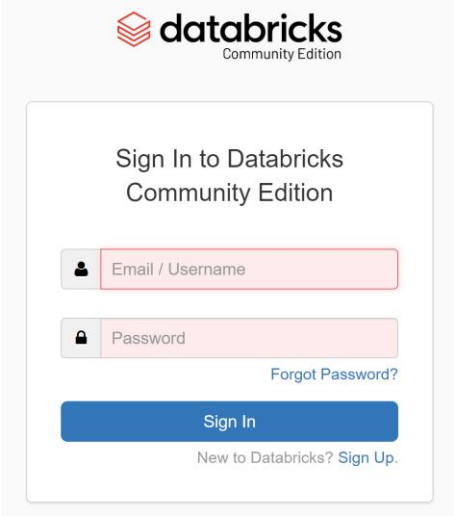
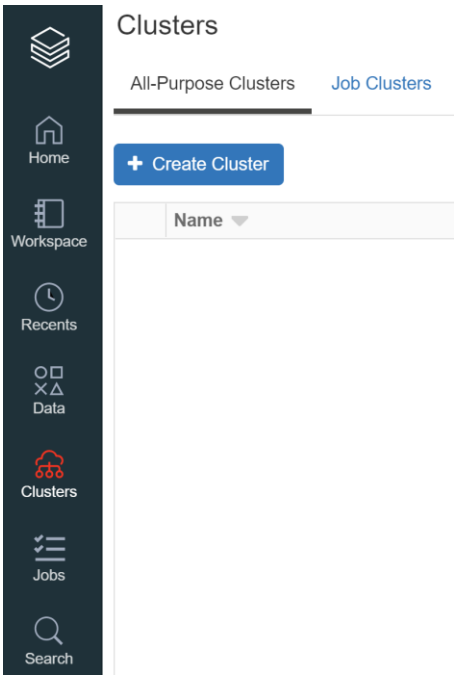
Using Databricks for Apache Spark

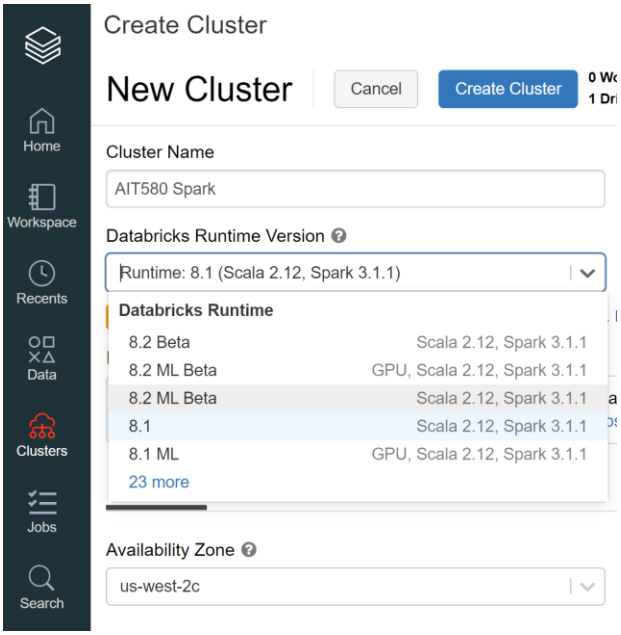
1. Go to https://community.cloud.databricks.com/



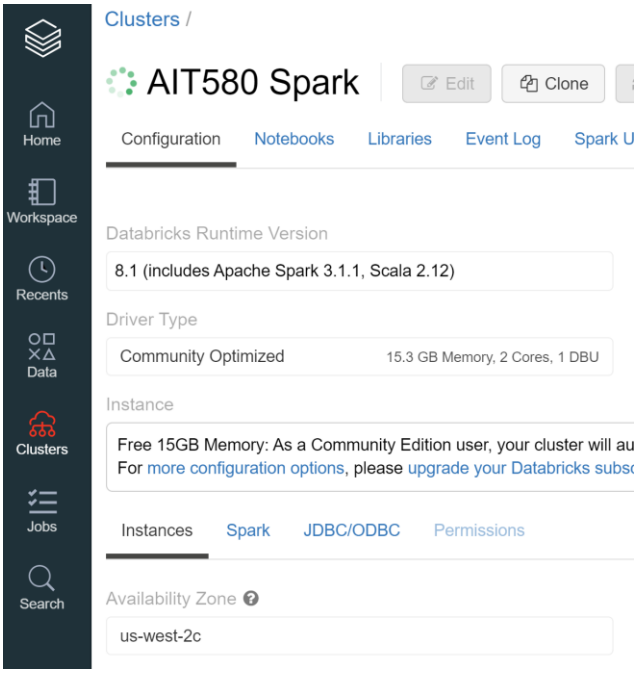
2. Sign Up and create your account
3. Login with your account
4. On the left menu, click Clusters



