**NAME :- ANIKET SANJAYKUMAR BIYANI**

**DATA ENGINEERING BATCH -1**

**AZURE DATABRICKS CODING ASSESMENT**

**Q3.Execute & explain, Azure datafactory and its copy activity.**

* Azure Data Factory is a cloud-based data integration service that allows you to create, schedule, and manage data pipelines for moving and transforming data across various data sources and destinations.
* It enables organizations to build scalable ETL (Extract, Transform, Load) and ELT (Extract, Load, Transform) workflows to ingest, transform, and process data for analytics, reporting, and other data-driven tasks.
* ADF provides a visual interface for building and monitoring data pipelines without writing complex code.
* Azure Data Factory provides a platform to create, schedule, orchestrate, and monitor data workflows. Here's an overview of its main features:
* **Data Integration**: ADF supports data integration from various sources such as Azure Blob Storage, Azure SQL Database, Azure Data Lake Storage, SQL Server, Oracle, Salesforce, and more.
* **Data Transformation**: ADF allows you to transform data using mapping data flows or custom activities. You can perform data cleansing, aggregation, enrichment, and more.
* **Monitoring and Management**: ADF provides monitoring and management capabilities to monitor pipeline runs, track data lineage, manage triggers, and handle errors and alerts.
* **Integration with Other Azure Services**: ADF integrates seamlessly with other Azure services such as Azure Synapse Analytics, Azure Databricks, Azure Machine Learning, etc., allowing you to leverage the full capabilities of the Azure ecosystem.

Azure Data Factory provides a visual interface for constructing data pipelines. Here are some key concepts:

**Pipelines**: Pipelines are a logical grouping of activities that perform a sequence of data processing or data movement tasks. These tasks can include data ingestion, transformation, and publishing.

**Activities**: Activities are the processing steps within a pipeline. There are different types of activities such as data movement, data transformation, control flow, and data flow activities.

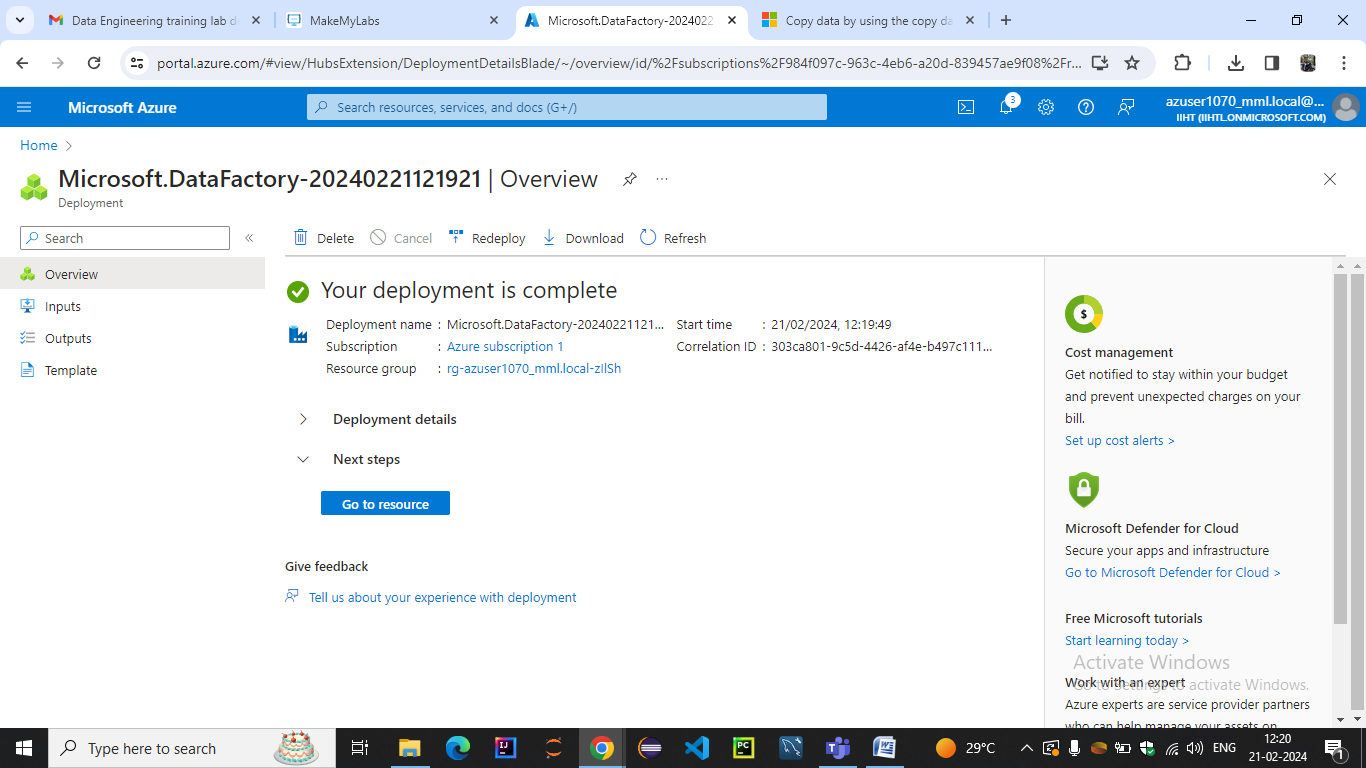
**Datasets**: Datasets represent the data structures within data stores. They define the schema and location of the data to be processed or moved.

**Linked Services**: Linked services are connection objects that define the connection information to external data stores or services. They are used by activities to interact with the data stores.

