Day 4

Azure Data Factory Masterclass

About Me

Alpa Buddhabhatti

/alpaBuddhabhatti/

/alpabuddhabhatti/

/AlpaB7

/@alpabuddhabhatti

/@meetalpa

Microsoft Certified:
Azure Data Engineer
Associate
Microsoft

Microsoft Certified:
Azure Data Scientist
Associate
Microsoft

Microsoft Certified:
Azure Developer
Associate
Microsoft

Microsoft Certified
Trainer 2021-2022
Microsoft

Agenda

1. Day 3 Recap

2. ADF Mapping dataflow

3. Azure Data Factory more Activities

   Data flow Activity

   Stored Procedure activity

2. Labs

   Creating Complex data flow using other transformation
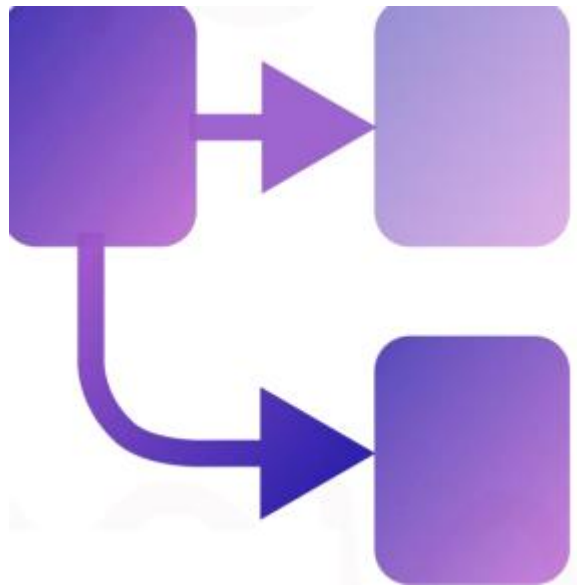
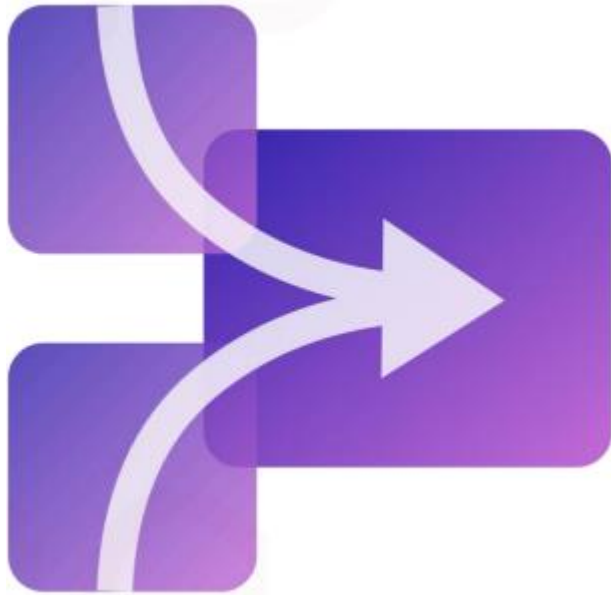3. Conclusions

Day 3 Recap

Branch Transformations

Self-joins

Aggregating on different levels

Join Transformations

Combine two data streams

- Full outer
- Inner
- Left outer
- Right outer
- Cross

Conditional Split Transformations

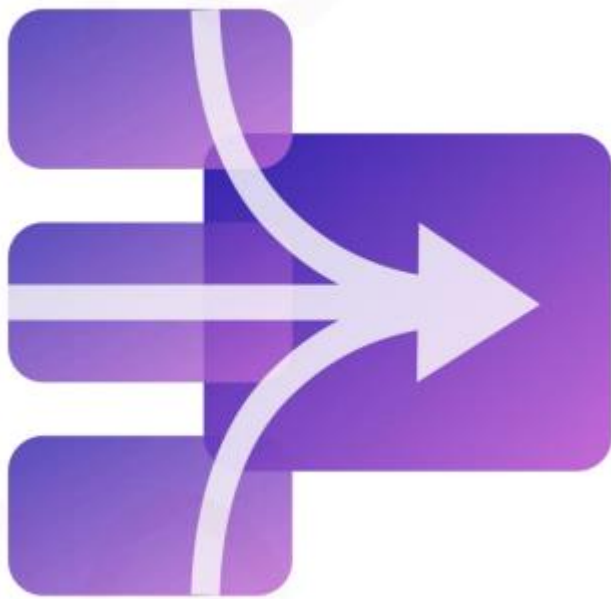Route rows to different streams based on conditions

Exists Transformations

Filter rows based on exists or not exist in different data stream

Union *Transformations*

Combine multiple data streams into one data stream

Lookup *Transformations (similar to left or right join)*

Add reference data from another source by matching on key fields

Aggregate Transformation

Define aggregations like:

- SUM

- MIN

- MAX

- COUNT

Surrogate Key

Add incrementing non-business key value to rows

Filter

Use expressions to filter which rows will be passed through the data stream

Flatten

When you have complex JSON/XML file

Labs