## Sources

1. SQL Server

* JDBC – Java Database Connectivity. It is a Java-based API that allows Java based applications to interact with RDBMS (MySQL, SQL Server, Oracle).

1. Oracle
2. Big query

* To read data from Big Query you need to follow certain steps:

1. Create service account in GCP with necessary permissions.
2. Get the service account keys and download that in a JSON file format.

Azure postgresSQL

Important queries

**Azure Synapse Analytics VS Azure data factory VS Azure Databricks**

* Differences

1. Data transformation – Both platforms use no-code capabilities.
2. Machine learning
3. Security and Access Control
4. Pricing -

How to get incremental data from a table?

## Azure Data Factory

If Condition activity

* It acts like if else statement in normal programming.
* If the expression resolves to **True**, then It executes a set of activitiesand if the expression is **False** then another set of activities gets executed.

ADF – ingest data from Big Query to ADLS Gen 2:

Copy Activity:

* Google big query connection is built on top of Big Query APIs.
* Connector configuration –
  + There are 2 ways to create connector for Big Query linked service:

1. User authentication
2. Service authentication

User Properties in ADF

* User activity is used for monitoring purpose in ADF and Azure Synapse analytics.
* User properties are key: value pair defined at the activity level.
* By adding user properties, you can view additional information about activities under activity runs window to monitor your activity executions.

Annotations

* It is used to identify them easily.
* An annotation allows you to classify or group different entities to easily monitor or filter them after an execution.

### Notebook activity

Linked Service for notebook activity

There are 3 ways to authenticate authenticate

## Databricks

Databricks compute

* Azure Databricks compute are computation resources and configuration on which you run data engineering, data science and data analytics workloads, such as production ETL pipelines, streaming analytics, ad-hoc analytics and machine learning.