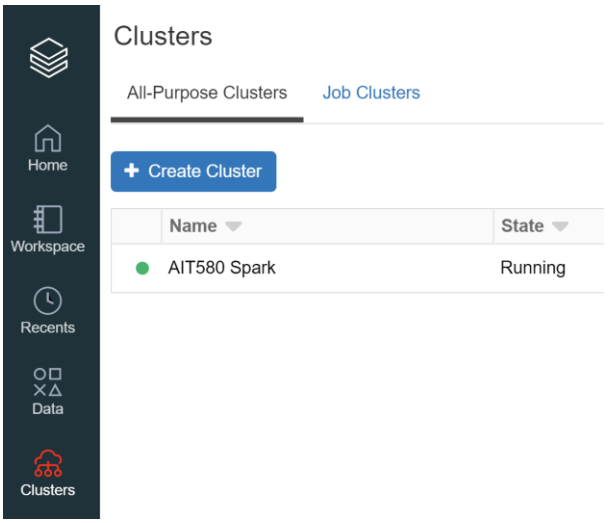
5. Click Create Cluster
6. Fill the required options. For Runtime, choose 8.2 ML Beta Scala (no GPU)



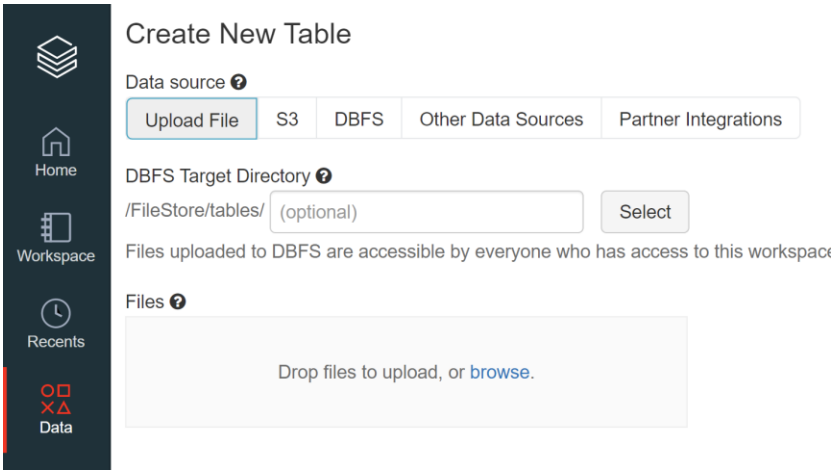
7. Click Create Cluster
8. You will see the screen below. It will take some time to get the cluster ready.



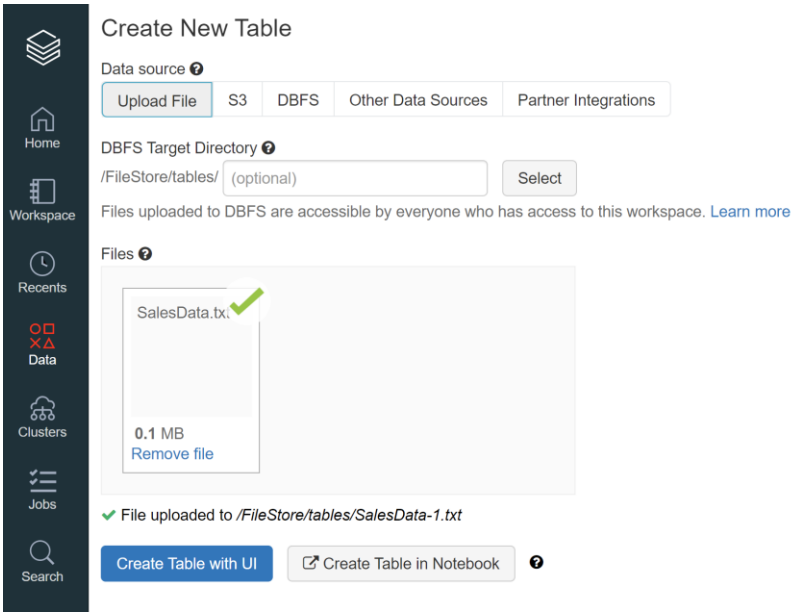
9. You will see green dot when cluster is ready. Click on the cluster again and make sure the status is ready