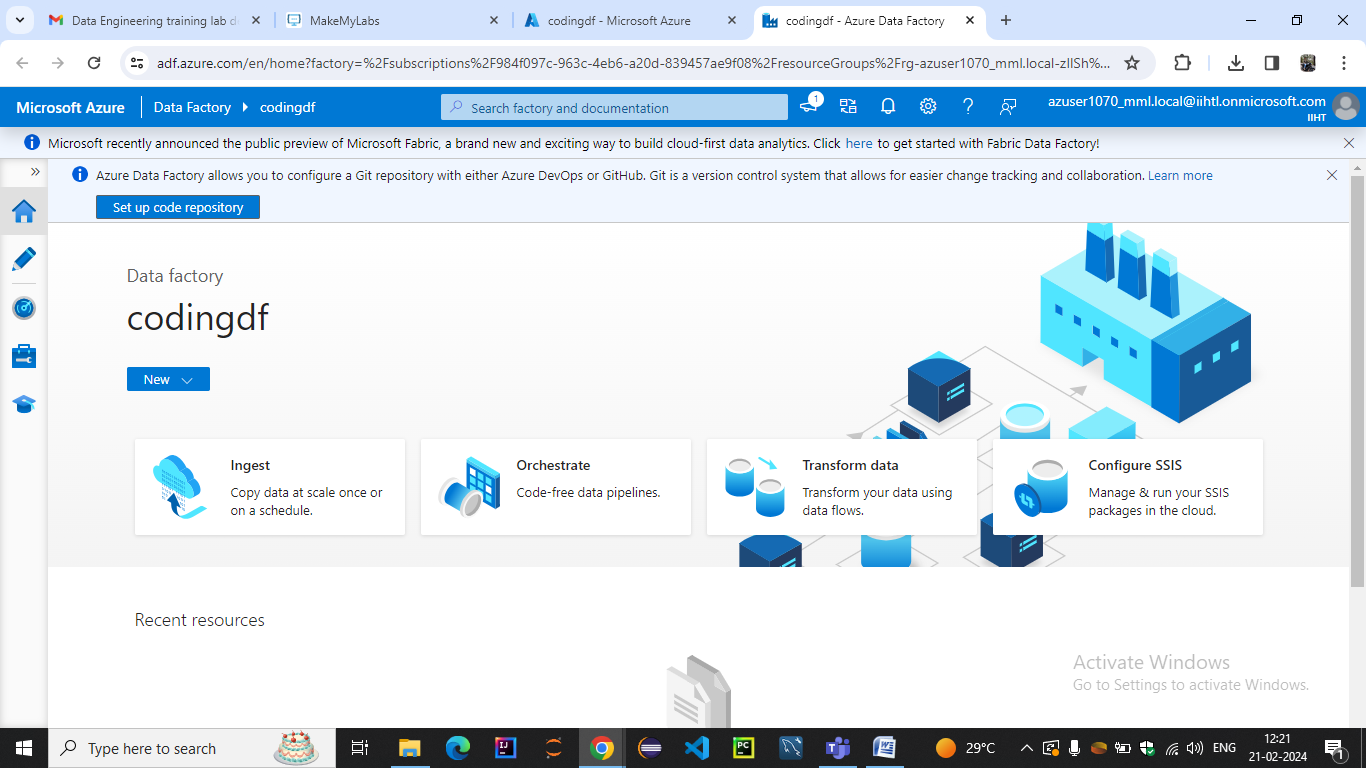
* One of the key components of Azure Data Factory is the Copy Activity.
* Copy Activity is a built-in activity in Azure Data Factory that enables data movement between supported source and destination data stores.
* It supports a wide range of data stores including Azure Blob Storage, Azure SQL Database, Azure Data Lake Storage, Amazon S3, SQL Server, Oracle, and more
* Copy Activity in Azure Data Factory provides a simple and efficient way to move data between different data stores, whether they are on-premises or in the cloud.
* It abstracts away the complexities of data movement, allowing developers and data engineers to focus on defining data pipelines and integrating data workflows without worrying about the underlying infrastructure.
* With features like performance optimization, fault tolerance, and monitoring capabilities, Copy Activity streamlines the process of data migration, replication, and synchronization in Azure Data Factory.

