NAME : AKULA SHARATH CHANDRA
BATCH : DATA ENGINEERING
DATE : 28-02-2024

Azure DevOps CODING CHALLENGE

QUESTIONS:

1)Create Azure Devops Environment and configure Azure Devops Git Repository ,configure on your local git to implement this upload few test files on the same.
2)Leverage the practices of CICD Using Azure Data Engineering and explain the architecture of the Azure synapse .

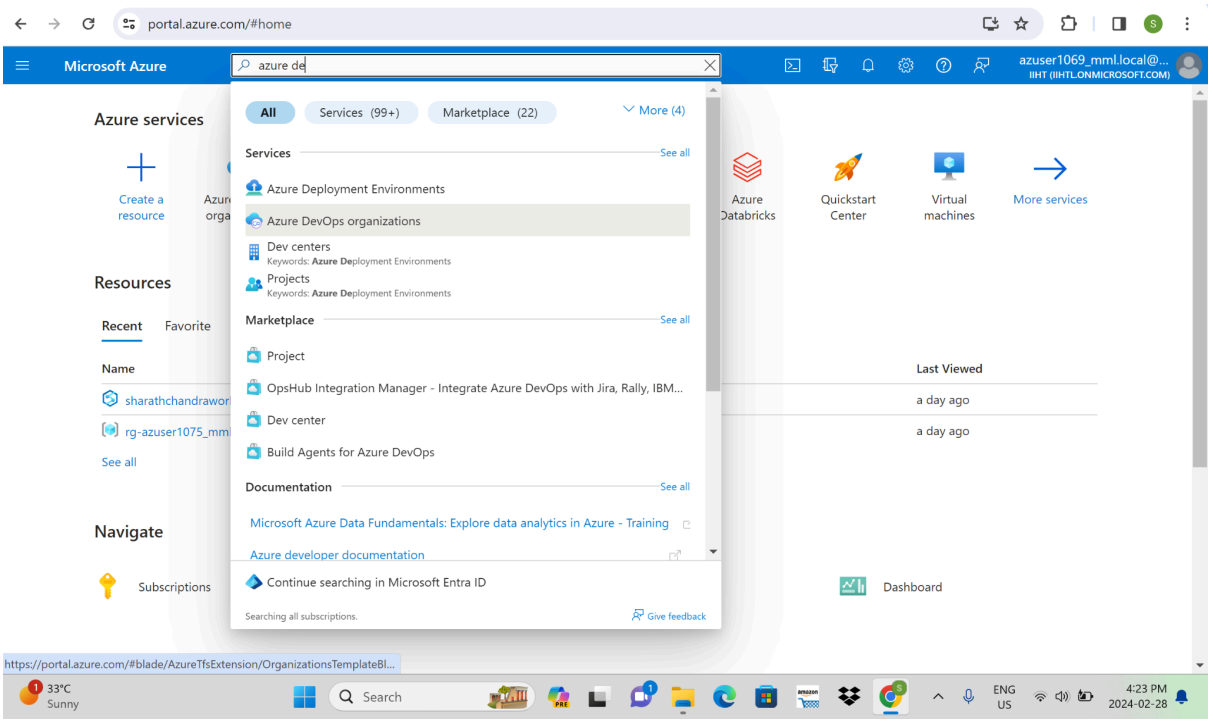
ANSWERS:

