Extract/source Transform/Processing Load/Target

ADLS STORAGE

Azure data lake storage



AZURE SQL

AZURE DATA FACTORY (PQ)

ETL thorough linked services

**Steps to create Pipeline:**

* Create Services – ADLS, ADF and ASQL using create resource

Note: While creating a storage account, enable hierarchical namespace which create the ADLS storage instead of BLOB.

Advantages of creating ADLS is it will have a option of add directory inside containers which will helps you to organize your files (like subfolders)

* Create a container in 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 (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 PQ under power query to get the data from dataset and do necessary transformations as similar like excel.
* Once the PQ 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 power query option under activities.
* Note: Sink option is available in pipeline activities for power query, Once the sink is done, one new dataset will be created automatically for the processed data to be sent to ASQL
* Points to remember:

1. Publish the activities whenever each task is completed
2. Add trigger to be clicked to execute the pipeline.
3. Ensure the datatypes are modifies while creating the datasets.

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















































