Azure Cloud Platform

Extract/source Transform/Processing Load/Target

BLOB

Binary large Object



AZURE SQL

AZURE DATA FACTORY (DF)

ADLS

Azure Data Lake storage



ETL thorough linked services

* Create Services – BLOB and ADLS, ADF and ASQL using create resource
* Create a container in BLOB and ADLS to upload the necessary source file.
* BLOB, ADLS and ASQL are different types of storage in Azure.
* Create a dataset to get the data from source (BLOB) through linked service and store them temporarily for the dataflow to process.
* Create a dataset to get the data from source (ADLS) through linked service and store them temporarily for the dataflow to process.
* Dataflow or Power Query needs to be created to do processing/transformation of data - >datatype change, addition or removal of columns with transformations/ transformations on existing columns can be applied by creating a new column and then delete existing one.
* Create a dataflow to get the data from datasets(both) and do necessary transformations. Once the transformations are done, then choose sink tab in DF to sink them (load them to target) to target DB through linked service by providing target DB details and appropriate table names.
* Use options like join to join the different datasets, then select the column, remove the unwanted columns as required.
* Once the sink is done, one new dataset will be created automatically for the processed data to be sent to ASQL
* Once the dataflow is created, new pipeline should be created to execute the actions of dataflow (Storage + processing) – based on the activities we choose. Here we have chosen the dataflow option under activities.
* Points to remember:

1. Publish the activities whenever each task is completed
2. Import projections and check the datatypes/ modify them uniformly across while creating the dataset to avoid any datatype errors while building expressions in dataflow creation.
3. Switch on dataflow debug mode on at least for 2 hrs to get the data preview in dataflow.
4. Add trigger to be clicked to execute the pipeline.

* Refer the following screenshots which gives all the details of above-mentioned steps.











































