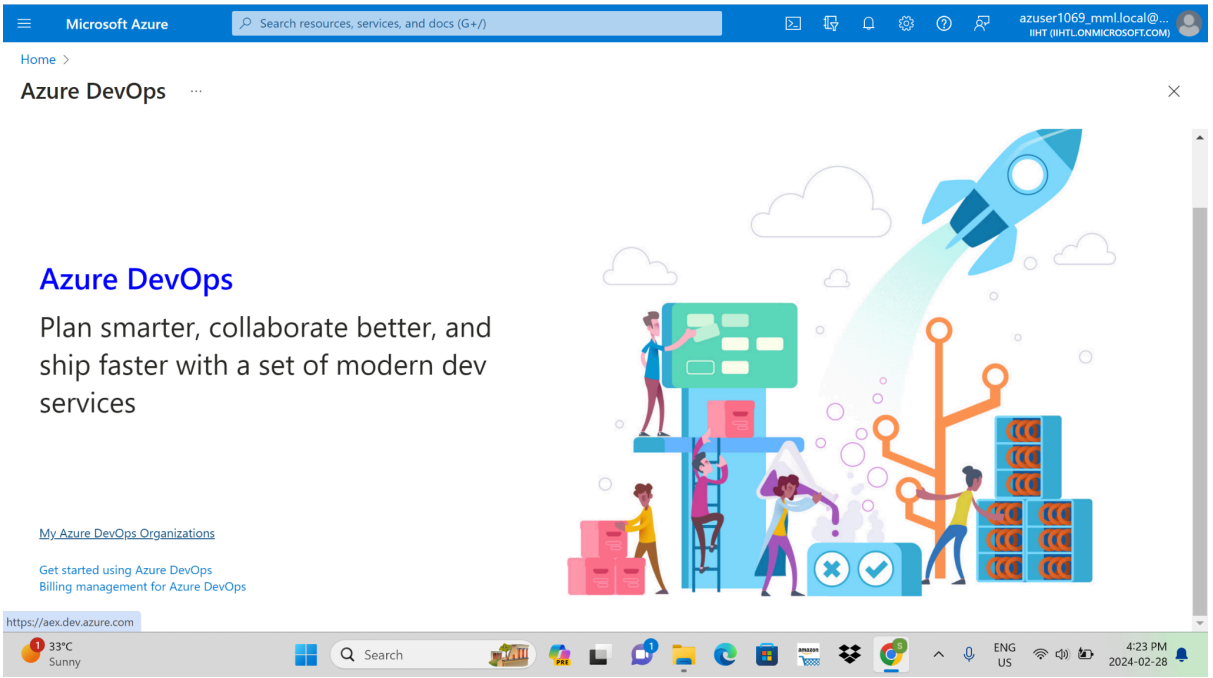
1)AZURE REPOS:

Azure Repos is a version control system that is part of Microsoft's Azure DevOps service family. This service is used to manage source code in software development projects. Azure Repos supports team-based software development processes, providing developers with a platform to collaborate, track code changes, and manage version history.
—>For configuring azure devops git repository Git Push Command is used. The git push command uploads content from a local repository to a remote repository. Pushing refers to the process of moving commits from one repository to another.

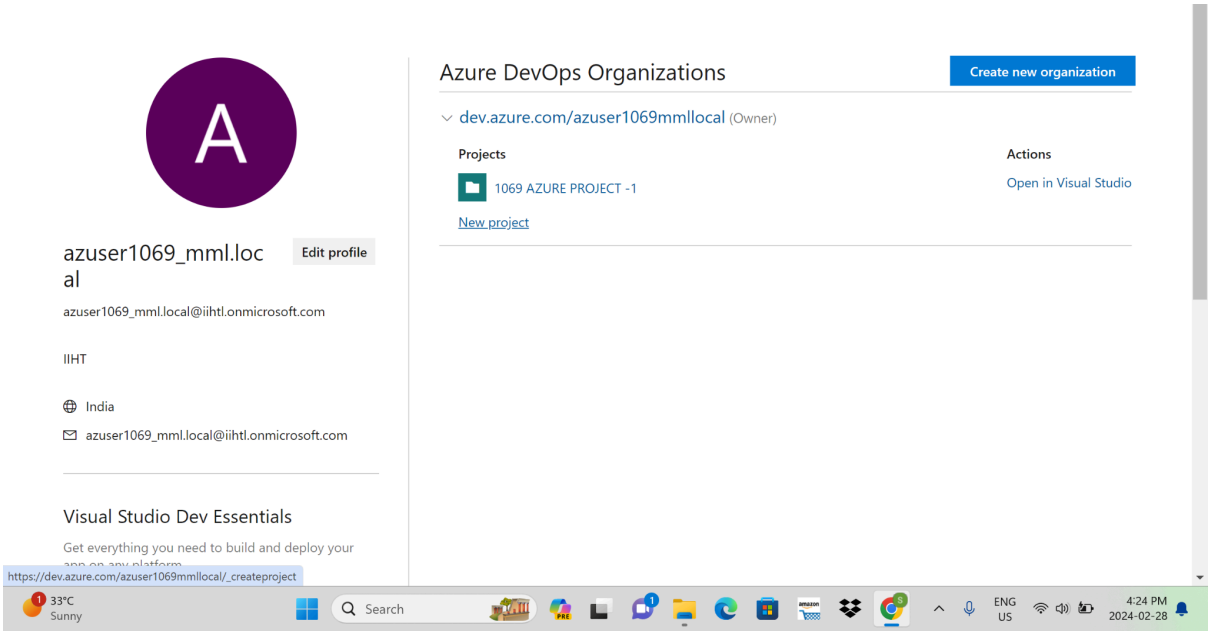
IMPLEMENTATION:

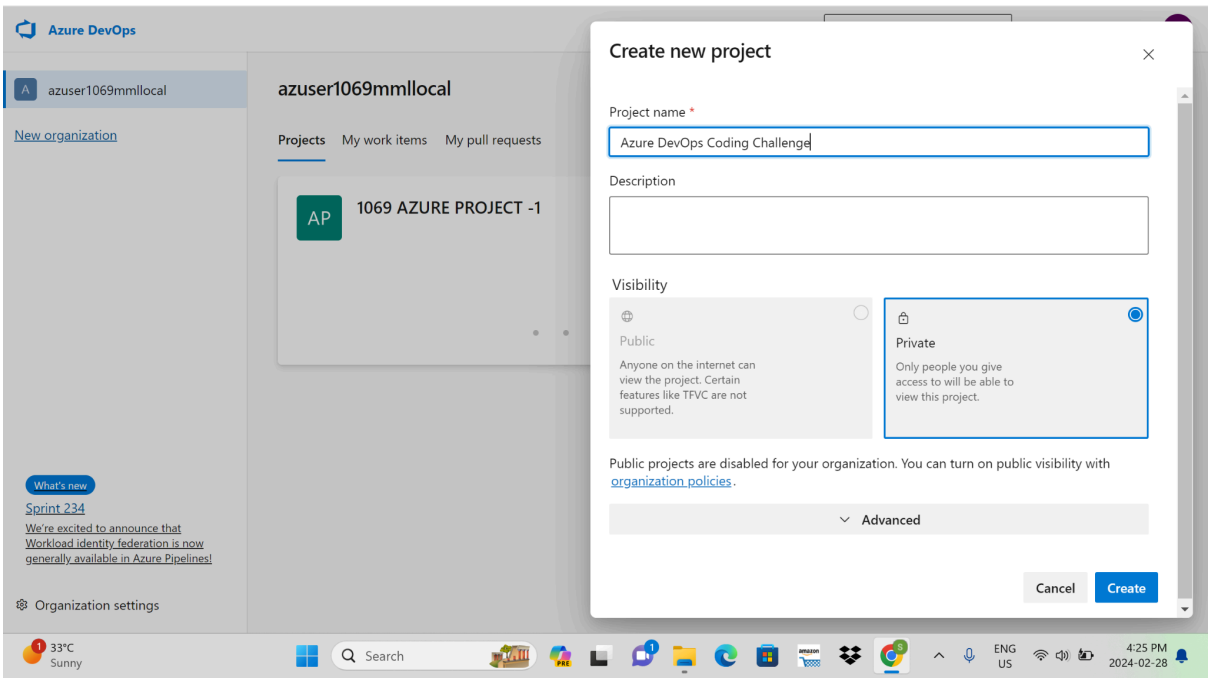
i) firstly login the azure portal and search for azure devops organizations and then click on my azure devops organizations.