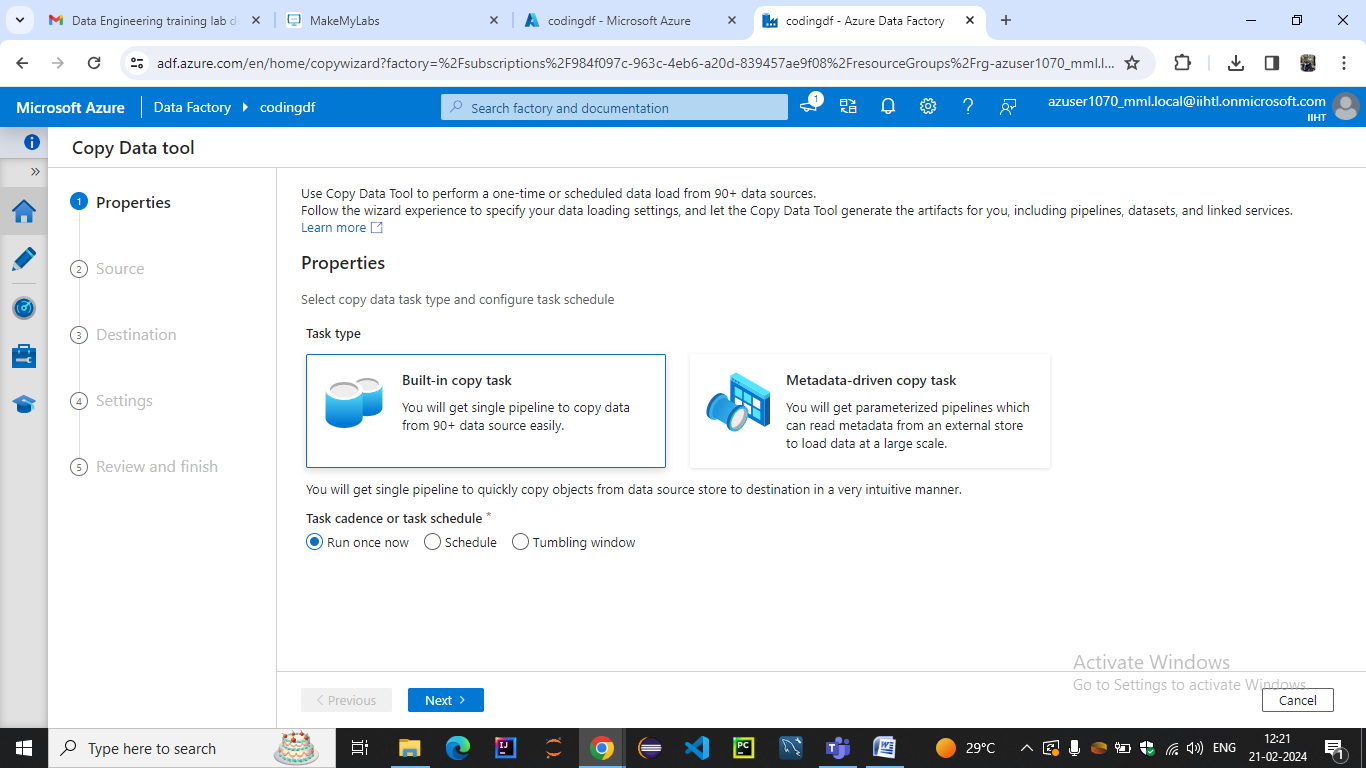
**Use the copy data tool to copy data**

The steps below will walk you through how to easily copy data with the copy data tool in Azure Data Factory.

