**Coding Challenge - 2**

**Name:** V Venkata Sri Prasad

**Batch:** Data Engineering

**Question:**

Leverage the practises of CICD Using azure Data engineering and explain the architecture of the Azure synapse.

**Azure Synapse:**

Continuous Integration and Continuous Deployment (CI/CD) practices can indeed be leveraged in Azure Data Engineering workflows, particularly in the context of Azure Synapse Analytics, Microsoft's cloud-based analytics service. Here's how CI/CD can be applied in Azure Data Engineering using Azure Synapse, along with an explanation of the architecture of Azure Synapse.

**What pipelines can do:**

* Ingest data from various data sources
* Process and transform the data
* Save the processed data to a staging location for others to consume

**Data pipeline overview**

Data pipelines in the enterprise can evolve into more complicated scenarios with multiple source systems and supporting various downstream applications.

**Data pipelines provide:**

* Consistency: Data pipelines transform data into a consistent format for users to consume
* Error reduction: Automated data pipelines eliminate human errors when manipulating data
* Efficiency: Data professionals save time spent on data processing transformation. Saving time allows then to focus on their core job function - getting the insight out of the data and helping business makes better decisions

**Azure Synapse Architecture:**

**Management:** Azure Synapse helps organize and control tasks like setting up resources and handling operations to make them easier to manage.

**Security:** It keeps your data safe and follows rules to ensure only authorized people can access it.

**Monitoring:** It watches over everything, checking how well things are running and letting you know if something needs fixing.

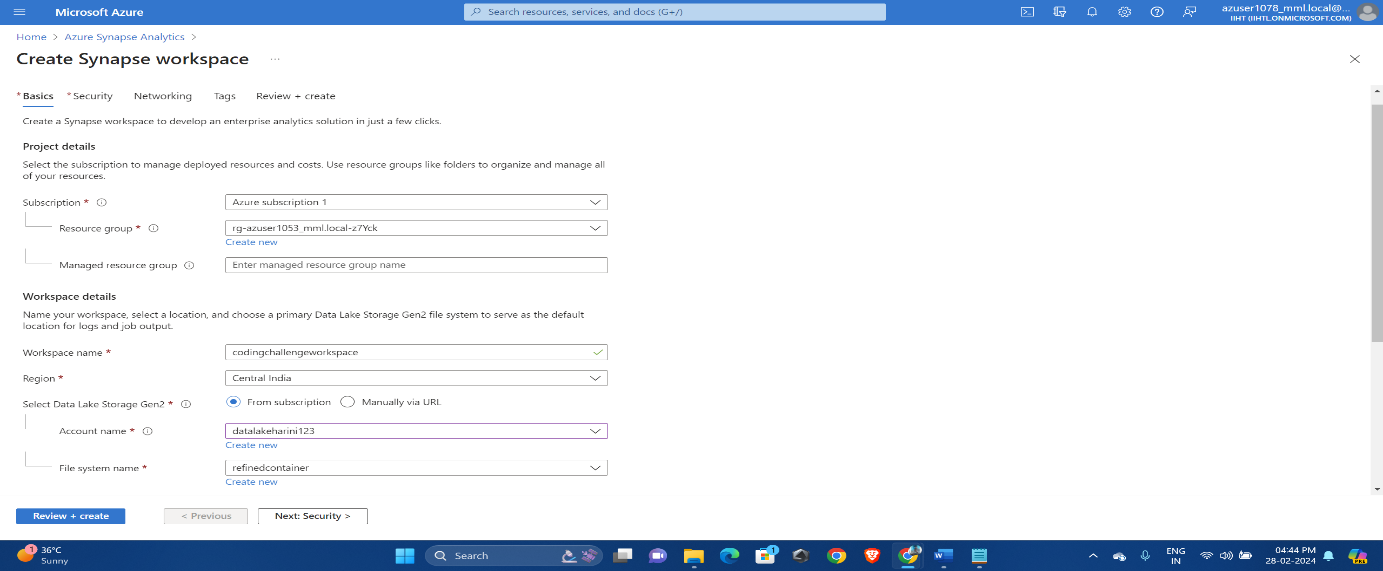
**Metastore:** It keeps track of information about your data, like where it's stored and how it's structured, making it easier to find and use.