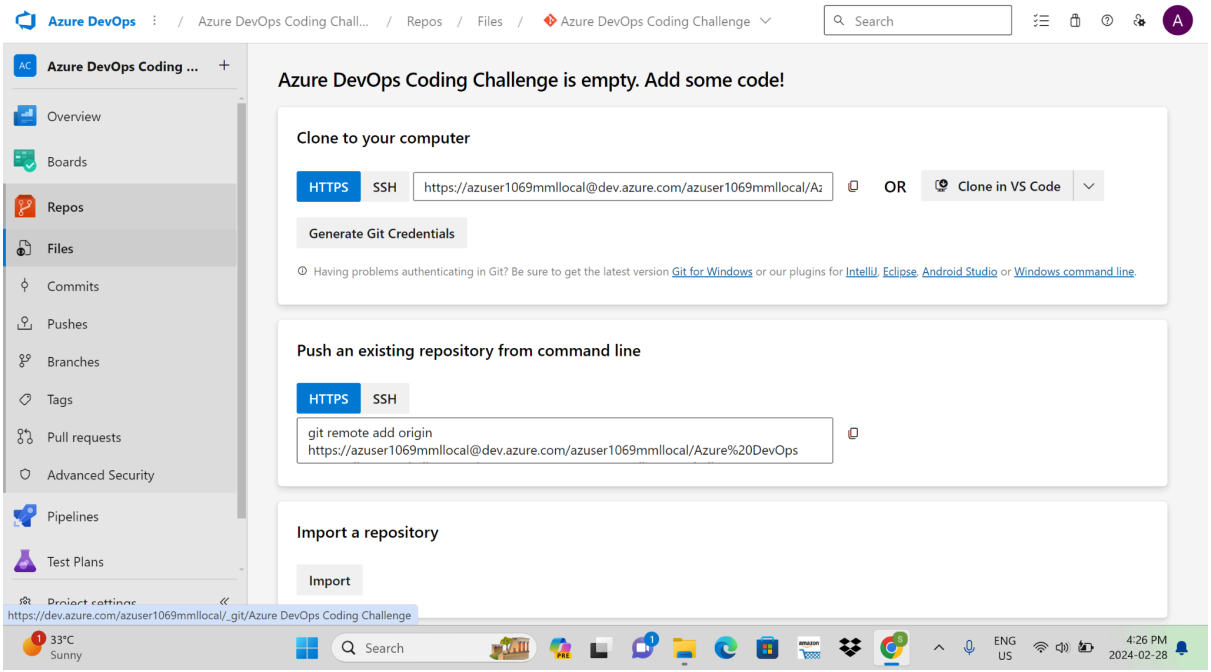


ii)Create a new project in Azure DevOps.



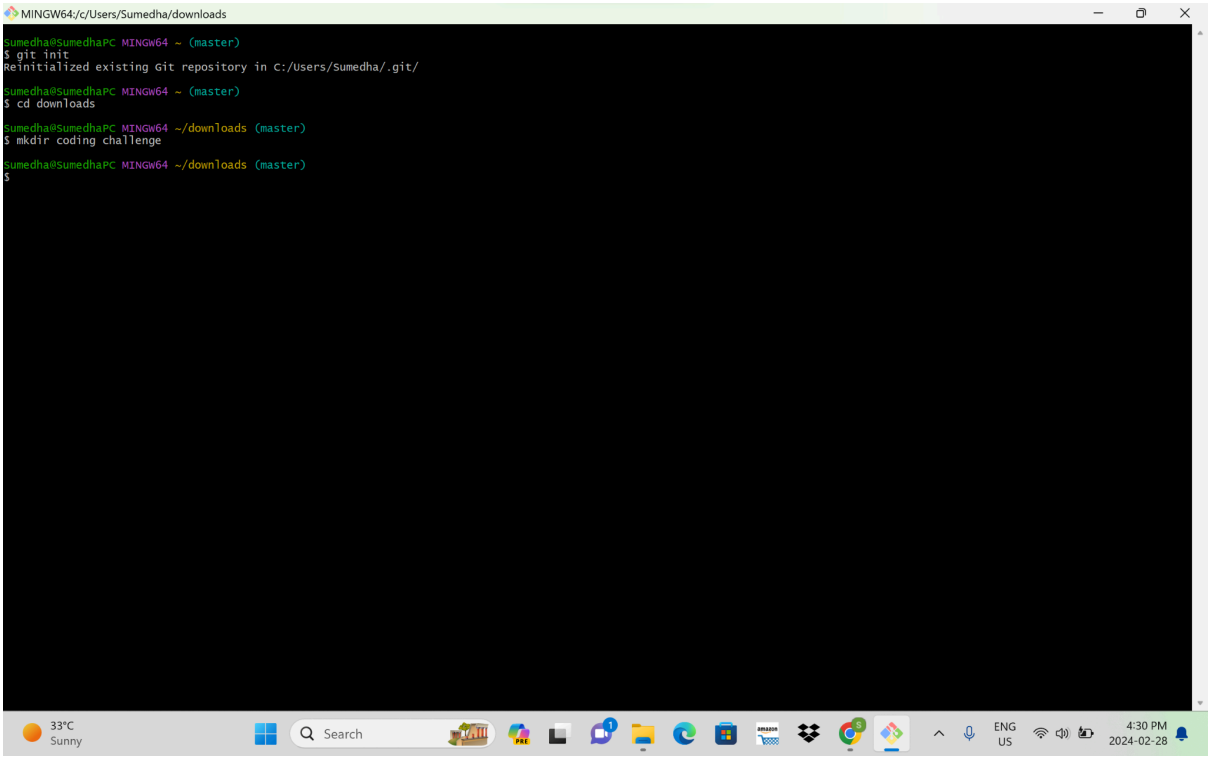


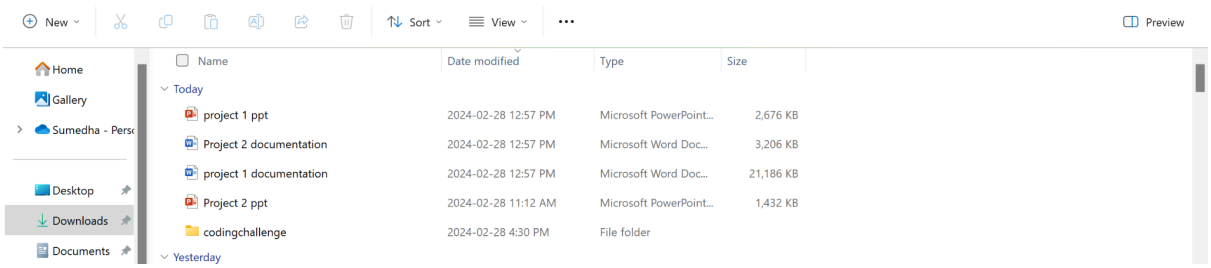
iii) Now Inside project navigate to Repos and create a new Git repository.



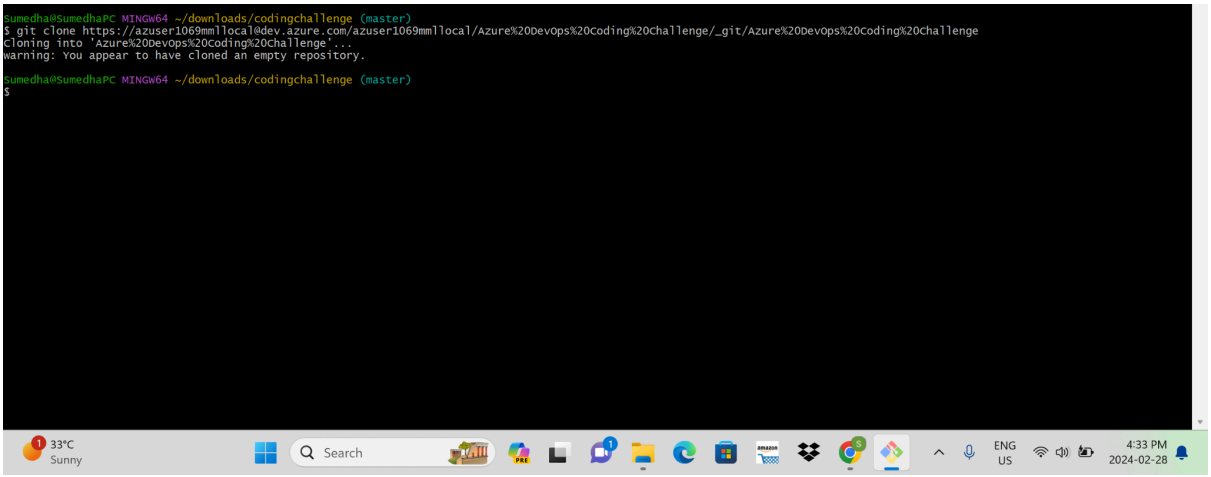
iv) Now open Git on your local machine and clone the Azure DevOps repository to your local environment.

→for that we need to first create a folder in our local

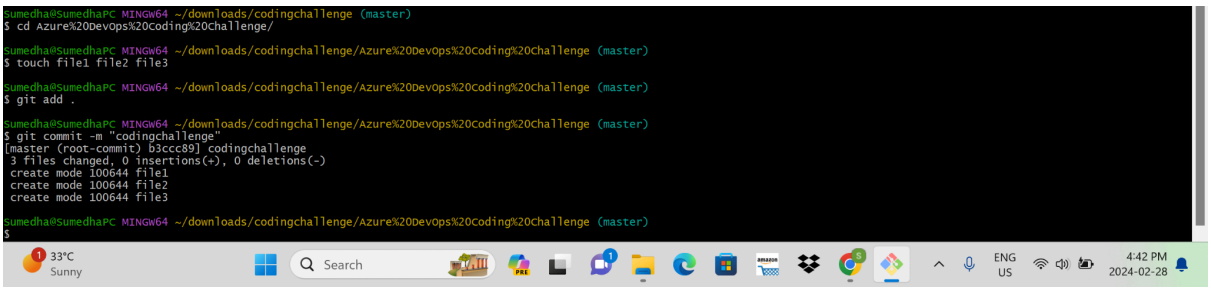




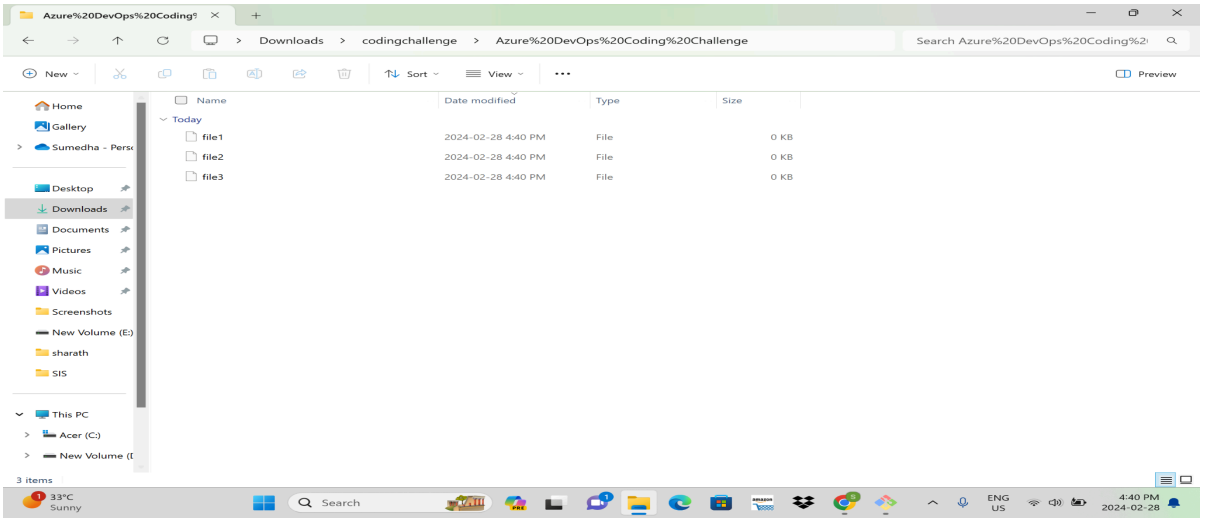
v) Then we need to clone the azure repository into that folder by git clone “that repository path”.



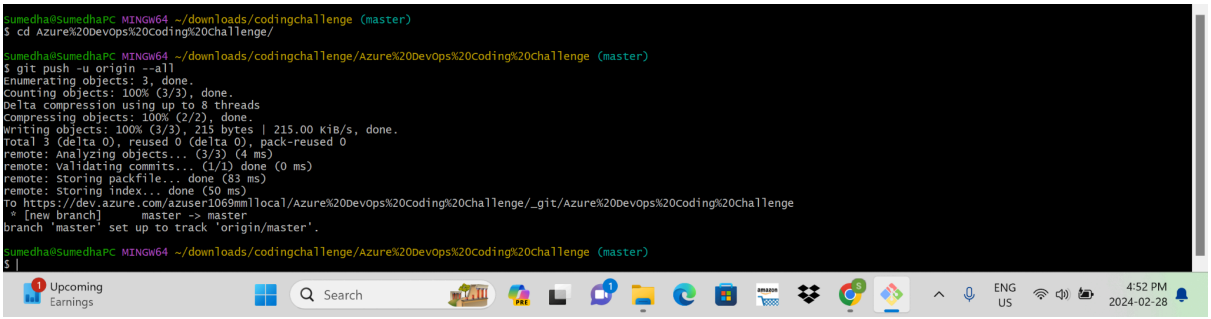
vi) Then we need to add the file into those repository in the local by git touch file will be created in local and by git add.. Files will added



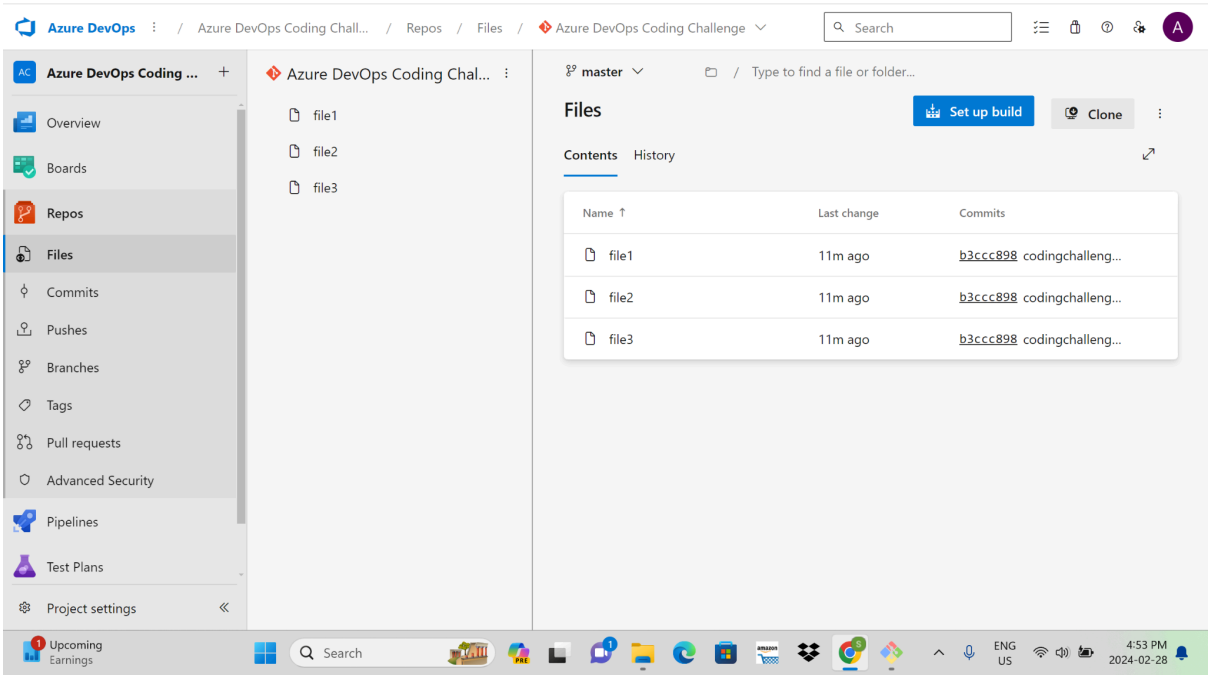
After adding the file in the local :



vii) Then we need to write the command to go to the azure project .
viii)Then we need to push the files into the repository from the local by “git push”.



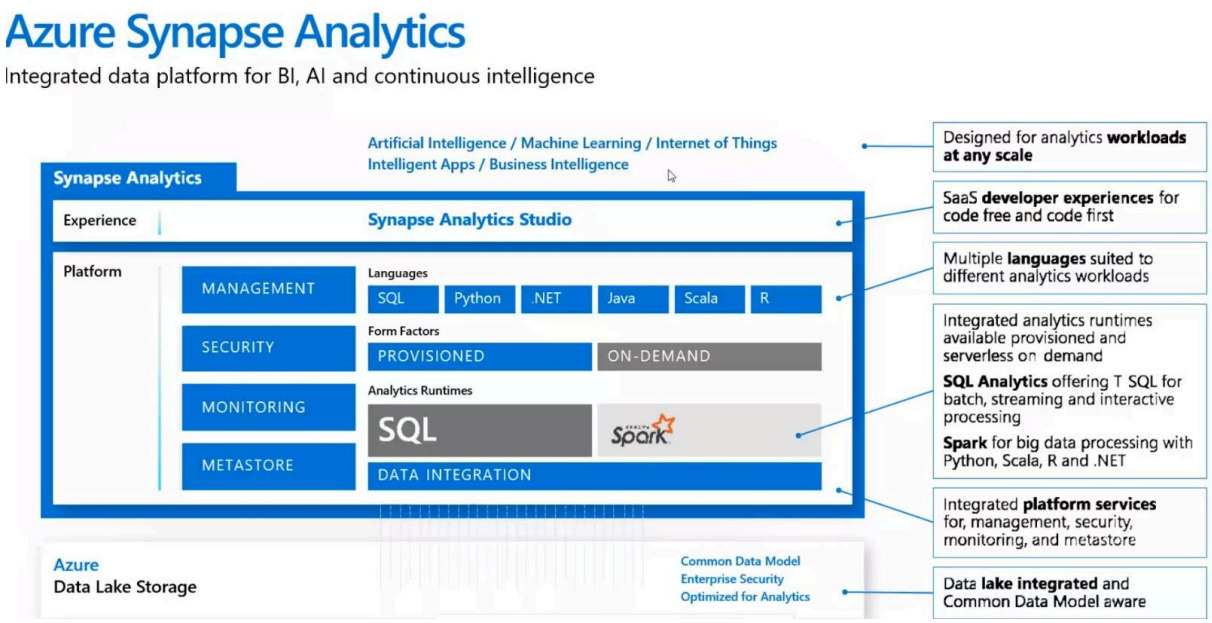
ix)Thereby all the files have been pushed from the local to the azure repository .



2)CI/CD PIPELINE:

Continuous Integration and Continuous Deployment (CI/CD) pipelines automate steps in software delivery, reducing errors and ensuring consistent releases. Automation spans development, testing, production, and monitoring phases, improving code quality, speed, and security. Manual execution is possible, but true value arises from automation.

AZURE SYNAPSE ARCHITECTURE:



Azure Synapse Analytics integrates data warehousing and Big Data analytics on a single cloud platform. Its architecture involves key components:

i)Integration Layer:

—>**SQL Pools:** Processes structured data with optimized analytics.

—>**Apache Spark Pools**: Uses Spark for big data analytics and machine learning.

—>**Data Lake Storage:** Integrates with Azure Data Lake Storage for flexible data storage.

ii)Data Ingestion and Integration:

—>**Azure Data Factory:** Orchestrates data pipelines for ingestion and transformation.

—>**Mapping Data Flows:** Enables visual design and execution of data transformation logic.

iii)Query Optimization:

—>**Query Processing Engine:** Executes SQL queries efficiently across the MPP architecture.

—>**Query Distribution:** Dynamically distributed query execution tasks, optimizing resource utilization.

iv)Security and Governance:

—>**Azure Active Directory Integration:** Manages identity and access control.

—>**Data Encryption:** Secures data at rest and in transit with industry-standard encryption.

—>**Auditing and Monitoring:** Tracks user activities, query performance, and system health.

iv)Data Visualization and Reporting:

—>**Power BI Integration:** Seamlessly connects with Power BI for creating interactive dashboards and reports.