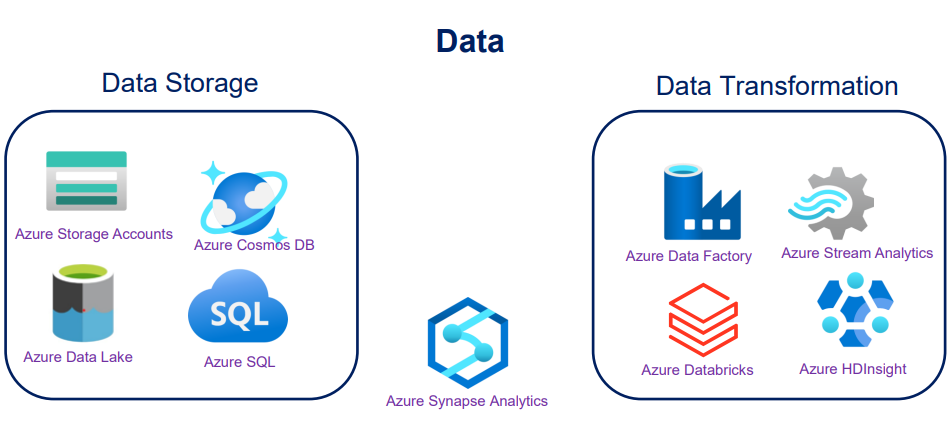
Azure Data Factory



**Azure Storage**

Azure Storage is a Microsoft-managed cloud service that provides storage that is highly available, secure, durable, scalable and redundant. Within Azure there are two types of storage accounts, four types of storage, four levels of data redundancy and three tiers for storing files

**Task-1**

**Copy Data from Azure Blob Storage to a SQL Database by using the Copy Data Tool (Copy Data Tool)**

**Task-2**

**Copy Data from Azure Blob Storage to a database in azure SQL Database by using Azure Data Factory (Copy Activity in Pipeline)**

**Task-3**

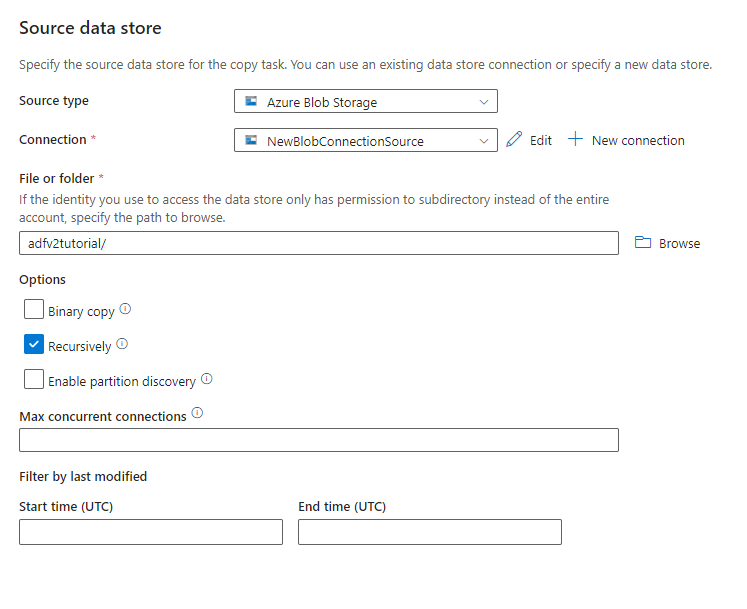
**Copy Data From A SQL Server Database to Azure Blob Storage by using the Copy Data Tool (Copy Data From On-Premise To Cloud)**

