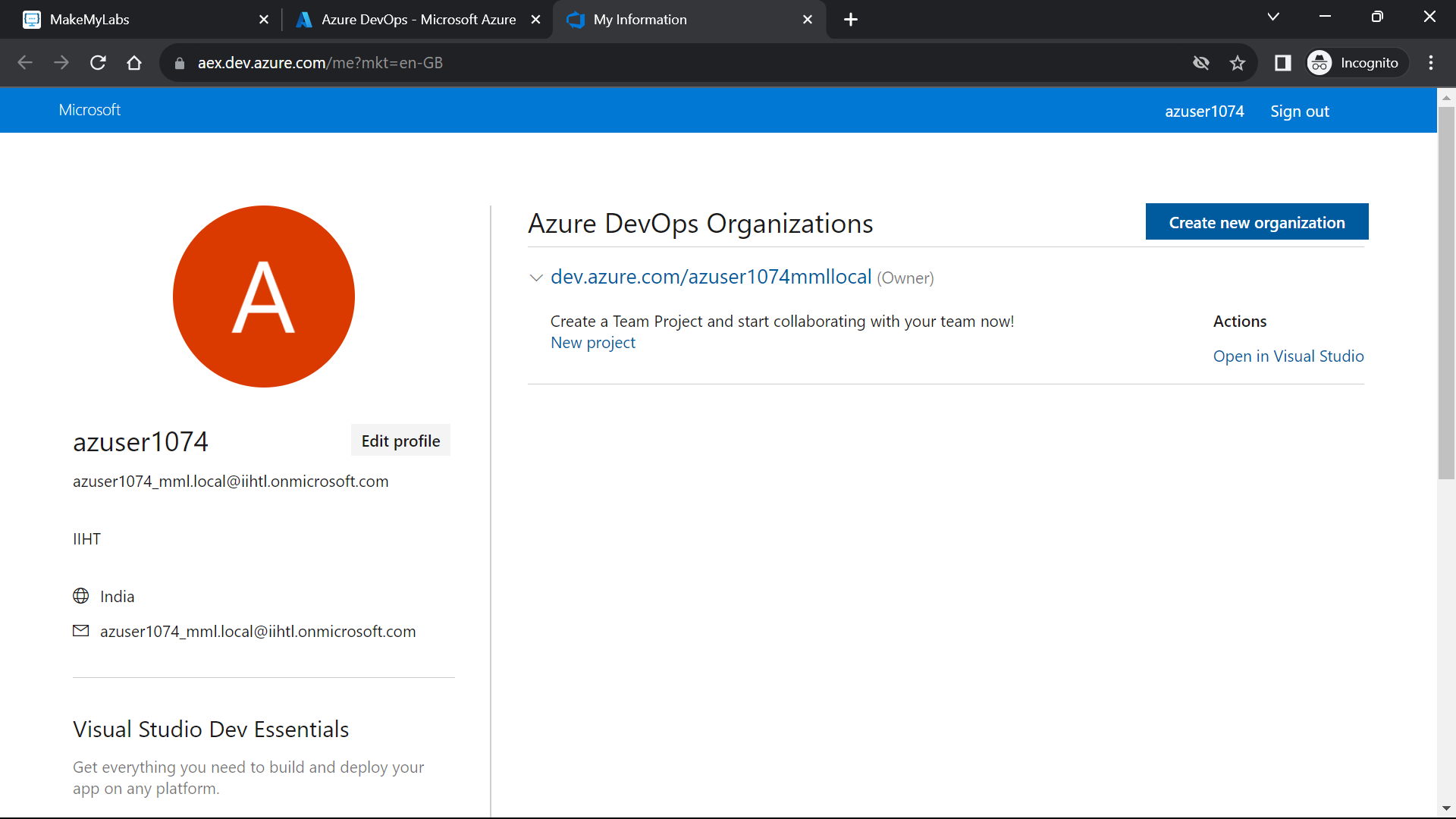
Name – Pujan Rastogi

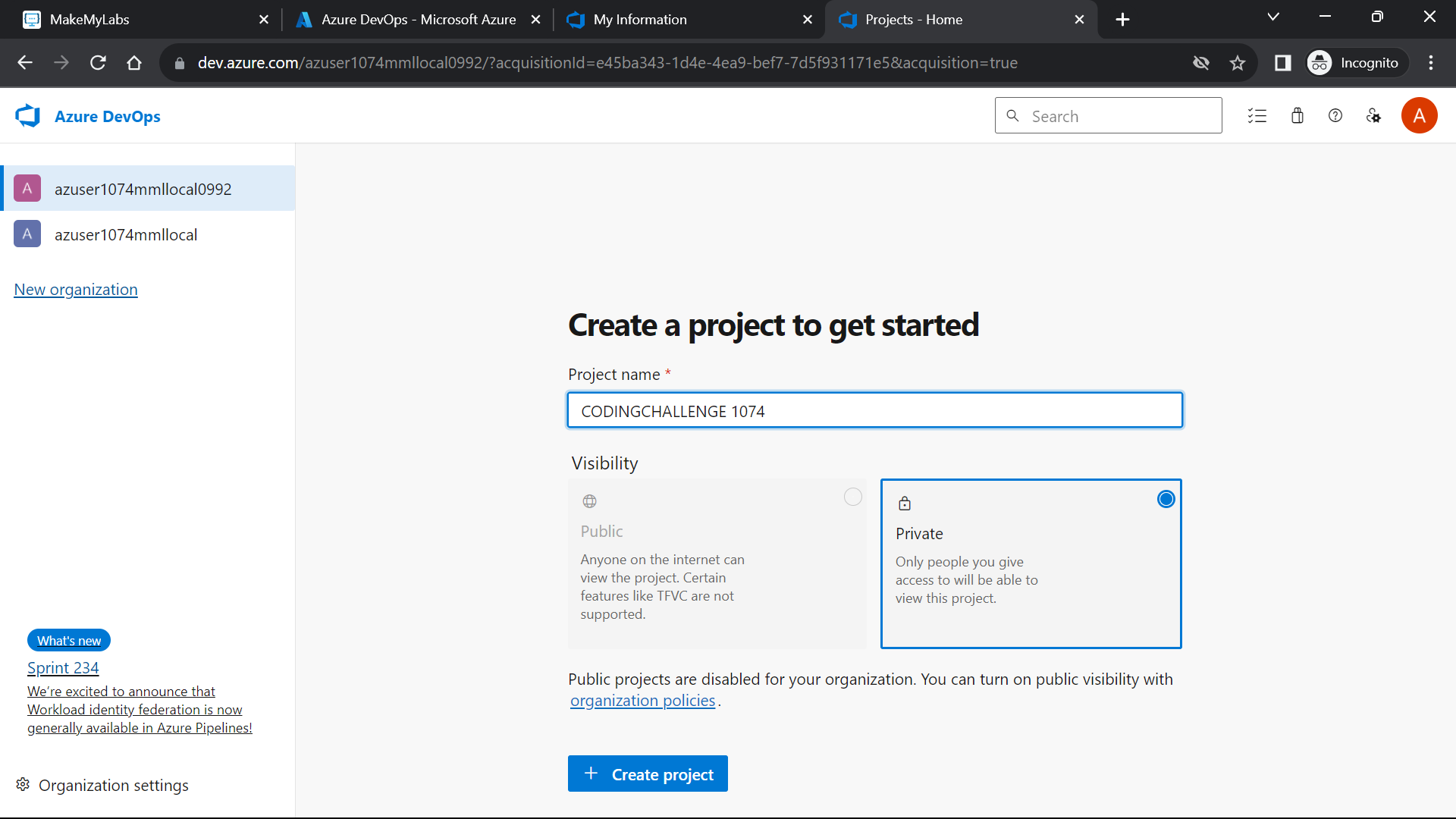
Azure DevOps Coding Challenge 28th feb 2024

**1.Create Azure Devops Environment and configuring Azure Devops Git Repository ,configure on your local git to implement this upload few test files on same.**

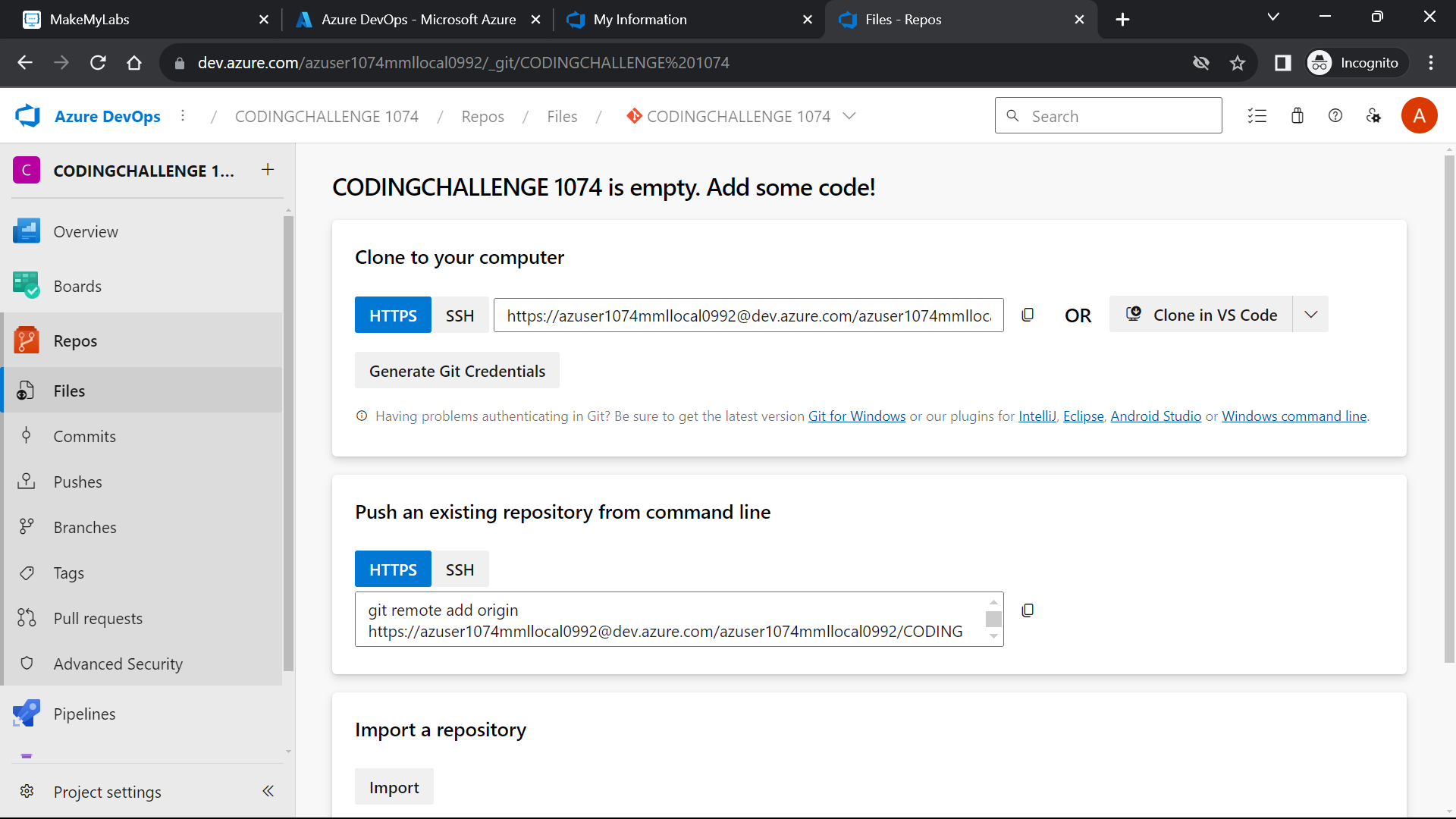
Creating DevOps account in Azure



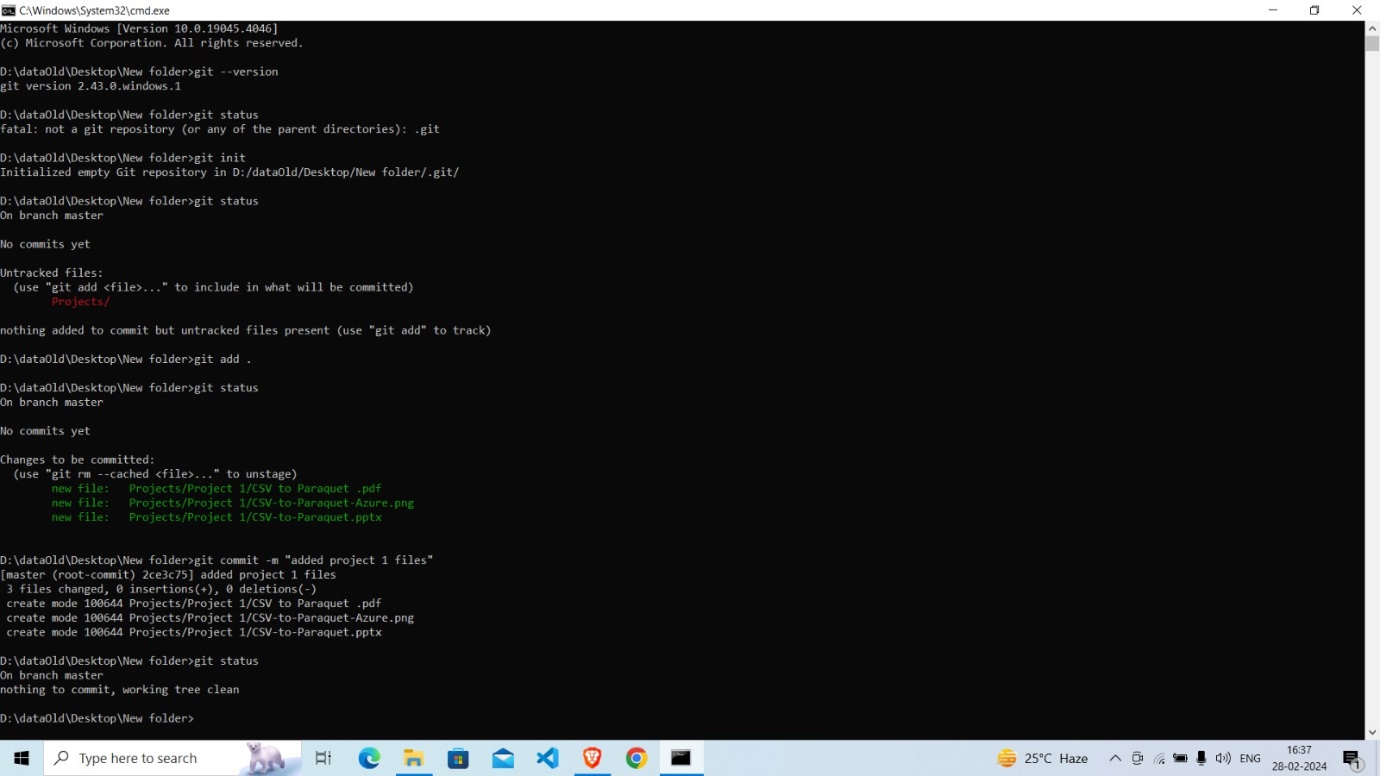
Creating new project

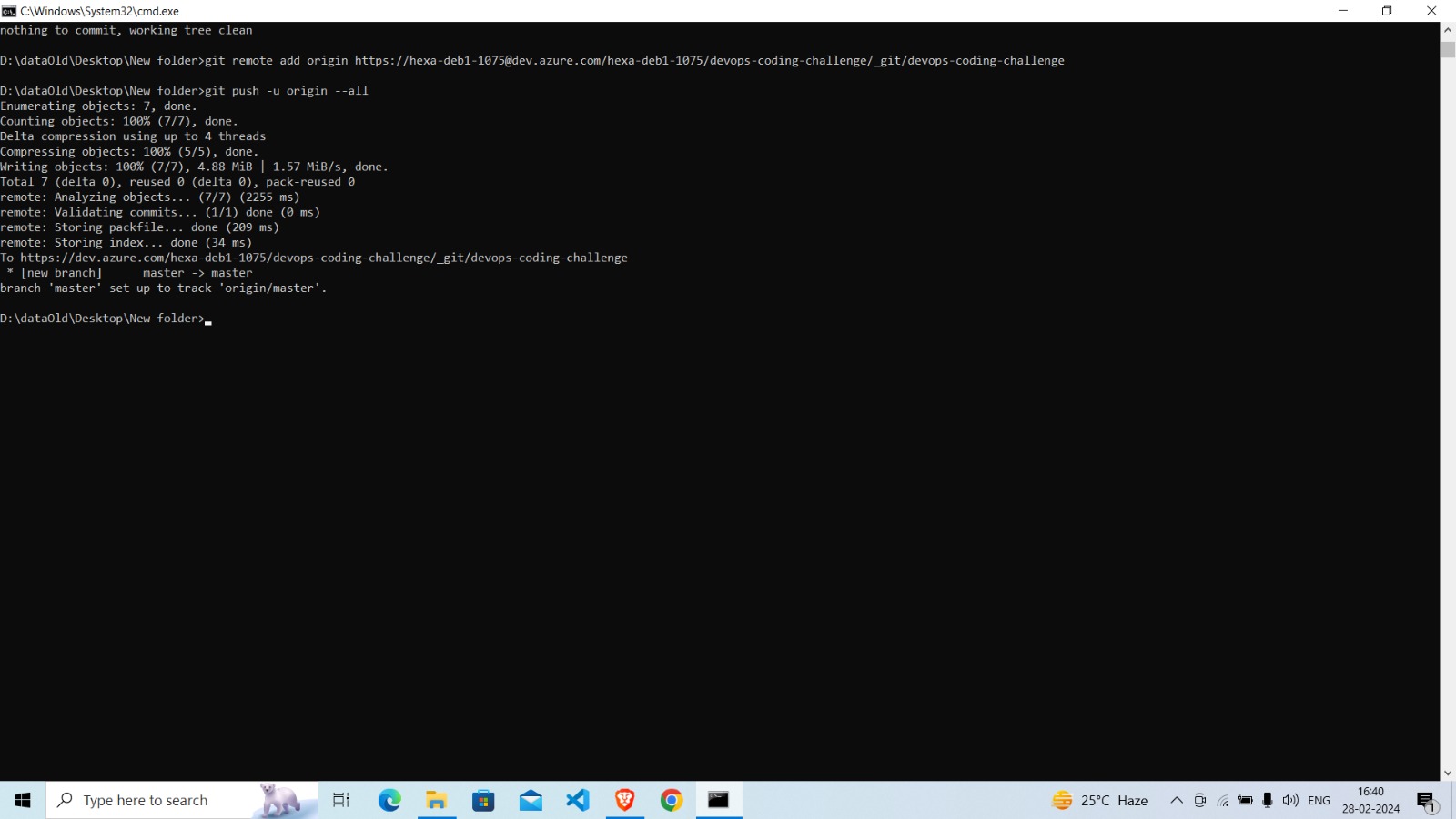


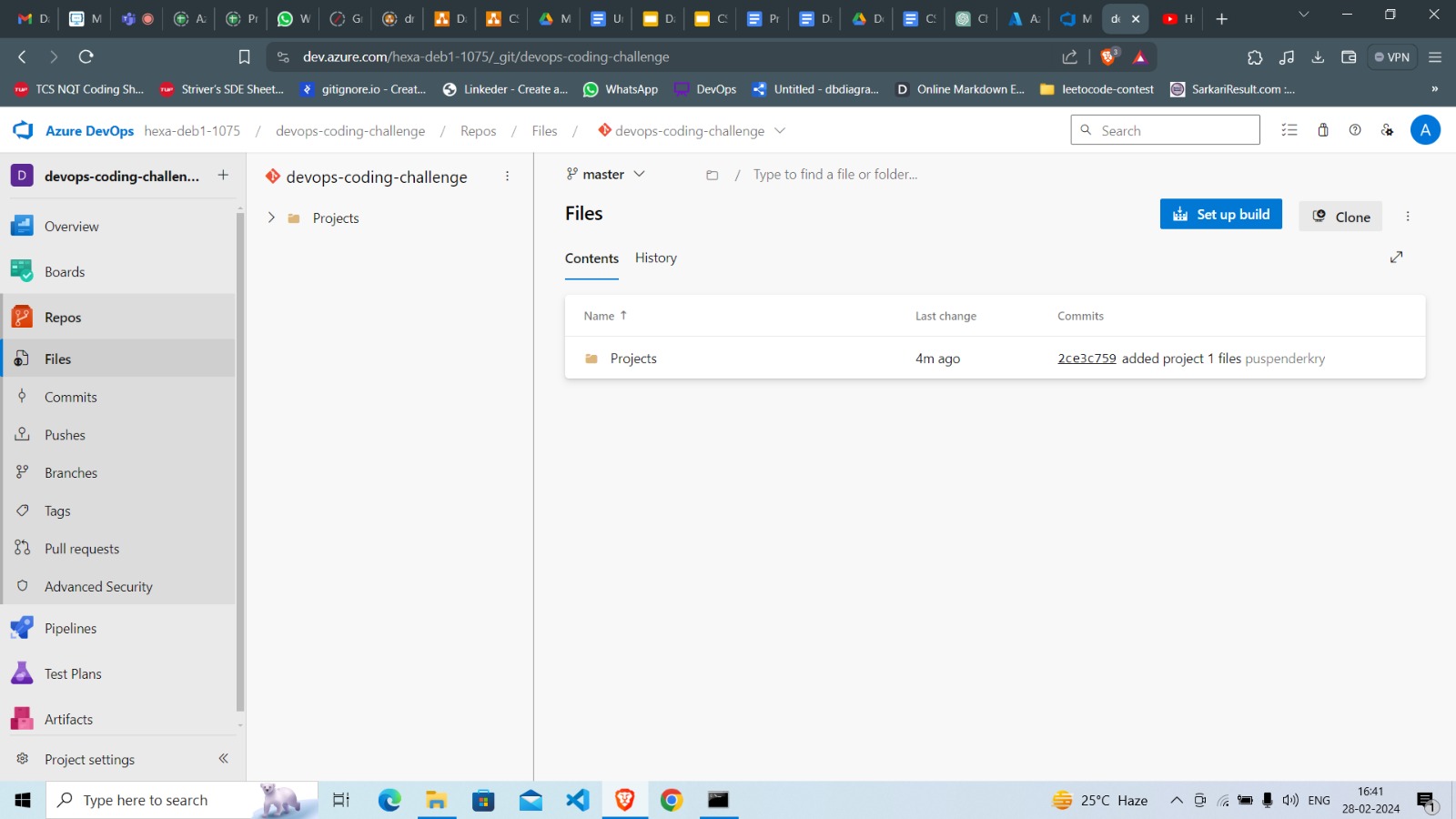
Go to repos and copy the clone to your computer URL.