Azure Synapse also contains the languages like (SQL , Python, Java, Scala), Form factors, Analytics Runtimes like SQL , Spark and Data Integration

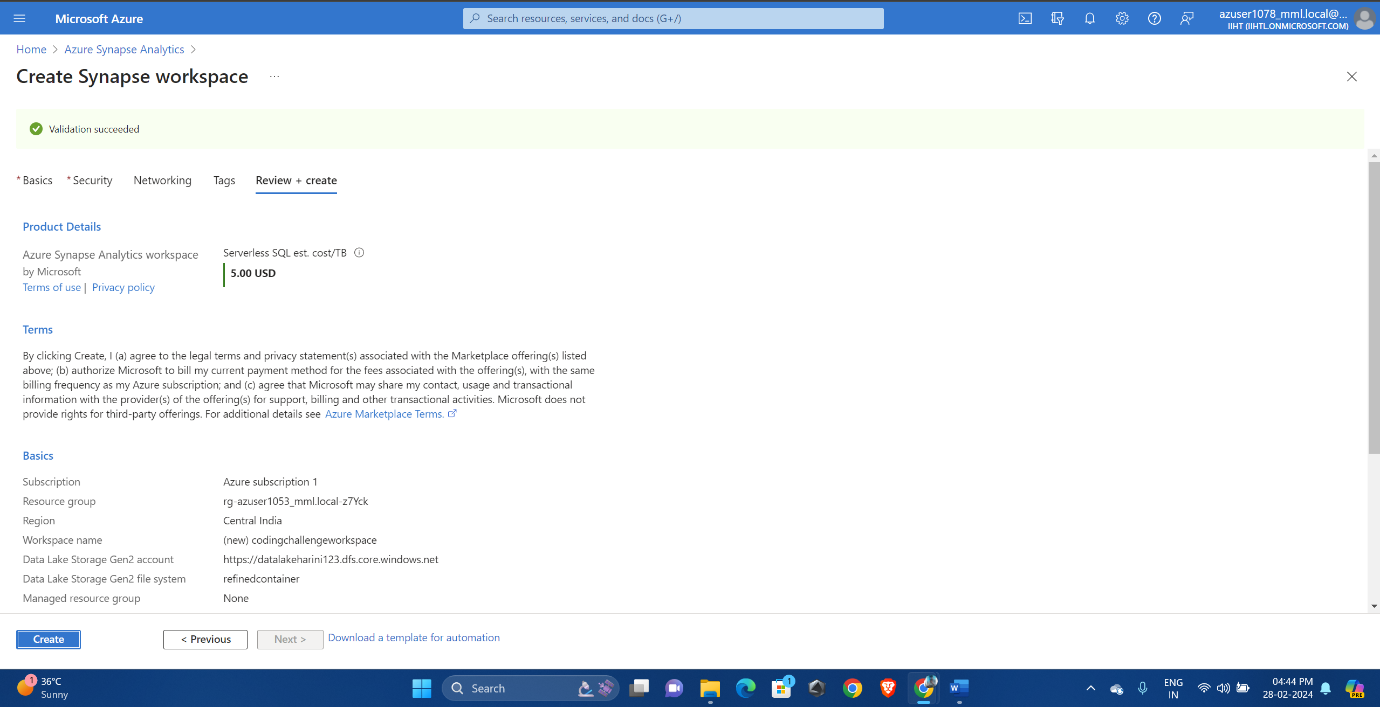
Its also uses Data lake storage to common data model , enterprise security, Optimized for analytics.