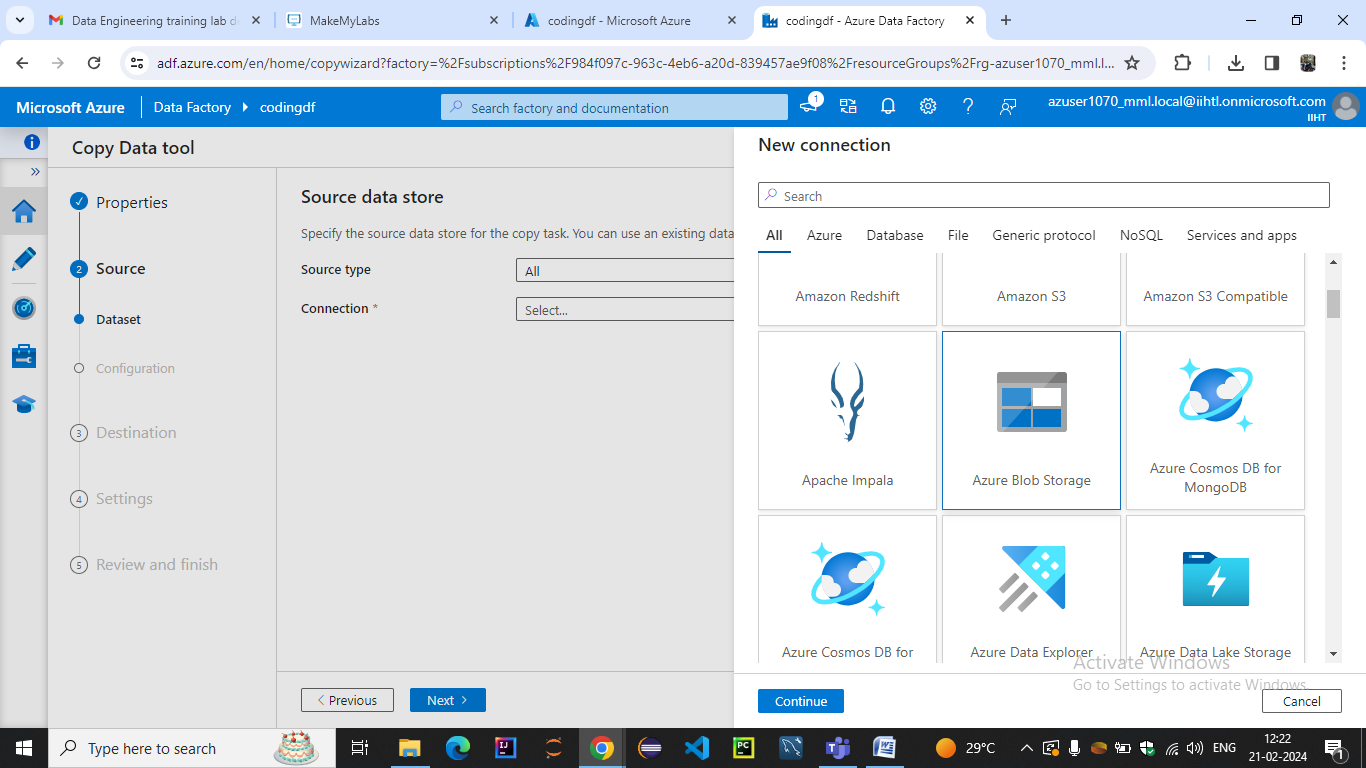


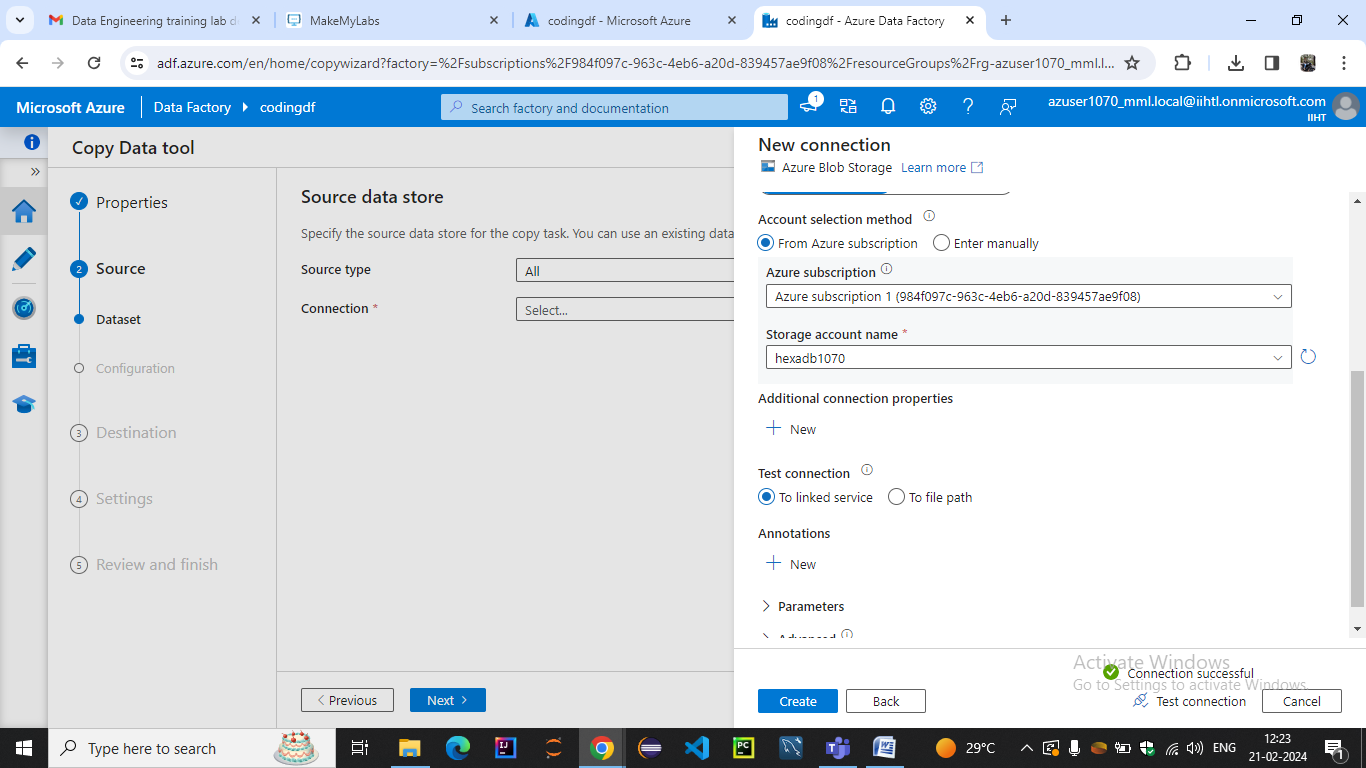
**Step 1: Start the copy data Tool**

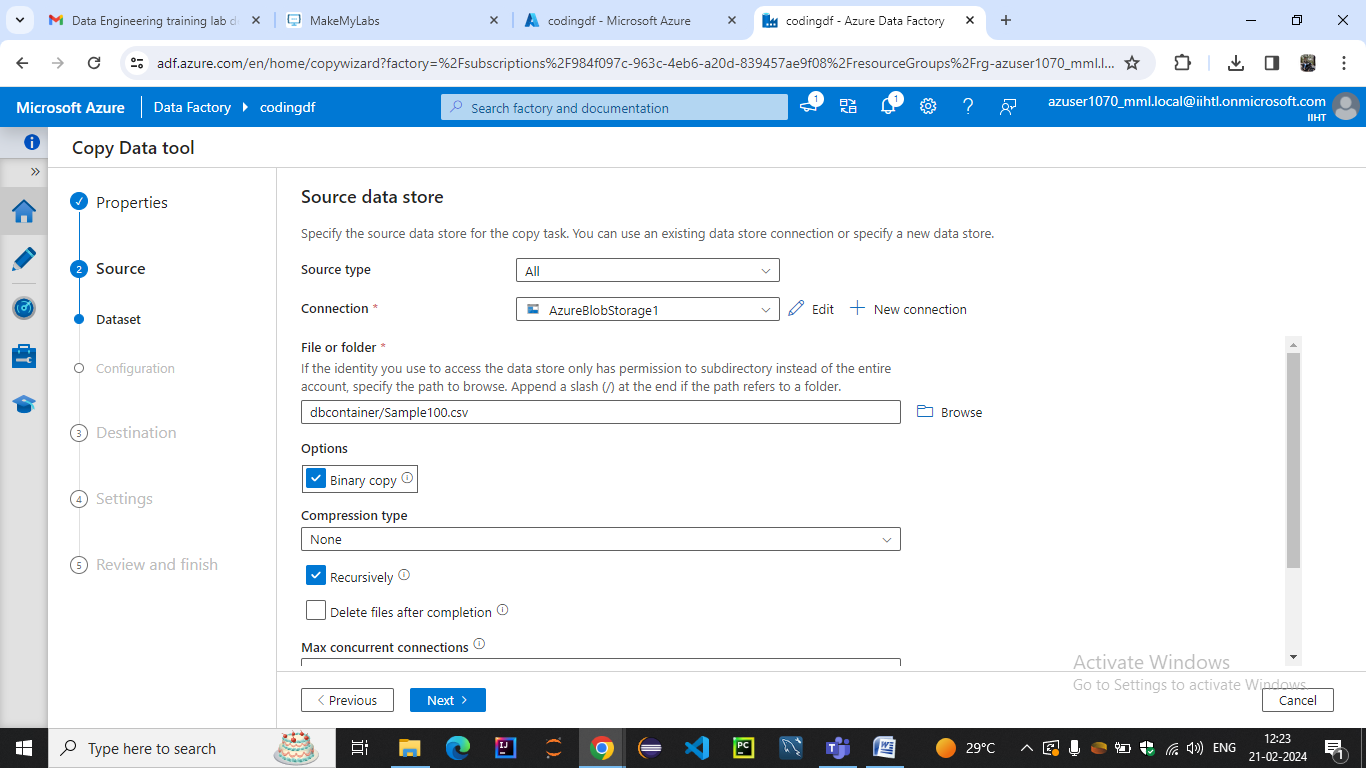
****

****

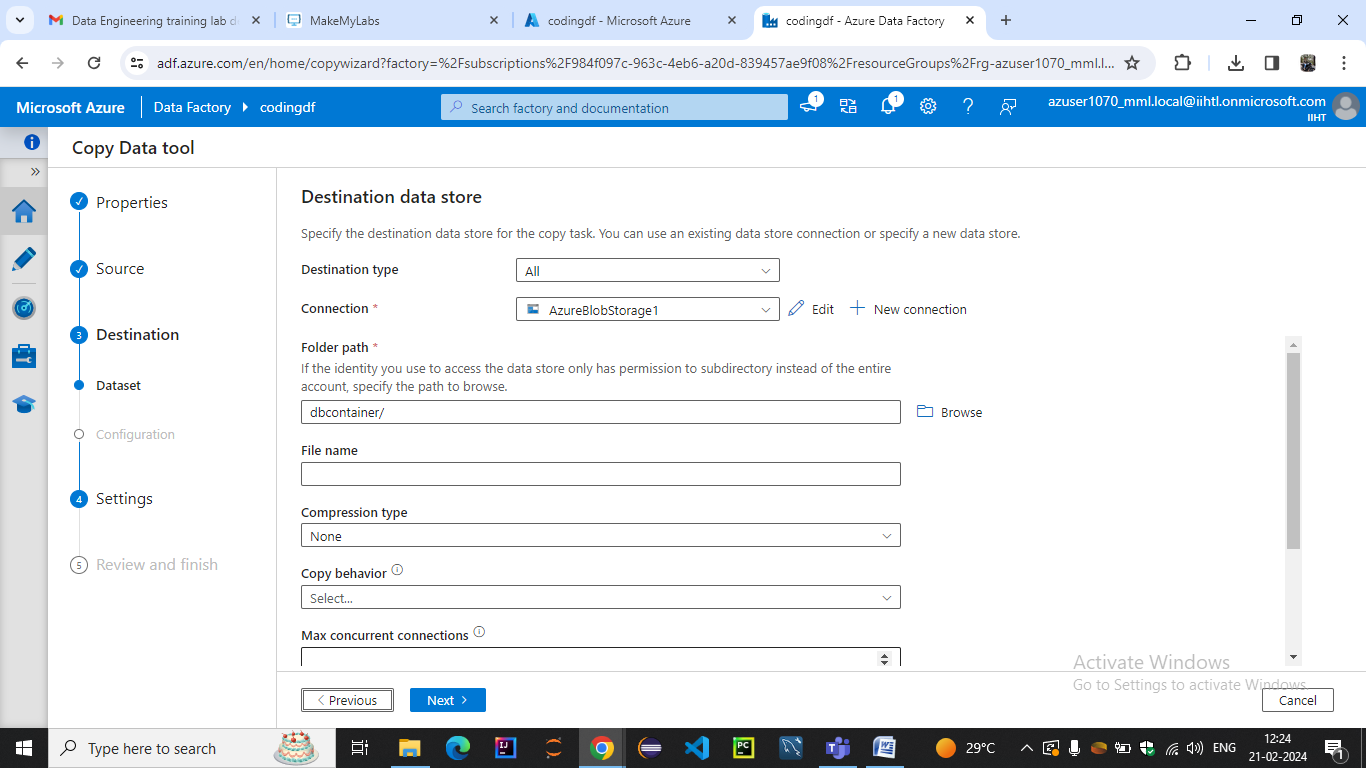
### Step 2: Complete source configuration

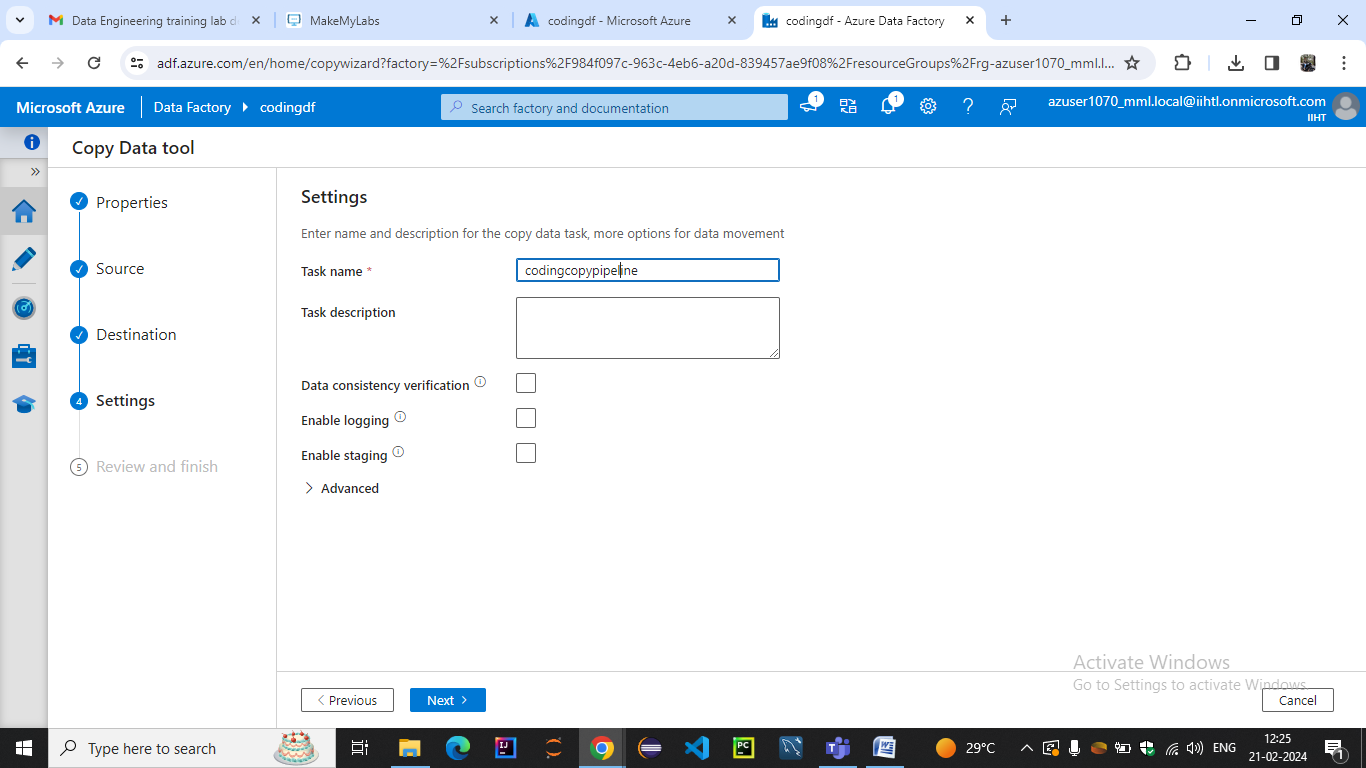
****

****

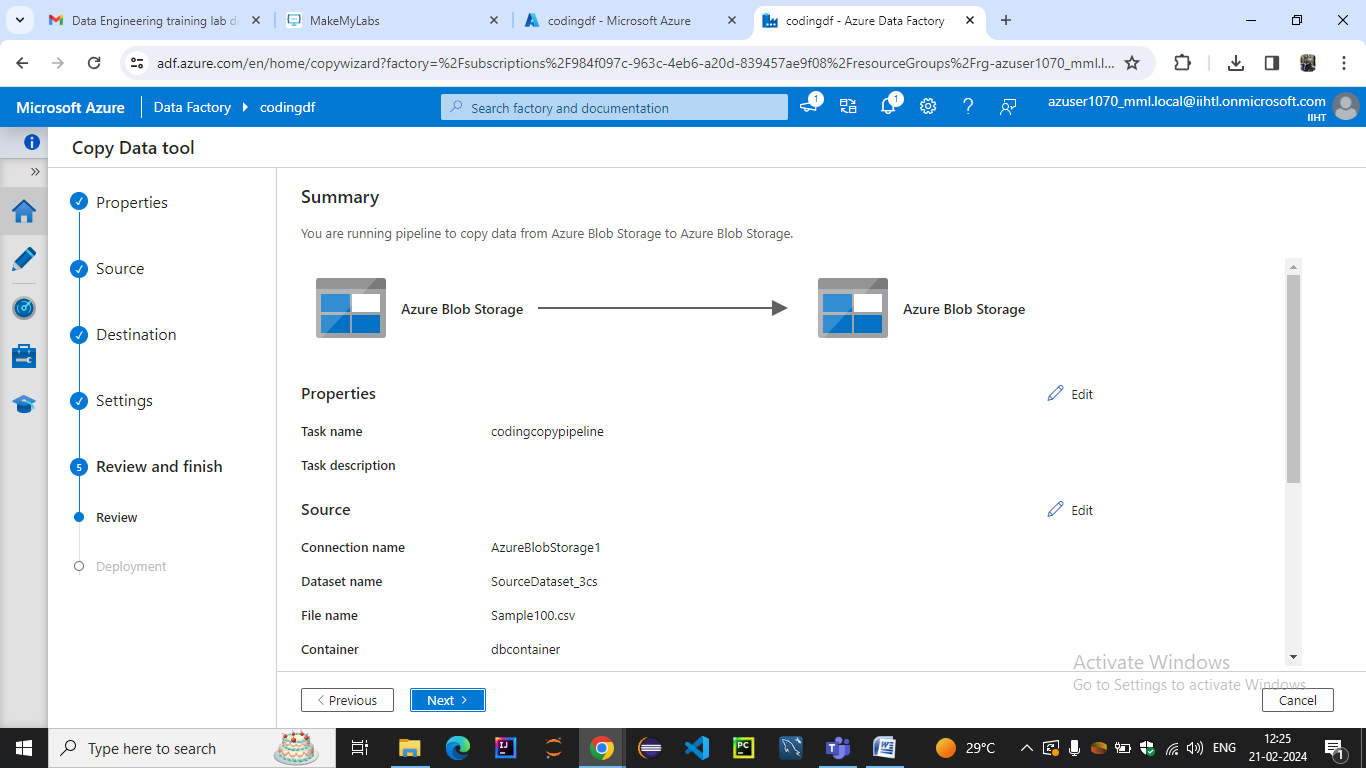
****

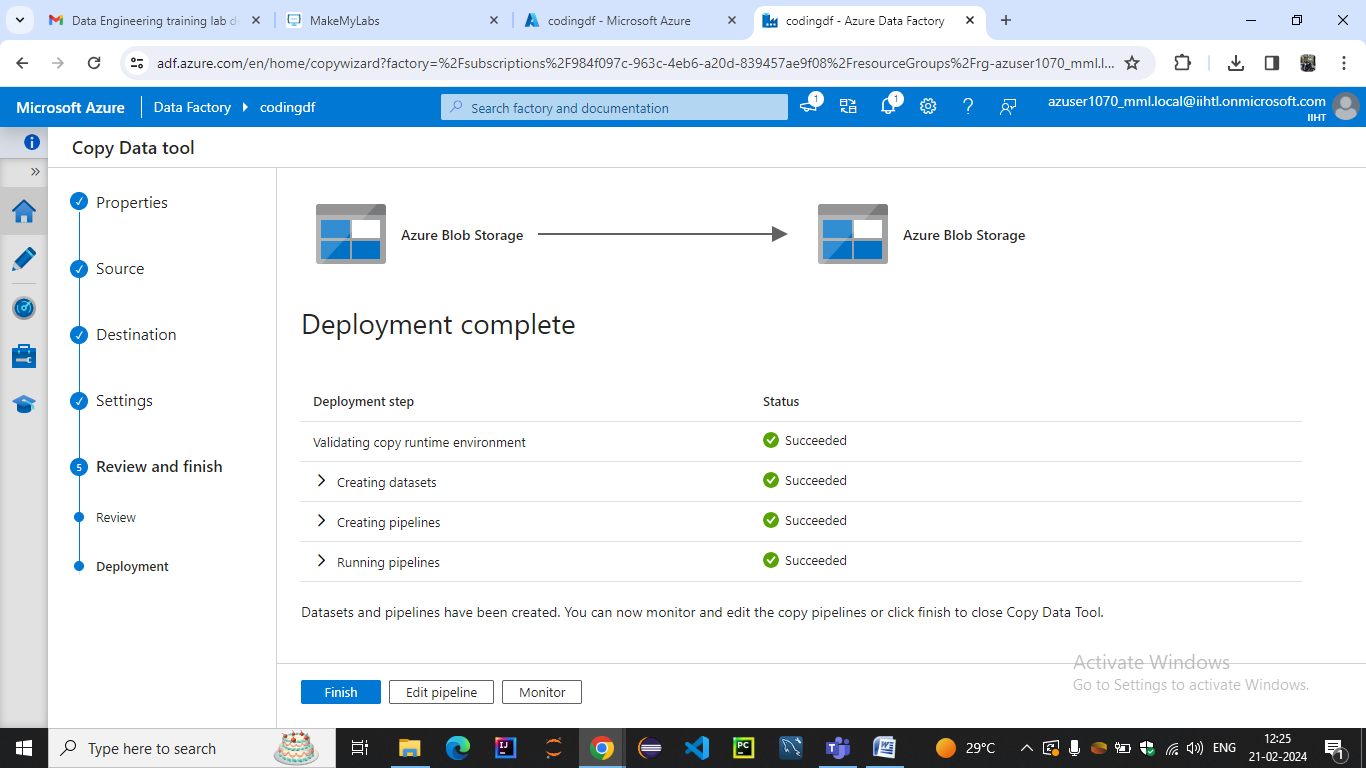
### Step 3: Complete destination configuration