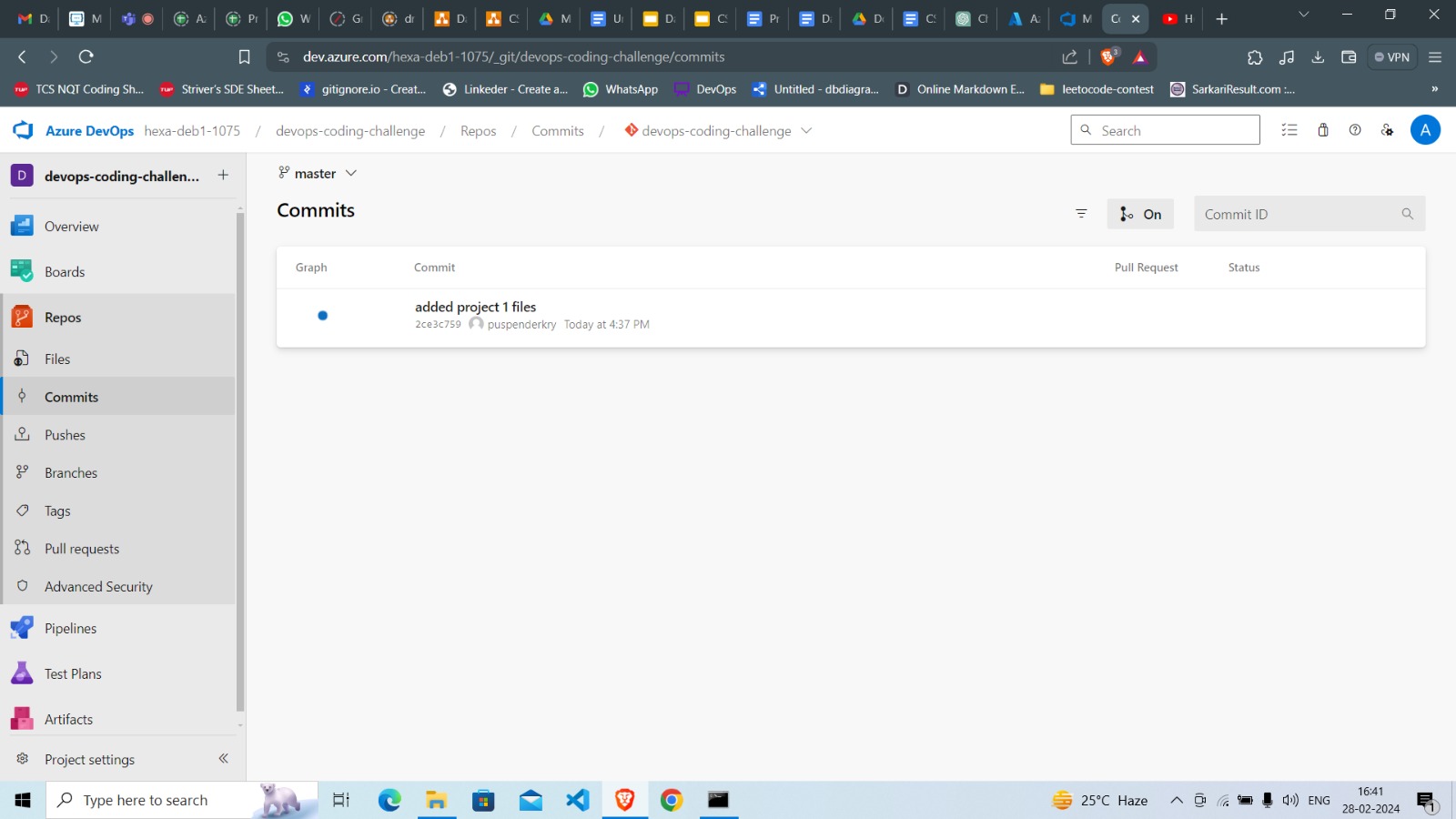


Now clone to git bash and perform the activities.









**2.Leverage the practises of CICD Using azure Dataengineering and explain the architecture of the Azure synpase.**

Continuous Integration and Continuous Deployment (CI/CD) practices are essential in modern software development, including data engineering. Azure provides a set of tools and services to implement CI/CD for data engineering workflows, and Azure Synapse Analytics is a powerful platform for big data and data warehousing.

**CI/CD Practices for Azure Data Engineering:**

**Source Control:** Store code and configurations in a version control system like Azure Repos or GitHub.

Manage changes, track history, and enable collaboration among team members.

**Automated Builds:** Use Azure Pipelines or other CI/CD tools to set up automated builds.

Ensure that code compiles successfully and dependencies are resolved.

**Automated Testing**: Implement unit tests, integration tests, and validation tests.

Automated testing ensures that changes do not introduce regressions.

**Artifact Management:** Use Azure Artifacts or other artifact management tools to store and version artifacts such as packages, libraries, and binaries.

**Deployment Automation:** Automate the deployment process to different environments (dev, test, production).

Use Azure DevOps or other CI/CD tools to orchestrate deployments.

**Infrastructure as Code (IaC):** Define and manage infrastructure using tools like Azure Resource Manager (ARM) templates or Terraform.

Infrastructure changes are versioned and deployed in a consistent manner.