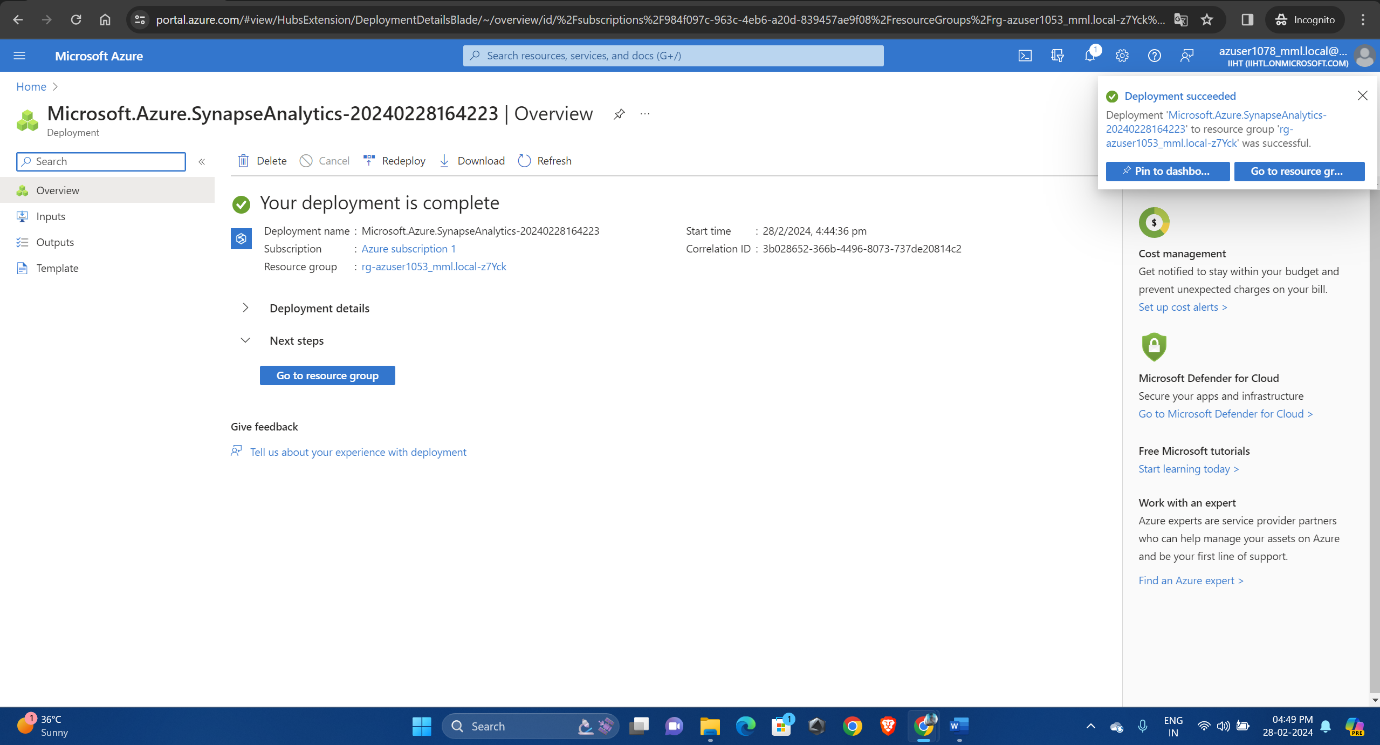
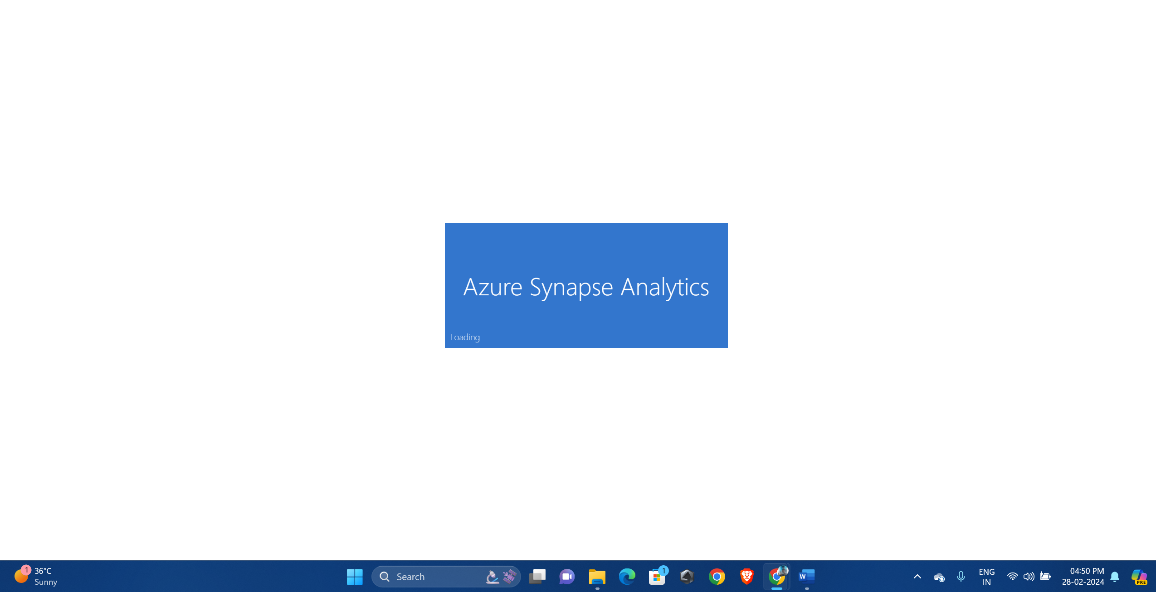
**Creation of Pipelines:**

