

ADF

Pipeline Name:

- The name of the Azure Data Factory pipeline.

Data Sources (Linked Services):

- **Linked Service Name:** The name of the linked service used in the pipeline to connect to external data sources.
- **Connection Strings:** Metadata about the connection, including server, database, authentication method, and credentials (e.g., using Azure Key Vault for secrets).
- **Shared vs. Embedded Linked Services:** Whether the linked service is shared across multiple pipelines or embedded in a single pipeline.

Datasets:

- **Dataset Name:** The name of the dataset that defines the data structure for a data source.
- **Source Type:** The type of data source (e.g., SQL Server, Azure Blob Storage, REST API).
- **Schema:** Metadata about the structure of the dataset, including the columns and their data types.
- **Parameters:** Input parameters required for configuring the dataset, such as file paths, query filters, or partition keys.
- **Data Format:** The data format of the dataset (e.g., CSV, JSON, Parquet).

Pipeline Activities:

- **Activity Name:** The name of the activity (e.g., Copy Data, Execute Pipeline, Data Flow, Lookup).
- **Source and Sink:** The source dataset and destination (sink) dataset used in the activity.
- **Transformation Logic:** Details about any transformations or mapping applied to the data during the activity.
- **Parameters:** Activity-level parameters required to process the data (e.g., filters, partition keys).
- **Error Handling:** Logic and actions taken if the activity fails (e.g., retry policies or failure handling paths).

Data Flows (Mapping Data Flows):

- **Data Flow Name:** The name of the data flow that defines the transformation logic.
- **Source Dataset:** The dataset providing the input data for the data flow.
- **Transformation Logic:** The transformation steps applied to the data (e.g., aggregations, joins, filters, derived columns).
- **Sink Dataset:** The destination dataset where the transformed data is written.

- **Schema Mapping:** The mapping of source columns to destination columns, including calculated columns and data transformations.

Parameters:

- **Parameter Name:** The name of the parameter used in the pipeline or activity.
- **Parameter Type:** The data type of the parameter (e.g., string, integer, array).
- **Default Values:** Default values for parameters, either static or derived from other pipeline components.
- **Optional vs Required Parameters:** Metadata about whether the parameter is optional or required.

Linked Services:

- **Linked Service Name:** The name of the shared or embedded linked service.
- **Connection String:** Information about the connection to external data sources (e.g., database server, credentials).
- **Security Settings:** Permissions, authentication methods, and access control (e.g., Azure Active Directory, Managed Identity).

Pipelines and Data Flows:

- **Pipeline Parameters:** Parameters used to configure pipeline behavior and logic.
- **Data Flow Transformations:** Transformations applied to the data, such as join, filter, aggregate, and derived column expressions.
- **Sink Datasets:** The destination datasets where transformed data is written.