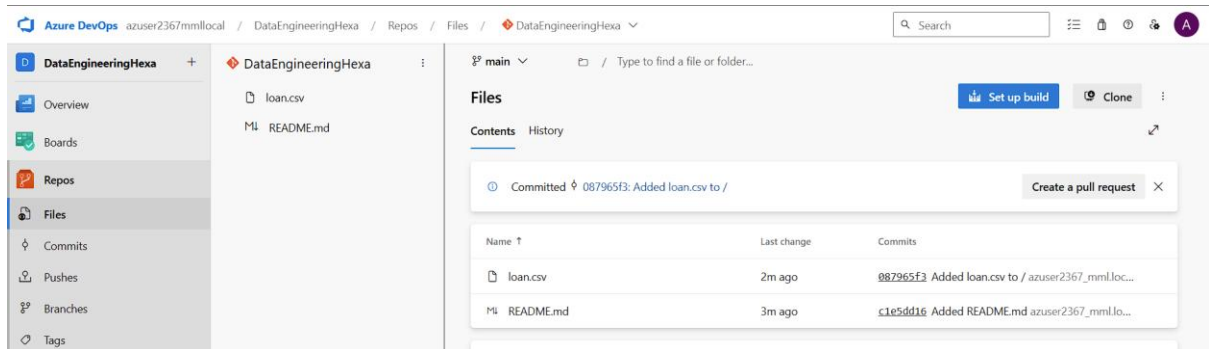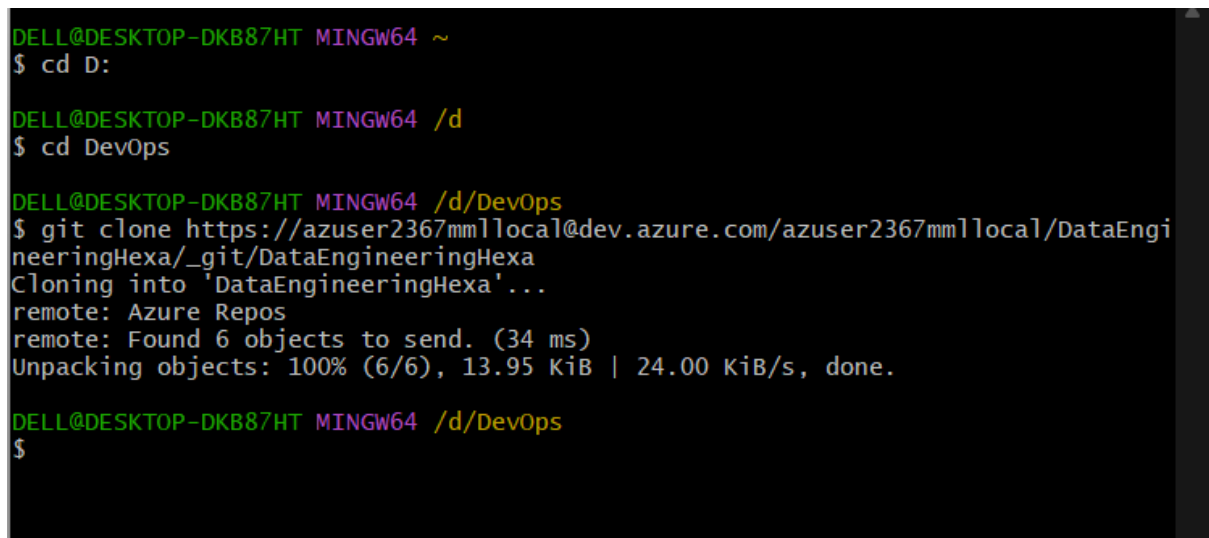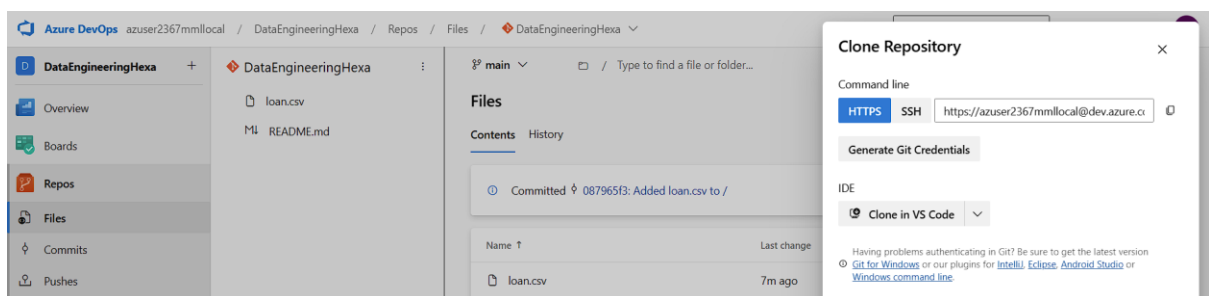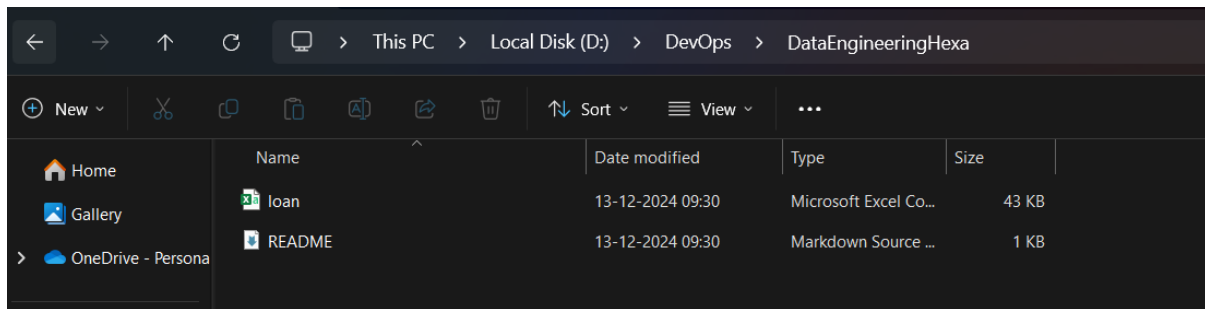Data Engineering Azure DevOps

Creating Repo

Go to Repos -> add the files and commit.

Click on the clone and copy the link and paste it in git bash. Type git clone

Give authentication -> using temporary access tokens -> (use different account & click on same account)

Name	Date modified	Type	Size
loan	13-12-2024 09:30	Microsoft Excel Co...	43 KB
README	13-12-2024 09:30	Markdown Source ...	1 KB

From local to repo

```
DELL@DESKTOP-DKB87HT MINGW64 /d/DevOps
$ cd DevOps
bash: cd: DevOps: No such file or directory

DELL@DESKTOP-DKB87HT MINGW64 /d/DevOps
$ touch MyFile Myfile.ipynb

DELL@DESKTOP-DKB87HT MINGW64 /d/DevOps
$ git add MyFile Myfile.ipynb
fatal: not a git repository (or any of the parent directories): .git

DELL@DESKTOP-DKB87HT MINGW64 /d/DevOps
$ cd /d/DevOps/DataEngineeringHexa

DELL@DESKTOP-DKB87HT MINGW64 /d/DevOps/DataEngineeringHexa (main)
$ touch MyFile.txt MyFile.ipynb

DELL@DESKTOP-DKB87HT MINGW64 /d/DevOps/DataEngineeringHexa (main)
$ git add MyFile.txt MyFile.ipynb

DELL@DESKTOP-DKB87HT MINGW64 /d/DevOps/DataEngineeringHexa (main)
$ git commit -m "Adding Files"
[main 2ae1aa6] Adding Files
 2 files changed, 0 insertions(+), 0 deletions(-)
 create mode 100644 MyFile.ipynb
 create mode 100644 MyFile.txt

DELL@DESKTOP-DKB87HT MINGW64 /d/DevOps/DataEngineeringHexa (main)
$ git push
Enumerating objects: 4, done.
Counting objects: 100% (4/4), done.
Delta compression using up to 4 threads
Compressing objects: 100% (2/2), done.
Writing objects: 100% (3/3), 318 bytes | 106.00 KiB/s, done.
Total 3 (delta 0), reused 0 (delta 0), pack-reused 0 (from 0)
remote: Analyzing objects... (3/3) (5 ms)
remote: Validating commits... (1/1) done (0 ms)
remote: Storing packfile... done (26 ms)
remote: Storing index... done (31 ms)
To https://dev.azure.com/azuser2367mmllocal/DataEngineeringHexa/_git/DataEnginee
ringHexa
   087965f..2ae1aa6  main -> main

DELL@DESKTOP-DKB87HT MINGW64 /d/DevOps/DataEngineeringHexa (main)
$
```
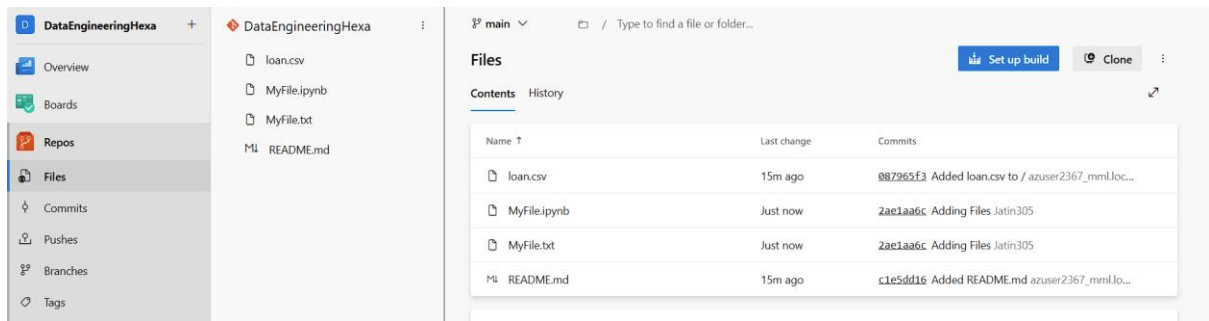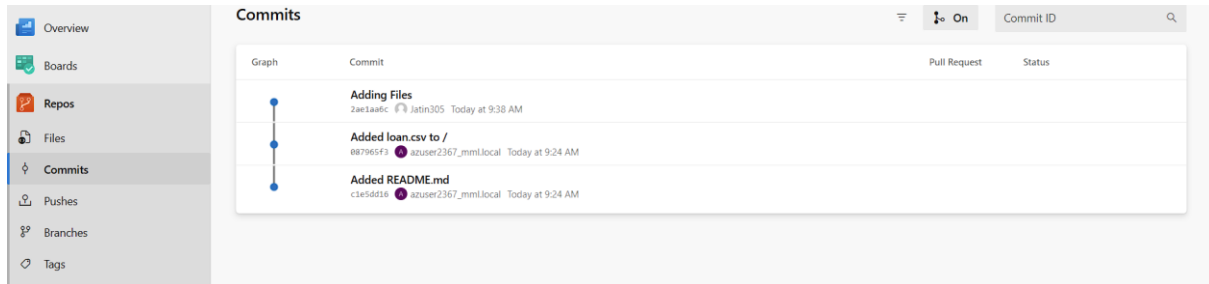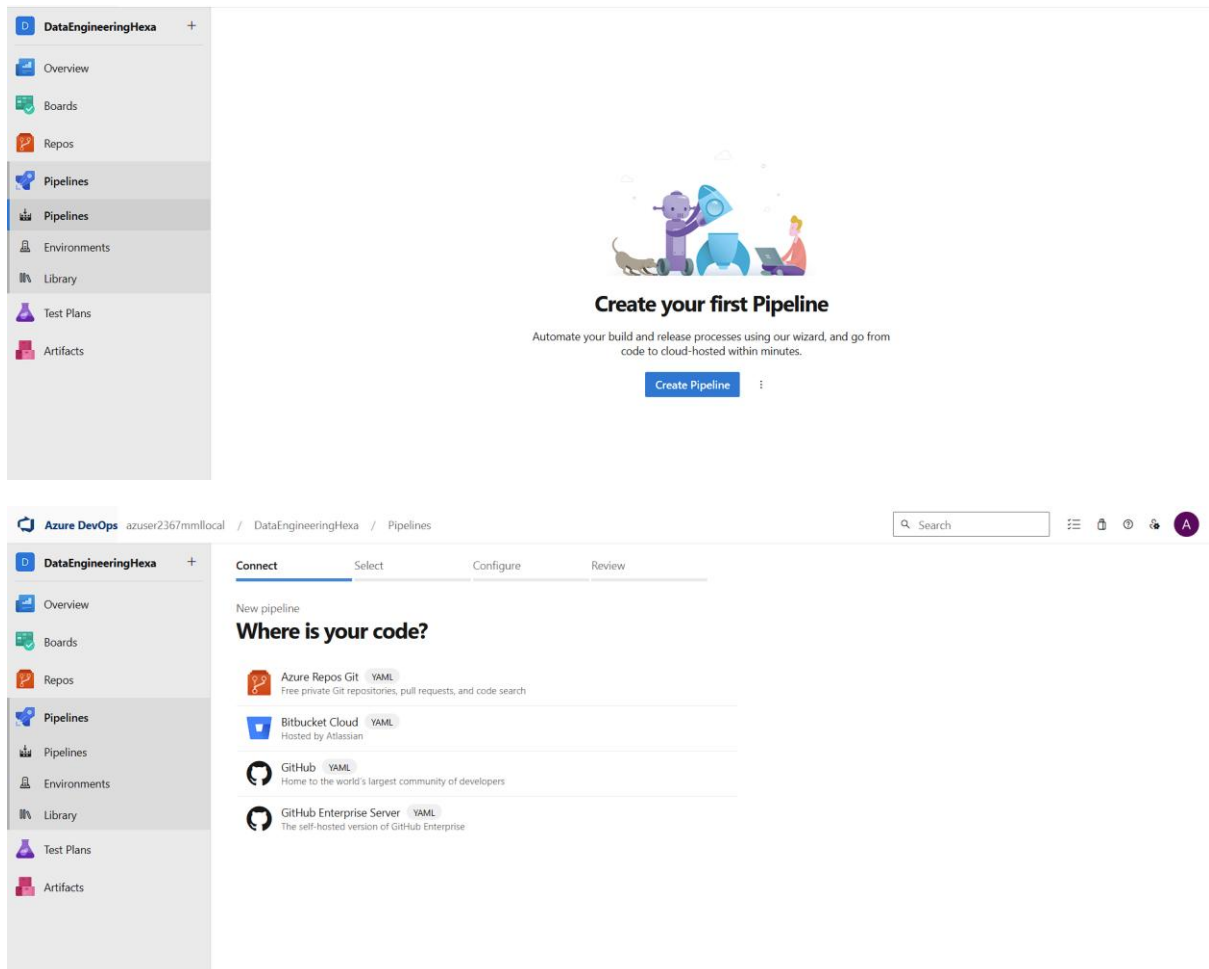
Check it in commits

## Pipeline creation

Choose starter pipeline


```
1   # Starter pipeline
2   # Start with a minimal pipeline that you can customize to build and deploy your code.
3   # Add steps that build, run tests, deploy, and more:
4   # https://aka.ms/yaml
5
6   trigger:
7   - main
8
9   pool:
10    vmImage: ubuntu-latest
11
12  steps:
13  - script: echo Hello, world!
14    displayName: 'Run a one-line script'
15
16  - script: |
17      echo Add other tasks to build, test, and deploy your project.
18      echo See https://aka.ms/yaml
19    displayName: 'Run a multi-line script'
20
```
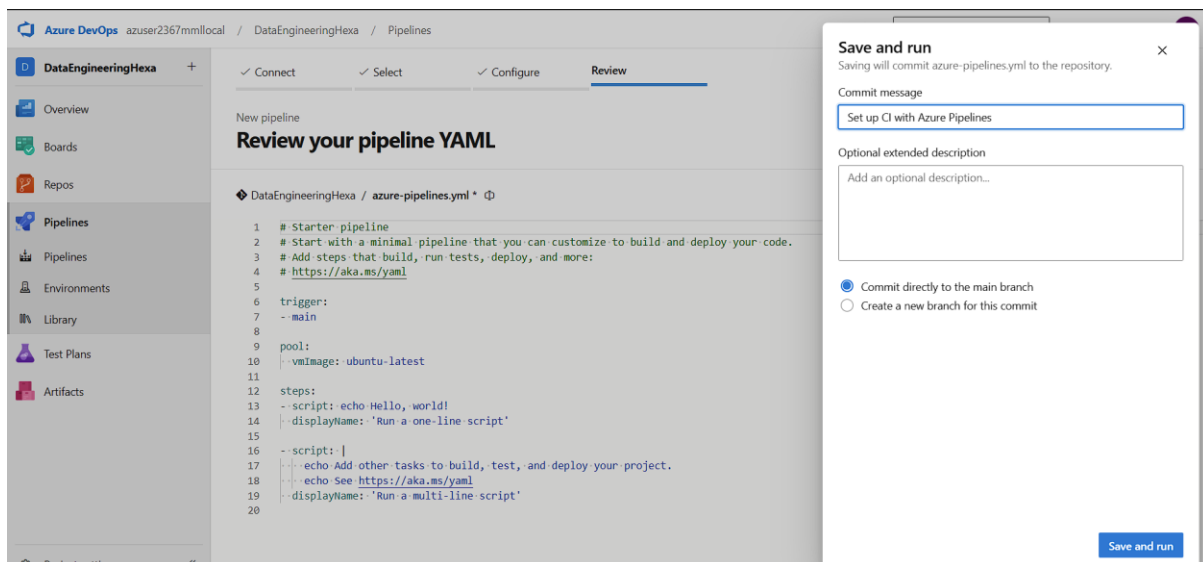
Pipeline running
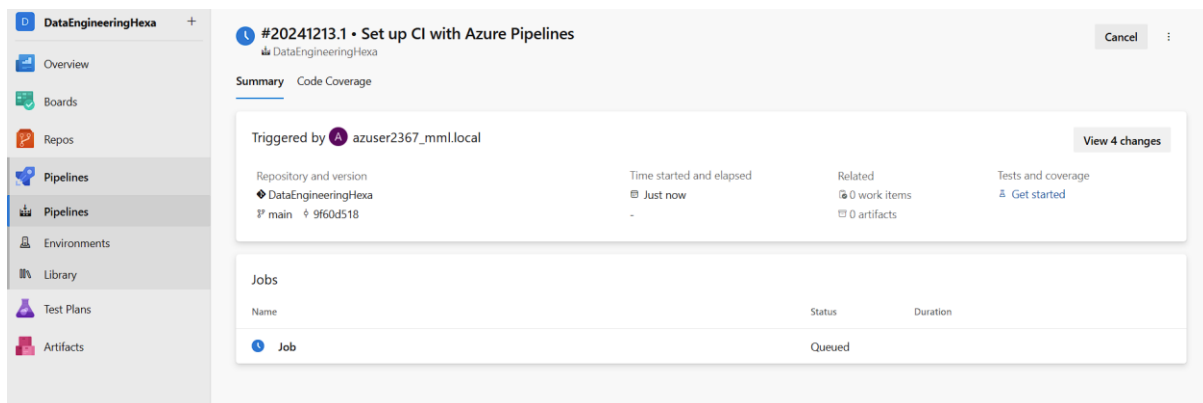
Pipeline ran successfully
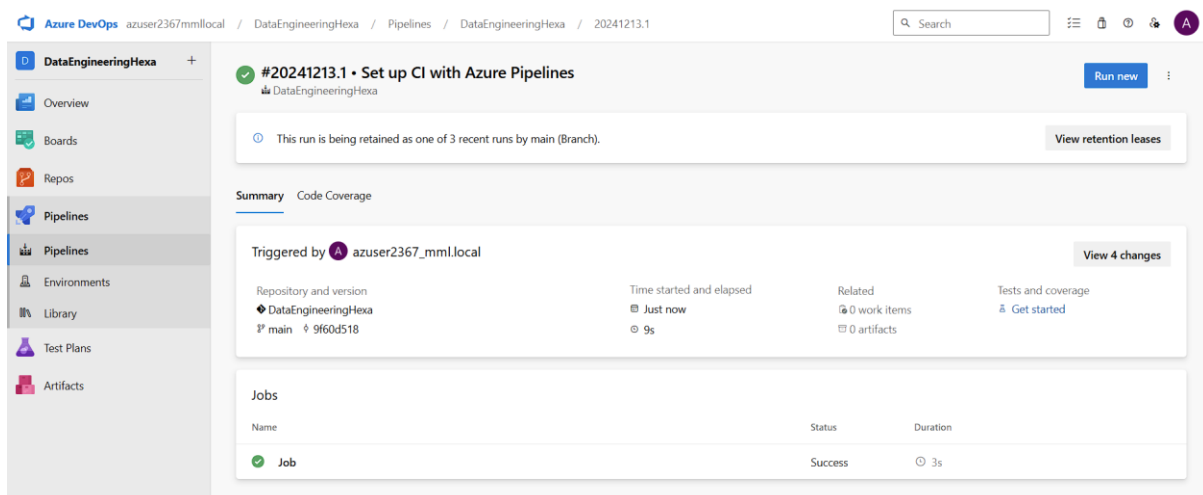
Connecting repo to databricks

In databricks go to settings
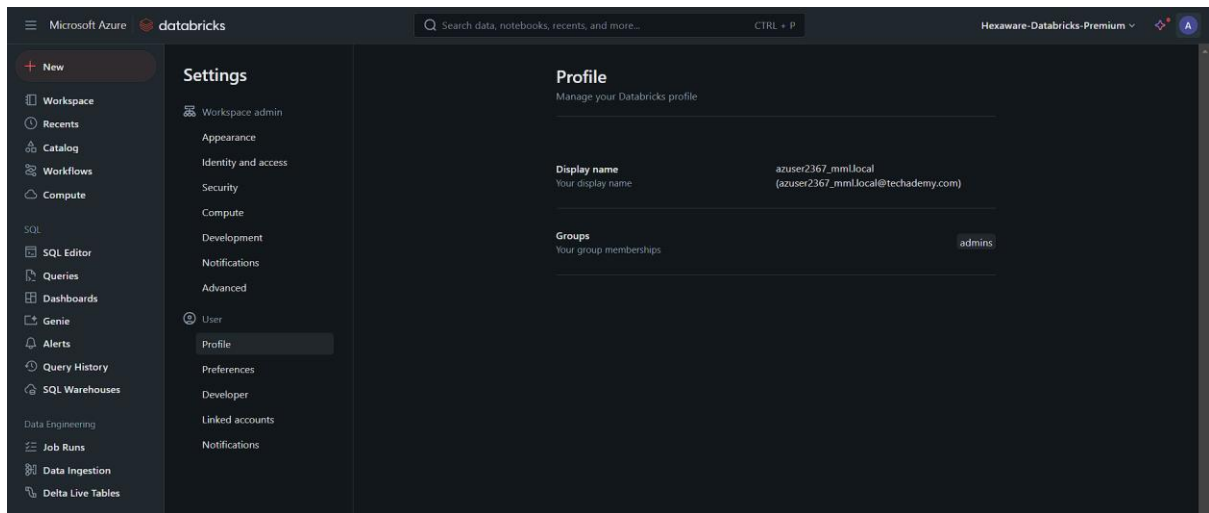
Go to linked accounts, here we have to choose the Azure devops services and give mail id, access token (taken from Azure repo)
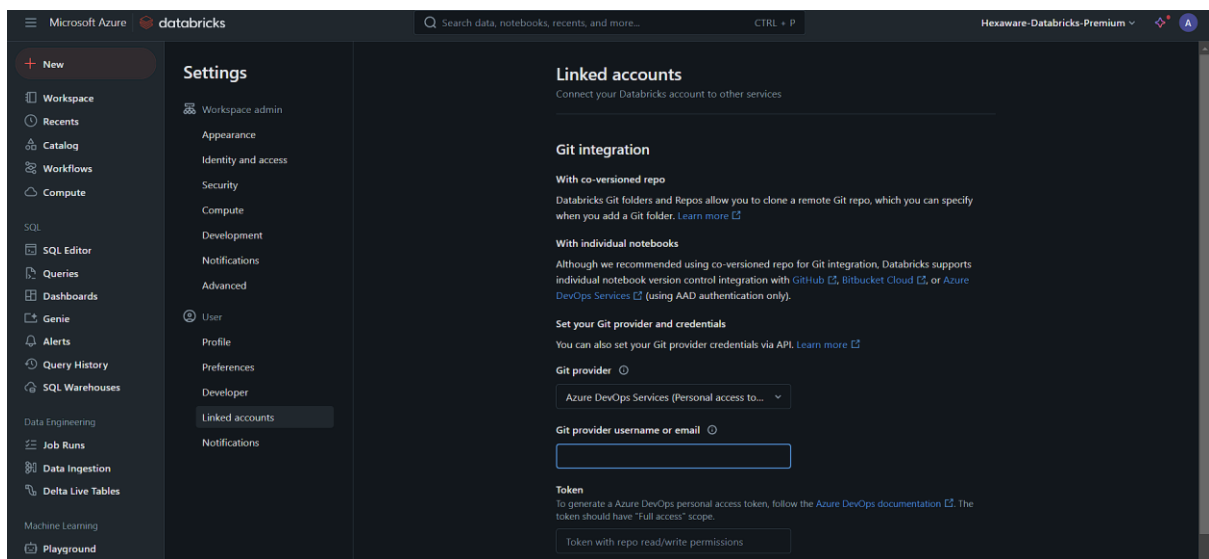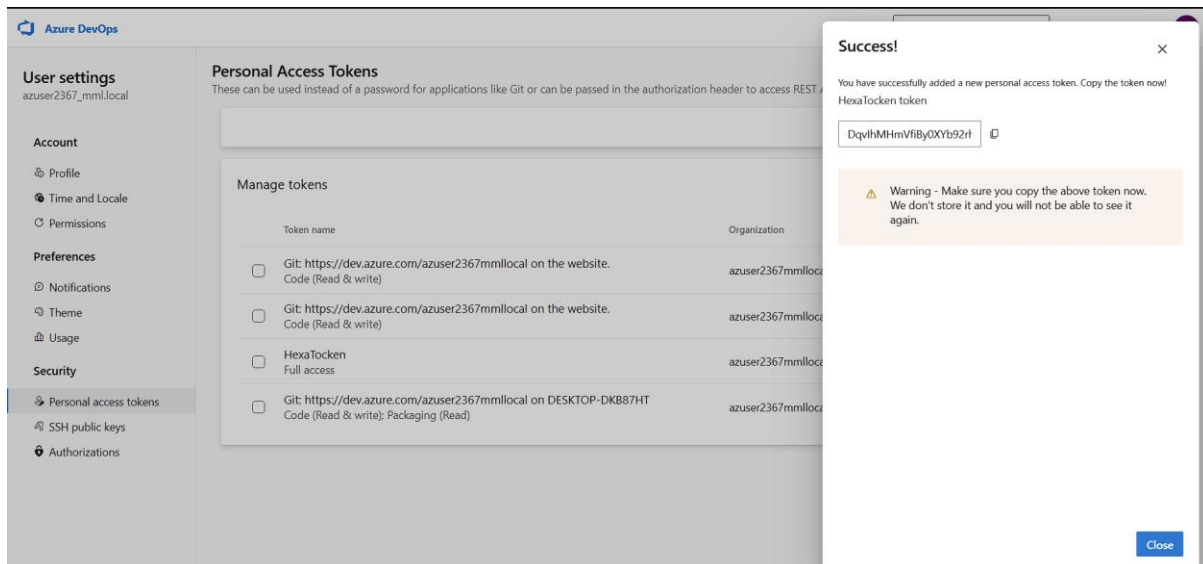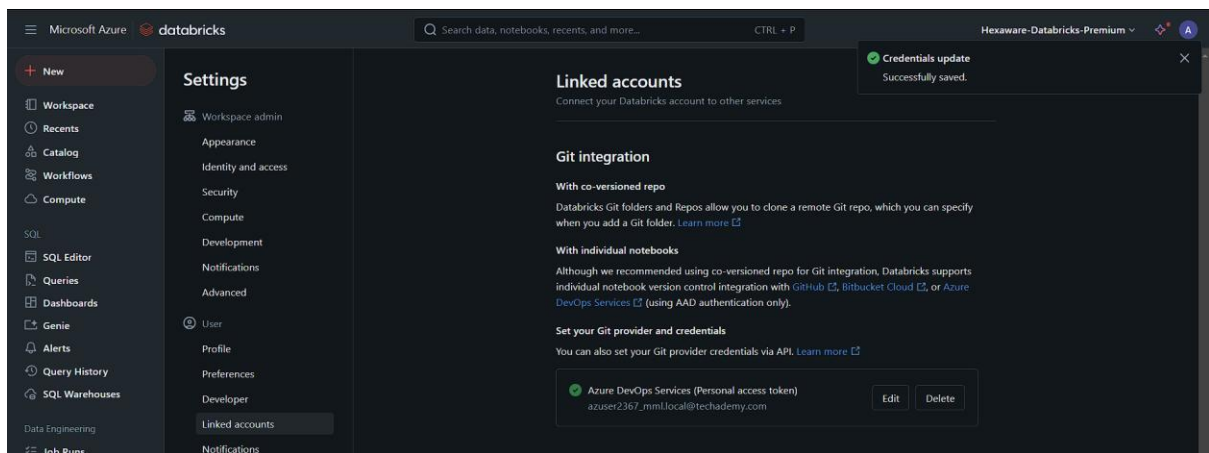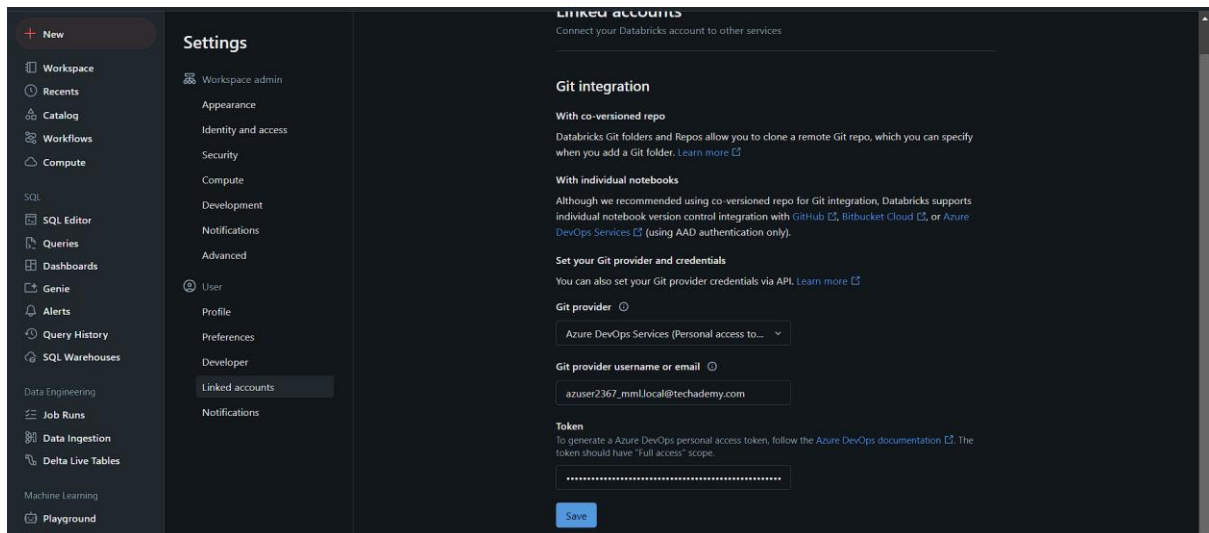

Go to azure repo -> personal access token
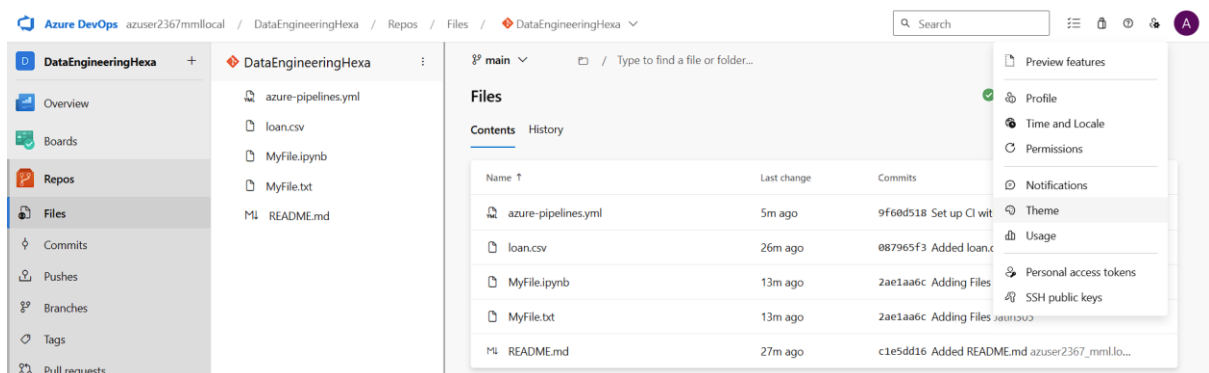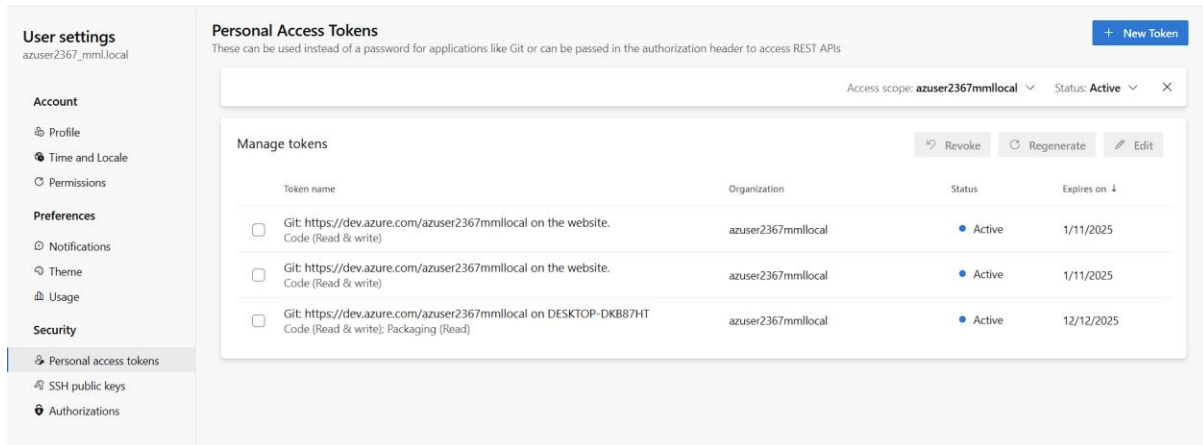
Click on + new token

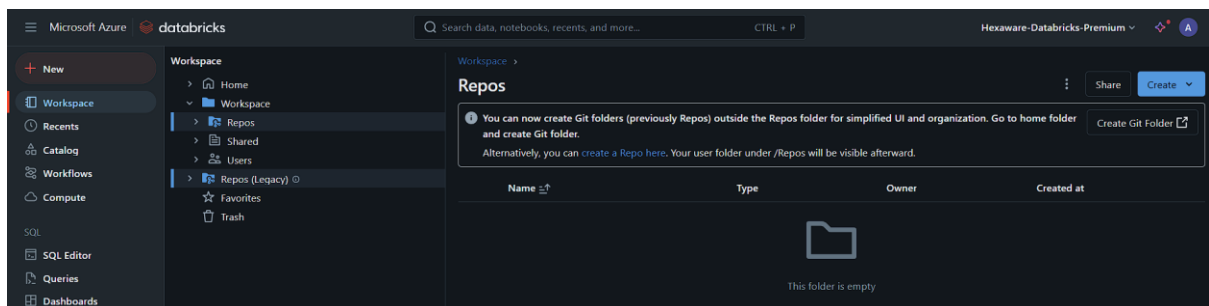Give name, access rights click on create.

Copy the token and save it for future uses.

Paste the token copied from azure repo in the linked accounts section of databricks and click on save

Now go to workspaces and create repos

Repo Cloned

Adding files:

Now refresh and check in Azure devops portal. Newly committed file fetched here.