* Go to Azure Home page ,click on Azure Synapse and click on create.
* Provide the resource group, region, workspace and click on ‘review + create’ .

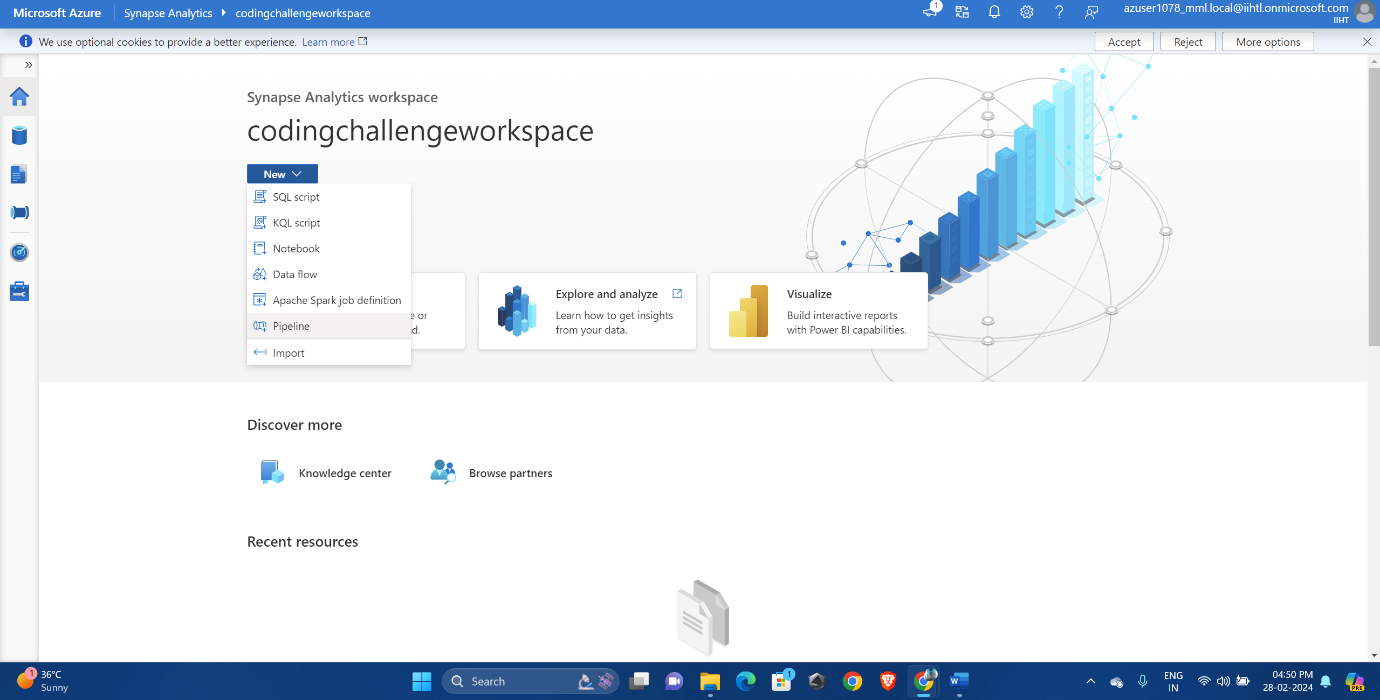


Once validation is successful click on create.

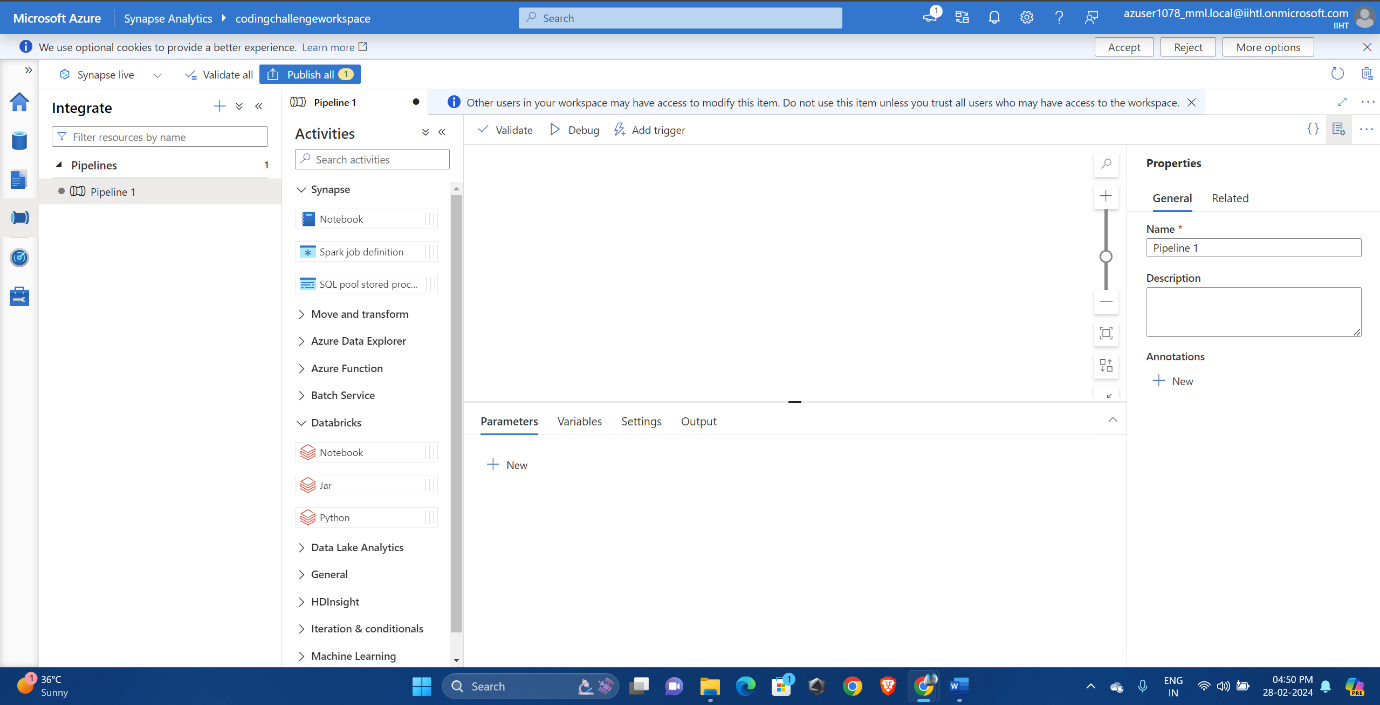


Click on ‘Go to resource group’.  

