ADF Interviews Questions & Answers

1. What is Azure Data Factory (ADF)?

   Azure Data Factory is a cloud-based ETL (Extract, Transform, Load) service and data integration service that allows organizations to move and transform data from various sources to target destinations. It helps build scalable and automated data pipelines. Using ADF, we can connect to both cloud and on-premises data sources, perform transformations using Mapping Data Flows or external compute services, and schedule or trigger pipeline executions as needed. It's a key tool in building modern data platforms on Azure.

2. What are the key components of ADF?

The key components of Azure Data Factory include Pipelines, which group related activities; Activities, which perform specific tasks like copying or transforming data; Datasets, which represent the structure of input/output data; Linked Services, which define connections to data sources; Triggers for scheduling pipeline runs; Integration Runtime, which provides compute resources for executing data movement and transformations; and Control Flow, which manages the order and logic of activity execution.

3. What is Integration Runtime in ADF? Types?

Integration Runtime in ADF is the compute engine that executes data movement, transformation, and SSIS workloads. There are three types: Azure IR for cloud data operations, Self-hosted IR for accessing on-premises or private data sources, and Azure-SSIS IR for running SSIS packages in the cloud. Choosing the right IR depends on where your data resides and what kind of processing is required.

4. What are the different types of activities in ADF?

Azure Data Factory provides various activity types to support complex data workflows. These include Data Movement activities like Copy Activity, Data Transformation activities such as Mapping Data Flows and Stored Procedures, Control activities like If Condition and ForEach for logic handling, and External Execution activities like Web and Lookup Activity to interact with external systems. This flexibility allows us to automate and orchestrate complete data pipelines across hybrid environments.

5. What is a trigger in ADF? Types of triggers?

A Trigger in Azure Data Factory is used to automatically start a pipeline based on three types Schedule Triggers for recurring time-based runs, Tumbling Window Triggers for fixed interval executions with dependency tracking, and Event-Based Triggers that respond to file events like uploads in Blob Storage. This allows for flexible and automated data workflows.

6. What is the use of parameters in ADF?

Parameters in Azure Data Factory are used to make pipelines and datasets dynamic and reusable. By passing values at runtime—such as file names, dates, or database names—we can avoid hardcoding and create more flexible data workflows. Parameters can be used in pipelines, datasets, linked services, and can even be passed through triggers, making ADF highly configurable and efficient

7.