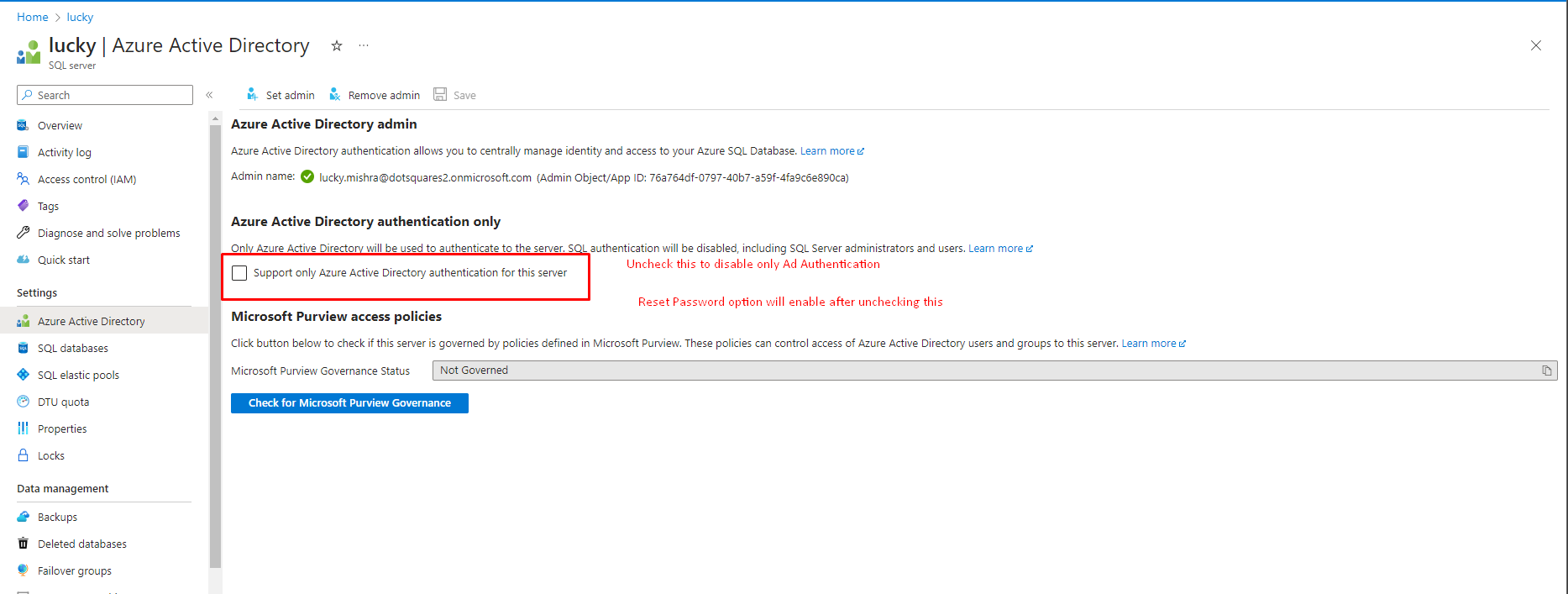
**Task-4**

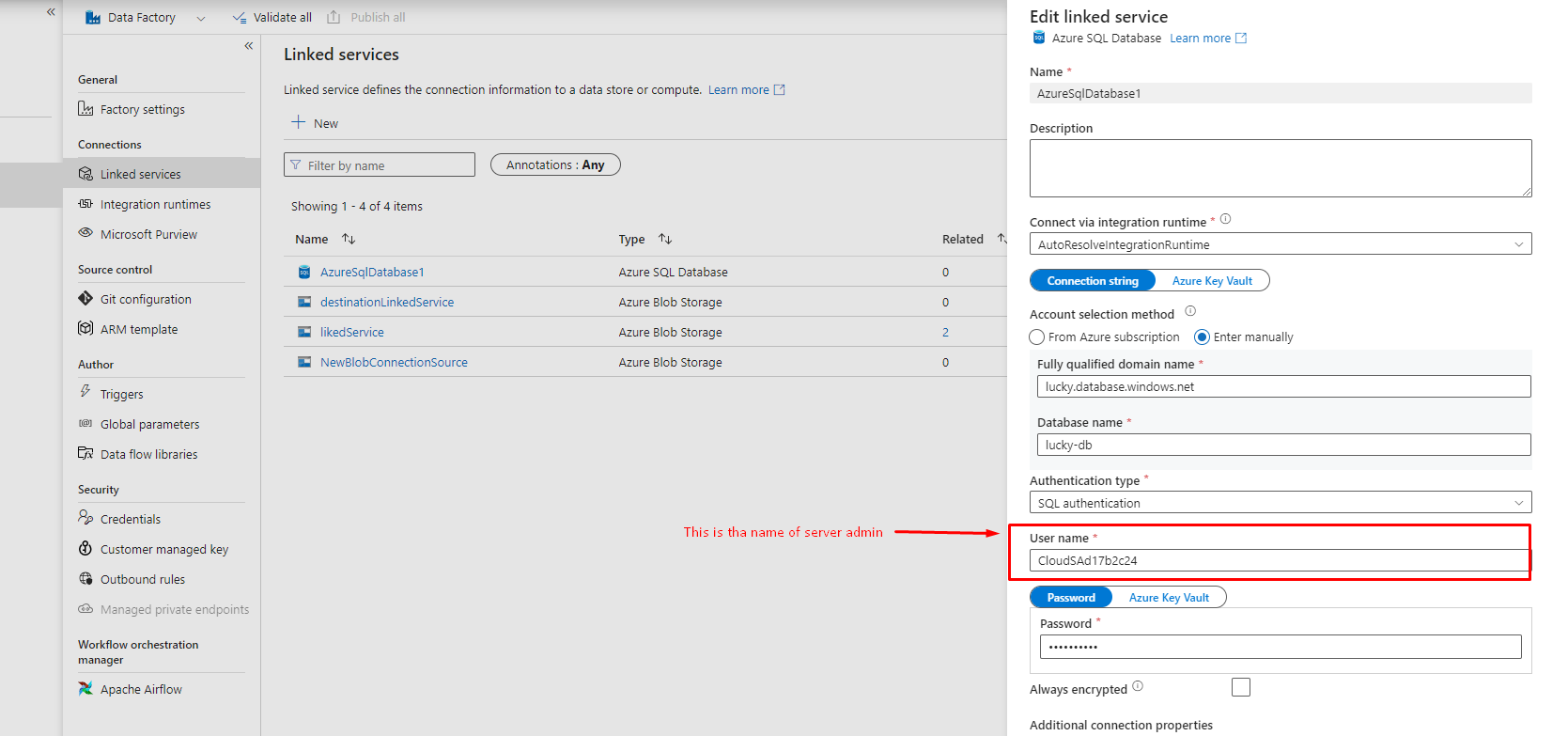
**Incrementally load data from Azure Sql database to Azure Blob Storage (Incremental copy Pattern)**

