* **INTRODUCTION**

Azure data factory is composed of below key components.

1- Pipelines

2- Activities

3- Datasets

4- Linked Services

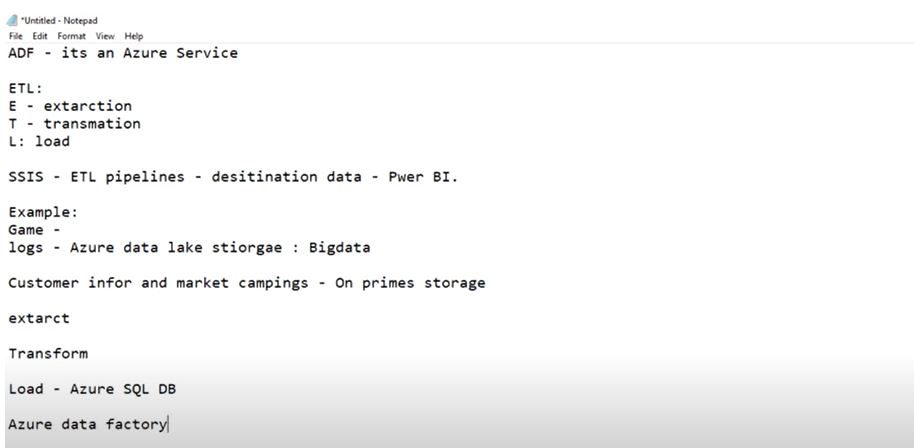
5- Data Flows

6- Integration Runtime

7- Triggers

These components work together to provide the platform on which you can compose data driven workflows with steps to move on and transform data.

**SSIS** -> SQL Server Integration Services



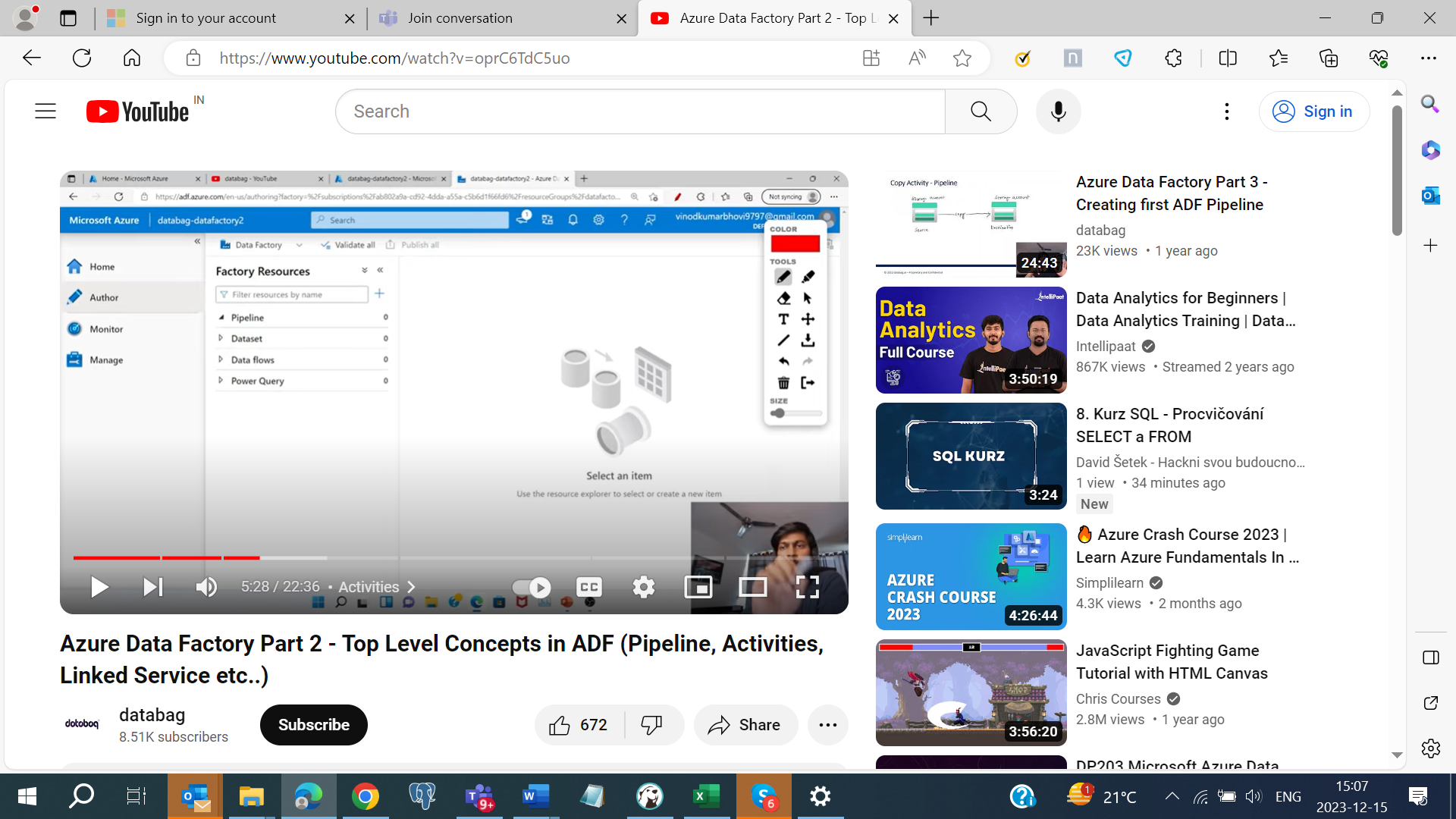
* **PIPELINE:**

A pipeline is a logical group of activities that perform a unit of work, together the activities in a pipeline perform a task.

**Example: -** A pipeline can contain a group of activities that ingests data from an Azure blob and then runs a Hive query on an HDInsight cluster to partition the data.

* **ACTIVITY:**

Activities represent a processing step in a pipeline. **For Example:-** you might use a copy activity to copy data from one data store to another data store.



* **DATASETS AND LINKED SERVICES: -**

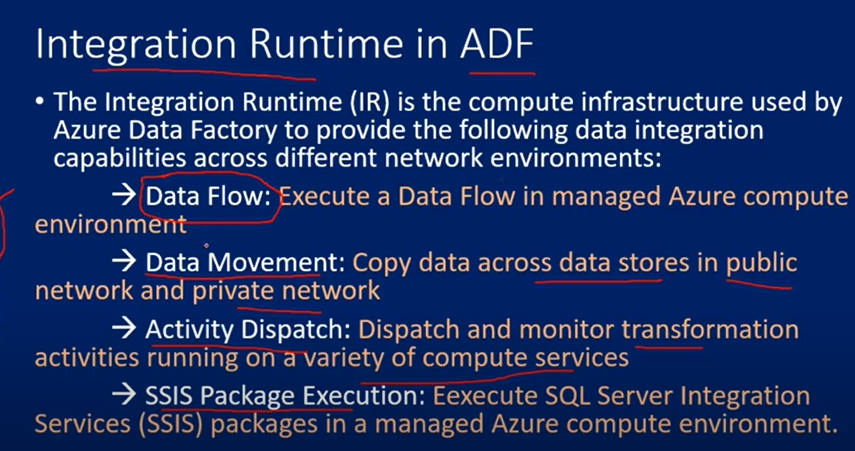
**Datasets** represents data structures within the data stores, which simply points to or reference the data store, which simply points to or reference the data you want to use in your activities as input or output.

**Linked services** are much like connection strings, which define the connection string that’s needed for data factory to connect to external resources. Think of it this way: a linked service defines the connection to data source, and a dataset represents the structure of the data. **For Example:-** An Azure storage linked service specifies a connection string to connect the Azure storage account. Additionally, an Azure blob dataset specifies the blob container and folder that specifies that data.