**Monitoring and Logging:** Integrate with Azure Monitor, Azure Log Analytics, and other monitoring tools to gain insights into the performance and health of the data engineering pipelines.

**Azure Synapse Analytics Architecture:**

Azure Synapse Analytics (formerly known as Azure SQL Data Warehouse) is an integrated analyticsservice for big data and data warehousing. Its architecture involves several components:

**SQL Pools:**

Synapse Analytics includes SQL pools that allow you to run analytical queries over large datasets.

These pools can be provisioned or scaled dynamically based on workload requirements.

**Apache Spark Pools:** Synapse also provides Apache Spark pools for big data processing.

These pools can be scaled up or down based on the processing needs.

**Data Movement:** Data can be ingested into Synapse Analytics using services like Azure Data Factory, PolyBase, or direct SQL data loading.

**Metadata Management:** Synapse keeps track of metadata using the SQL Data Warehouse catalog.

Metadata includes information about tables, views, and other database objects.

**Security:** Integration with Azure Active Directory for authentication and authorization.

Role-based access control (RBAC) ensures that users have appropriate permissions.

**Data Integration:** Integration with various data sources and data formats, enabling data engineers to work with diverse datasets.

**Monitoring and Management:** Azure Portal provides a user interface for monitoring and managing Synapse Analytics.

Azure Monitor and Azure Log Analytics can be used for detailed monitoring and logging.

**Integration with Other Azure Services:** Synapse Analytics can be integrated with other Azure services like Azure Data Lake Storage, Azure Blob Storage, and Azure Key Vault.