**Task-5**

**Create a Pipeline that copy data from a container in Azure Blob storage to another in the same storage account. If the activity succeeds or fails ,the pipeline sends details to the email**







**Task-5**

**Incrementally copy data from Azure SQL Database to Blob Storage by using change tracking in the Azure Portal (incremental pattern with change tracking)**

**Task-6**

**Incrementally load data from multiple tables in SQL Server to a database in Azure SQL Database using Azure Portal (Incremental SQL Db Multiple Tables)**

**------Data Flows-----**

**Task-1**

**Transform Data using Mapping Data flows done**

**Task-2**

**Transform Data in Delta Lakes using Mapping Data flows**

**Task-3**

**Prepare data with data wrangling done**

**Task-4**

**Best Practices for Azure Data Lake Gen2**

**Task-5**

**Understanding of Control Flow done**

**Task-6**

**Data Integration using Azure Data Factory and Azure Data Share done**

**Task-7**

**Push Data Factory lineage data to Microsoft Purview**

**Task-8**

**Understanding of SSIS Integration Runtime done**

**Task-1**

**Data Wrangling:**

**Data Wrangling in data factory allows you to build interactive power query mashups natively in ADF and then execute those at scale inside of an adf pipeline.**

**Data Wrangling is the process of cleaning , transforming and organizing raw data to make it suitable for analysis and other related tasks**

**Data Lake :**

**A Data Lake is a centralized repository that stores vast amount of raw and unprocessed data in its native format for flexible analysis and processing.**

**It can store structured , semi-structured and unstructured all types of data.**

**Ex. Social media feeds , database , audio , video files.**

**Data Warehouse :**

**A Data Warehouse is a structured and integrated repository that stores historical and consolidated data for reporting , analysis and business intelligence purposes.**

**A Data warehouse stores structured and formatted data to support analytical querying and reporting.**

**Ex : transactional data , financial data , sales data etc.**

**Azure Data Factory :**

**Azure Data Factory is a cloud-based data integration service that enables users to create and manage data workflows for ingesting, transforming, and moving data across various sources and destinations.**

**Delta Lakes :**

**Delta Lake is an open-source data storage layer that provides reliability, scalability, and performance enhancements on top of existing data lakes. It introduces ACID (Atomicity, Consistency, Isolation, Durability) transactions and schema enforcement capabilities to enable more robust data management and data quality in data lake environments. In simple terms, Delta Lake improves the reliability and usability of data stored in a data lake by adding transactional and schema enforcement features.**

**Control Flow in ADF :**

**In ADF , Control Flow refers to the sequence and order of activities in the pipeline.**

**Control Flow is represented by a series of activities that are executed in a specified order. Each activity performs a specific operation data , such as copying data from a source to a destination , transforming data , running scripts or invoking external services.**

**What is SSIS?**

**SQL Sever Integration services is a tool used to perform data integration tasks.It allows you to extract data from various data source , transform it and load it into different destinations.**

**Traditionally SSIS executes on - premise or within your local network.**

**What is Azure SSIS IR ?**

**Azure SSIS IR stands for Azure SQL Server Integration Services Integration Runtime**

**It is a service provided by Microsoft cloud that allows you to run your SSIS package in the Cloud.**

**It acts as an bridge between ssis and cloud.**

**Azure Factory Tutorials**

**<https://learn.microsoft.com/en-us/azure/data-factory/data-factory-tutorials?source=recommendations>**