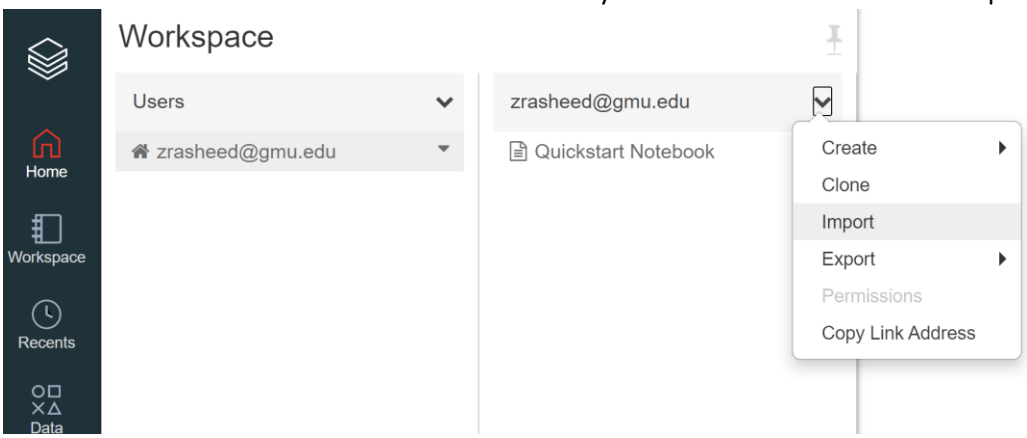
10. On the left, click on Data, and then click on Create Table



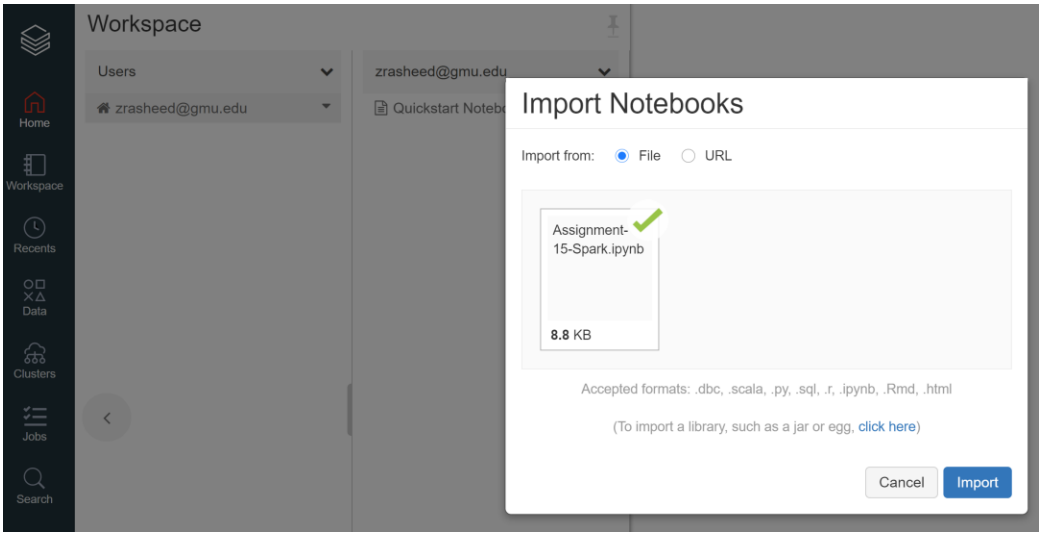
11. Click Browse and upload your data file. Make sure to note the File Uploaded Path. This path will be used in your Spark Notebook for reading data



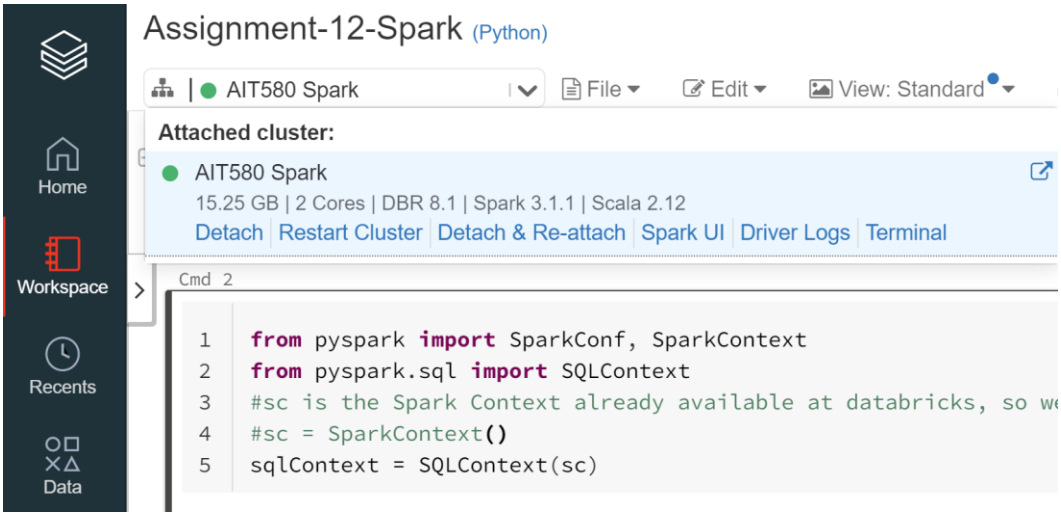
12. Click on Home. Click on down arrow at the end of your username and then click Import



13. Browse and Import your Notebook.



14. Your notebook will open. If not, then you can always click on Workspace on left menu and open the notebook.
15. Before running anything, you need to attach the running cluster to your notebook



16. Happy Coding