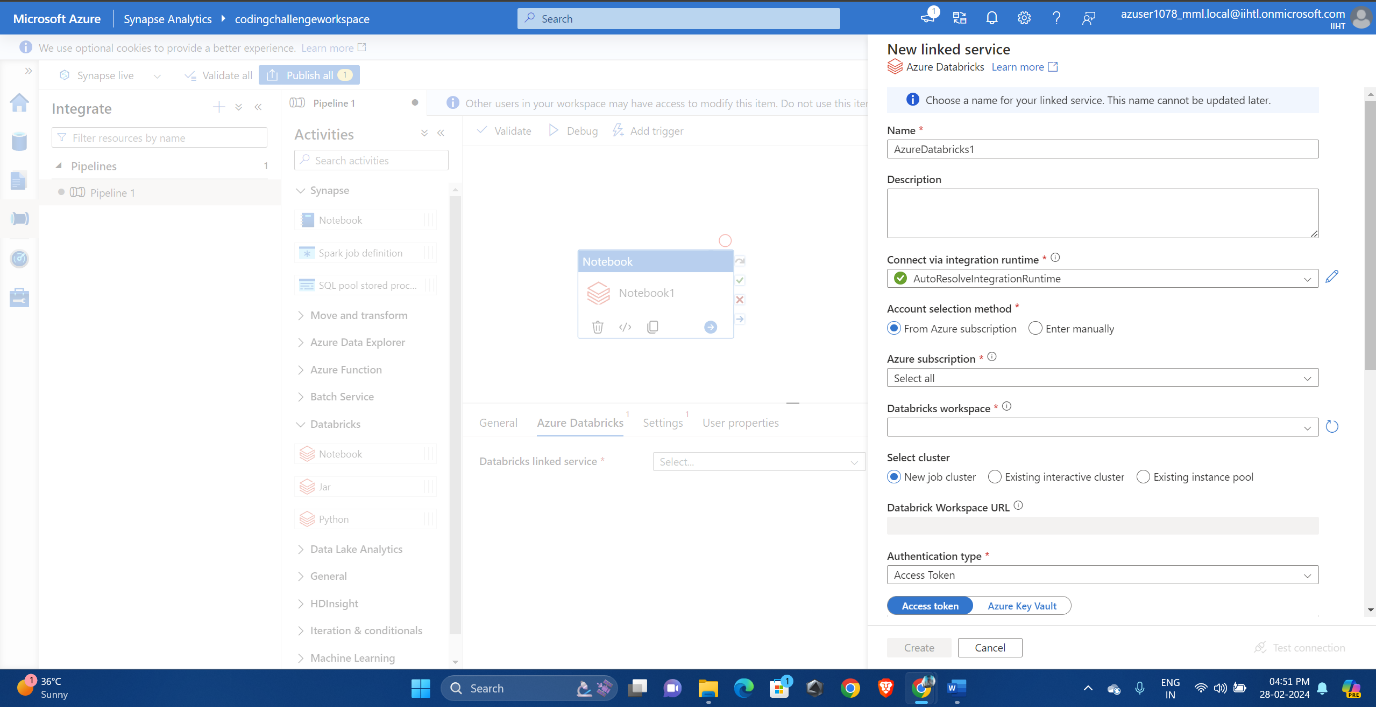
Click on new and select pipeline.



We can see there are several azure services which can be linked and connected through pipelines.



Provide proper paths and access tokens to which the pipelines need to be connect.



So, the pipeline connecting to multiples services is now created.

