**Slowly Changing Dimension (SCD) Type 1 - Azure Data Factory Mapping Data Flow**

**Overview**

This document provides details on the **Dataflow3** Mapping Data Flow in Azure Data Factory, which is responsible for processing employee data by performing lookups, updates, and inserts into the **DIM\_EMPLOYEE\_TYPE2** table in an Azure SQL Database.

**Data Sources & Destinations**

**Sources:**

1. **source1** (File Source - CSV)
   * Linked Service: synapseshubha-WorkspaceDefaultStorage
   * Reads employee data from PaperTowel2.csv.
   * Schema:
     + ID (integer)
     + NAME (string)
     + CITY (string)
     + PHNO (long)

### **A screenshot of a computer Description automatically generated**A screenshot of a computer Description automatically generatedA screenshot of a computer Description automatically generatedA screenshot of a computer Description automatically generatedA screenshot of a computer Description automatically generatedA screenshot of a computer Description automatically generatedA screenshot of a computer Description automatically generated **2. Select Transformation (select1)**

* Renames columns with a SRC\_ prefix for clarity.

### A screenshot of a computer Description automatically generatedA screenshot of a computer Description automatically generatedA screenshot of a computer Description automatically generated **3. Derived Column (derivedColumn1)**

* Creates a SRC\_HASHKEY using a CRC32 hash function on ID, NAME, CITY, and PHNO:  
  SRC\_HASHKEY = crc32(concat(toString(SRC\_ID), SRC\_NAME, SRC\_CITY, toString(SRC\_PHNO)))

A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated  **source2** (Azure SQL Database)

* Linked Service: AzureSqlDatabase1
* Reads active employee records from **DIM\_EMPLOYEE\_TYPE2**.
* SQL Query:

SELECT EMP\_ID, EMP\_HASHKEY FROM DIM\_EMPLOYEE\_TYPE2 WHERE IsActive = 1;

* Schema:
  + EMP\_ID (integer)
  + EMP\_HASHKEY (long)

### A screenshot of a computer Description automatically generatedA screenshot of a computer Description automatically generatedA screenshot of a computer Description automatically generatedA screenshot of a computer Description automatically generated **Lookup (lookup1)**

* Matches SRC\_ID from source1 with EMP\_ID from source2.

### A screenshot of a computer Description automatically generatedA screenshot of a computer Description automatically generatedA screenshot of a computer Description automatically generated **Split (split1)**

* **New Records (Insert):** If EMP\_ID is NULL, it is a new employee.
* **Updated Records (Update):** If SRC\_ID exists but SRC\_HASHKEY differs, the record is updated.

### A screenshot of a computer Description automatically generatedA screenshot of a computer Description automatically generatedA screenshot of a computer Description automatically generated **Derived Column (derivedColumn2) - Updating Records**

* If a record is modified, the following values are updated:
* SRC\_UPDATED\_BY = "Dataflow-Updated"
* SRC\_UPDATED\_DATE = currentTimestamp()

SRC\_IsActive = 0

### A screenshot of a computer Description automatically generatedA screenshot of a computer Description automatically generatedA screenshot of a computer Description automatically generated **8. Alter Row (alterRow1)**

* Configures update rules for sink1.

A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated  **sink1** (Update Existing Records in Azure SQL Database)

* Linked Service: AzureSqlDatabase1
* Table: DIM\_EMPLOYEE\_TYPE2
* Updates records based on keys: EMP\_ID, EMP\_HASHKEY
* Operations: Update existing employee data.

### A screenshot of a computer Description automatically generatedA screenshot of a computer Description automatically generatedA screenshot of a computer Description automatically generatedA screenshot of a computer Description automatically generatedA screenshot of a computer Description automatically generated **Union (union1)**

* Merges new and updated records.

### A screenshot of a computer Description automatically generatedA screenshot of a computer Description automatically generatedA screenshot of a computer Description automatically generated **Derived Column (derivedColumn3) - Preparing Data for Insert**

* Assigns UPDATED\_BY = "Dataflow-Updated" for new records:
* SRC\_CREATED\_BY = "Dataflow"
* SRC\_CREATED\_DATE = currentTimestamp()
* SRC\_UPDATED\_BY = "Dataflow-Updated"
* SRC\_UPDATED\_DATE = currentTimestamp()

SRC\_IsActive = 1

A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated  **sink2** (Insert New Records in Azure SQL Database)

* Linked Service: AzureSqlDatabase1
* Table: DIM\_EMPLOYEE\_TYPE2
* Inserts new employee records.

A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated