Metadata-Driven Azure Data **Factory**

Course Duration

1 Day

This course will provide delegates with the skills to define and setup a metadata-driven approach to using Azure Data Factory to create the necessary ETL (Extract, Transform and Load) processes to extract data from various sources and use it to populate a data warehouse.

The processes created within Azure Data Factory should be highly parameterised in order to make them as flexible, efficient and as reusable as possible. In a metadata-driven approach, the values of the parameters are held in a Metadata Control table which alleviates the need to supply detailed and complicated parameters to the pipelines.

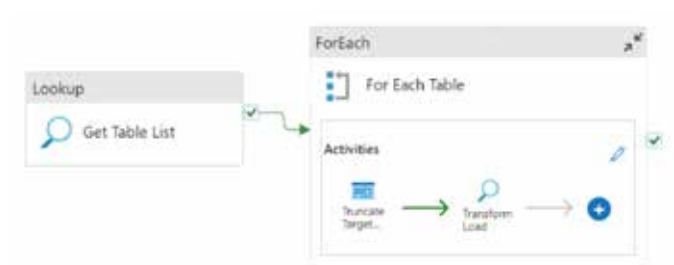## On completion, delegates will be able to:

- Setup a Metadata Control table and populate it with the values required by the Data Factory pipelines to run the ETL processes.

- Create Azure Data Factory pipelines that will read the Metadata Control table and use the values held within it to perform the required ETL processes.

- Incorporate delta extract and slowly-changing dimension processing.

- Setup orchestration pipelines to run the individual pipelines in the correct order and with the correct dependencies.

Who should attend?

IT professionals who will be using Azure Data Factory in order to populate a data warehouse.

Prerequisites

This course is an add-on to the standard 2-day Azure Data Factory course which the delegates must have attended beforehand.

Azure Data Factory Studio

Course Duration

1 day

Practical work

Practical sessions are used within the course to reinforce the topics covered. These allow delegates to have direct hands-on practice when using Azure Data Factory.

Course Content

See below

Related and follow-on courses

- Data Warehouse Standards
- Azure Data Warehouse Resources

Course Content

- **Defining the Metadata.**

  - Setup the Metadata Control table and populate it with the required data.
  - Review the stored procedures required to transform and load the necessary data.

- **Pipeline to Extract Data from a Flat File.**

  - Create a pipeline to extract all required data from flat files held in an Azure Storage account using the data contained within the Metadata Control table.

- **Pipeline to Extract Data from a DB Table.**

  - Create a pipeline to extract all required data from tables held in an Azure SQL Database using the data contained within the Metadata Control table.

- **Pipeline to Transform and Load Data.**

  - Create a pipeline to transform all required data held in the landing tables and load it into the data warehouse using the data contained within the Metadata Control table.

- **Amend the Metadata for Delta Loads and SCDs.**

  - Amend the Metadata Control table to source the data from alternate tables for which a delta extract can be performed and to run alternate stored procedures which cater for slowly-changing dimensions.

- **Orchestration Pipelines.**

  - Create Orchestration pipelines to sequence the running of several pipelines in the correct order with correct dependencies.

Acuma Education

All Acuma courses are delivered by trainers who are highly experienced consultant practitioners. This rich project experience coupled with professional training skills helps delegates to focus on how the course contents can be put into practice in their workplace. Acuma delivers courses from dedicated training facilities in Manchester, Birmingham and London, but also deliver training offsite using suites of laptops, with the option to customise courses to use your own business data and tailored to your exact training needs. The most appropriate training format depends upon your training objectives, number of learners, timescales, budgets and delivery logistics. Whatever your needs, Acuma will find the approach most suited to them.

Acuma Education's seal of quality is demonstrated through consistent achievement of the prestigious "Gold Standard" from the Learning and Performance Institute.