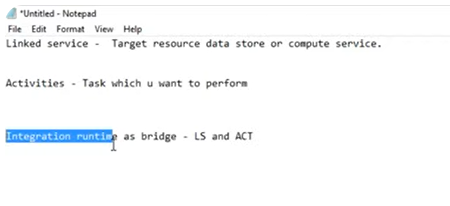


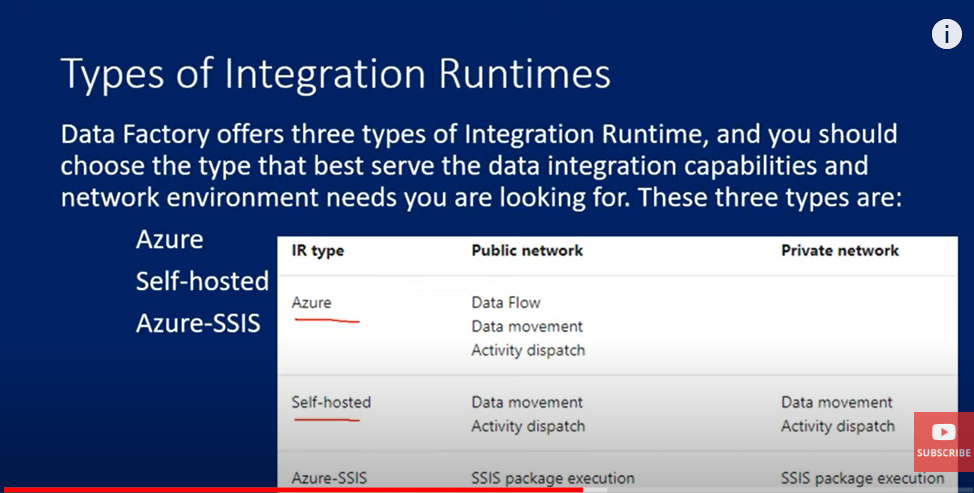
* **INTEGRATION RUNTIME: -**

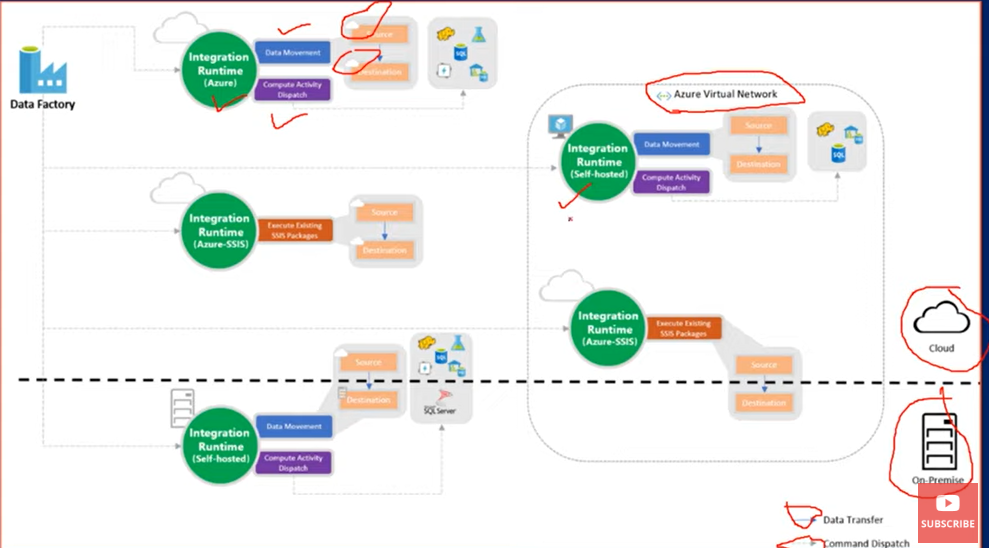
In Data factory, an **Activity** defines the action to be performed. A **Linked Service** defines a target data store or a compute service. An **Integration Runtime** provides a bridge between the **Activity and Linked Services**. It’s referenced by Linked Service or Activity and provides the compute environment where the Activity either runs on or get dispatched from. This way, the Activity can be performed in region closest possible to target Data Store or compute service in the most performant way while meeting security and compliance needs.



**AUTO RESOLVE INTEGRATION RUNTIME** is default integration runtime on Azure Data Factory.







**TYPES OF RUNTIME INTEGRATION**

* Azure Type
* Self Host
* Azure SSIS