****

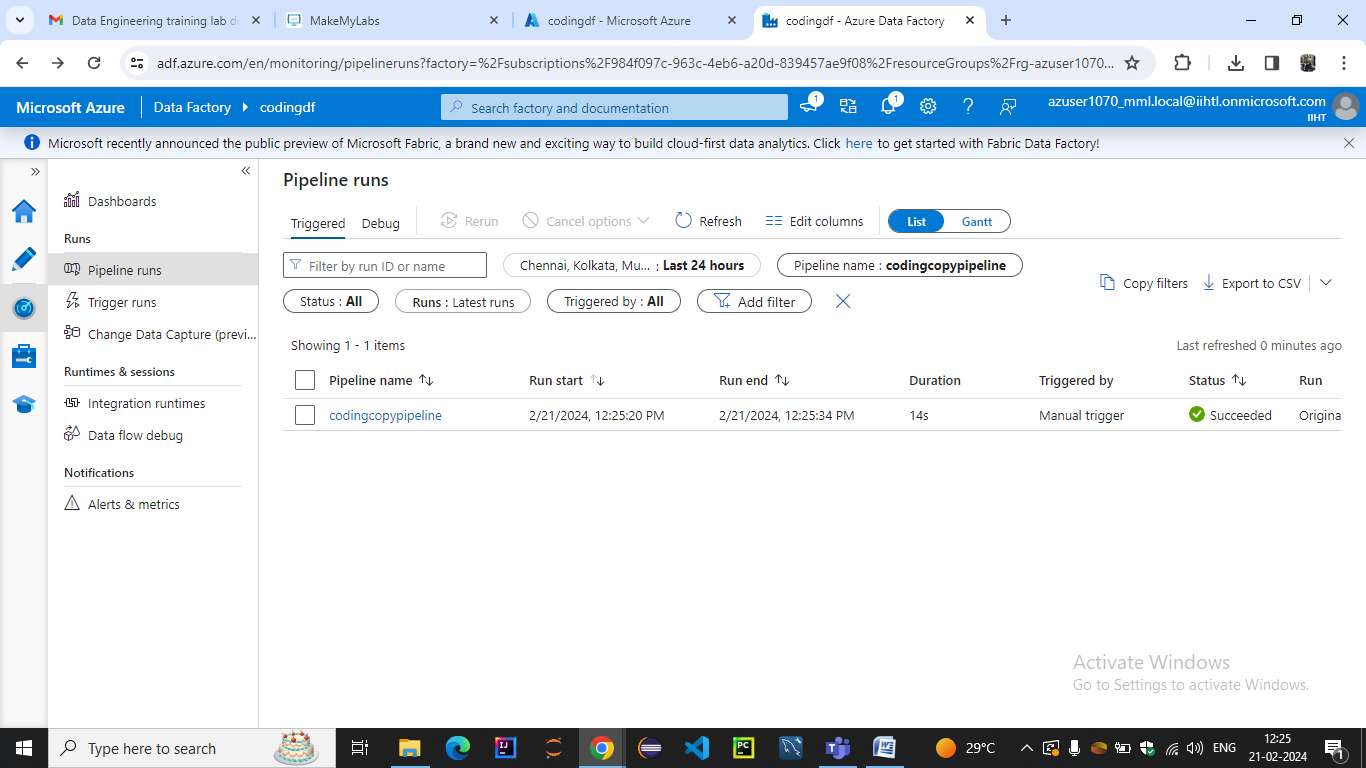
****

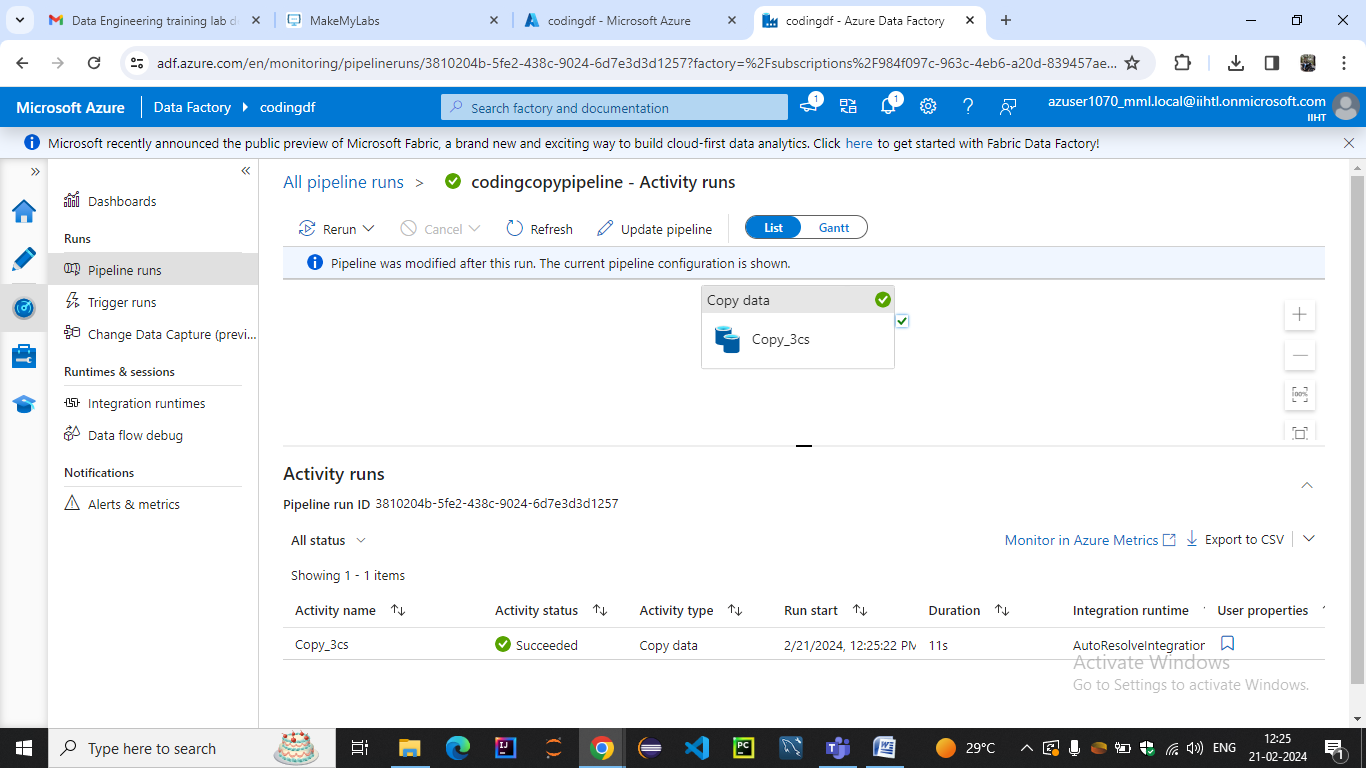
### Step 4: Review all settings and deployment

****

****

### Step 5: Monitor the running results

****

****