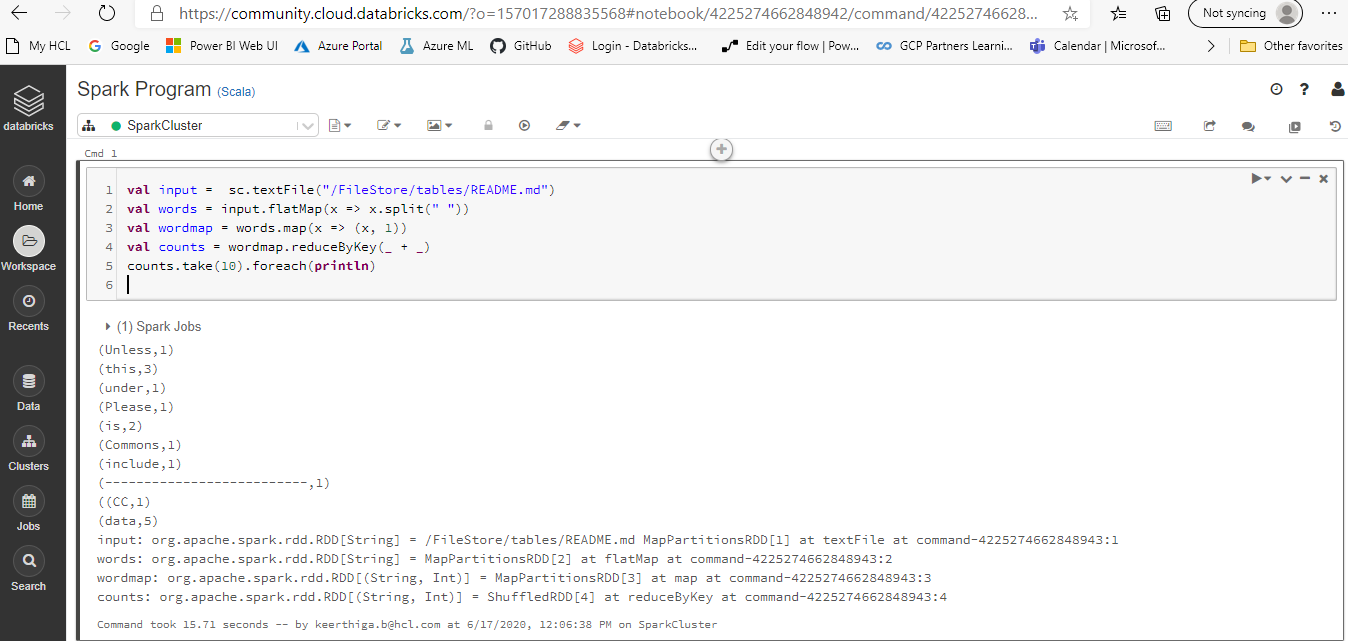
1. In Notebook, provide the name and choose language as Scala and then click Create.



1. Here you can execute your spark program.

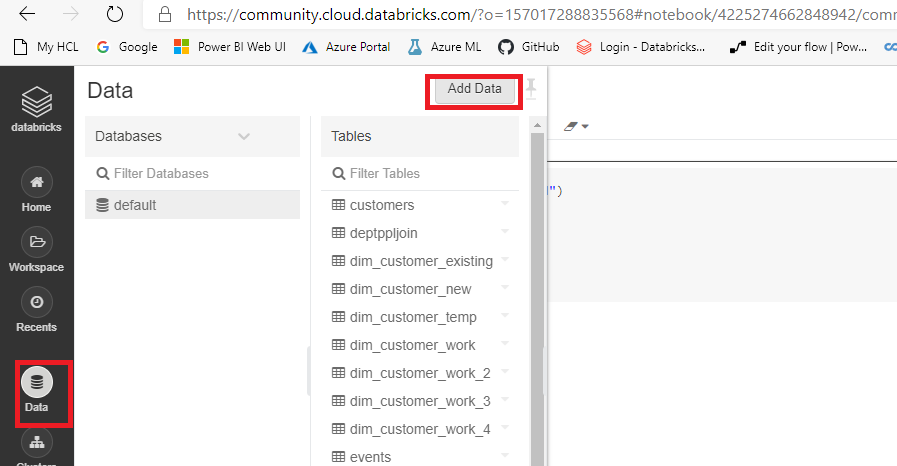


1. Click Run All, to run the program

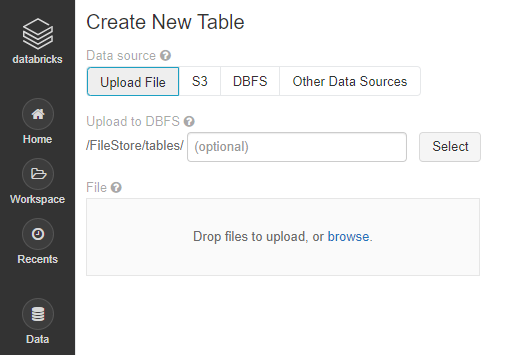


1. To add files

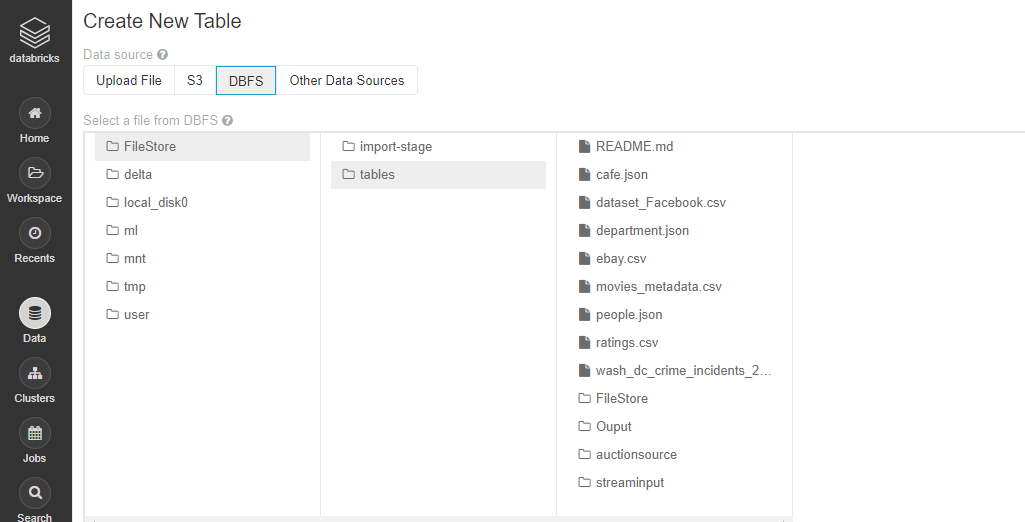
Click Data > Add Data



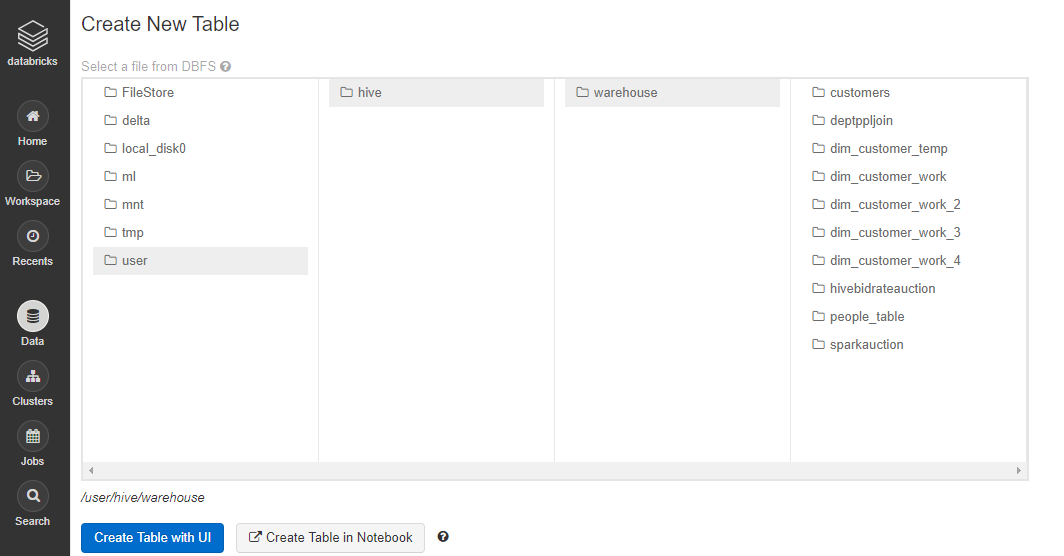
Files will be loaded in Filestore/Tables as default location. Sub directories can be created if needed and files can be uploaded



All the files uploaded is available below



User/Hive/Warehouse directory is also there for Hive Integration



1. Few DBFS Commands (Use it if needed)

**To remove file or directory:**

%fs rm -r "/FileStore/tables/posts"

**To list the contents of directory:**

%fs ls /FileStore/tables/posts/

**To see the contents of the File:**

dbutils.fs.head ("FileStore/tables/streamout-1562600954000/part-00001")