**Database Table Creation Script**

* Creating a table in SQL DB

CREATE TABLE CUST\_SCDTYPE1

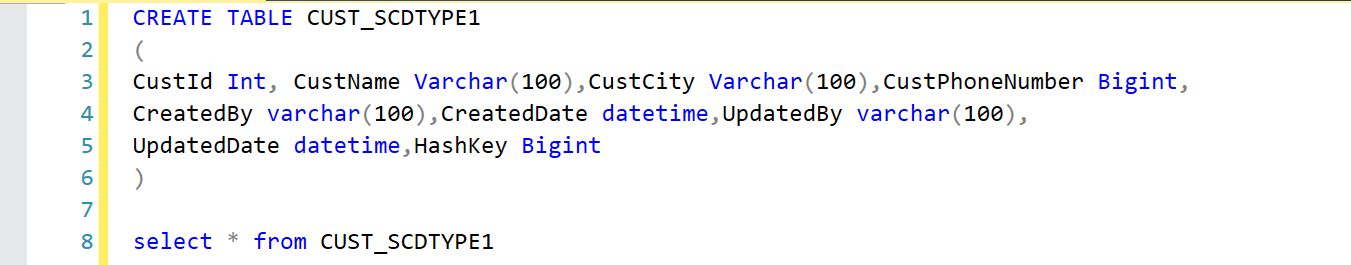
(

CustId Int, CustName Varchar(100),CustCity Varchar(100),CustPhoneNumber Bigint,

CreatedBy varchar(100),CreatedDate datetime,UpdatedBy varchar(100),

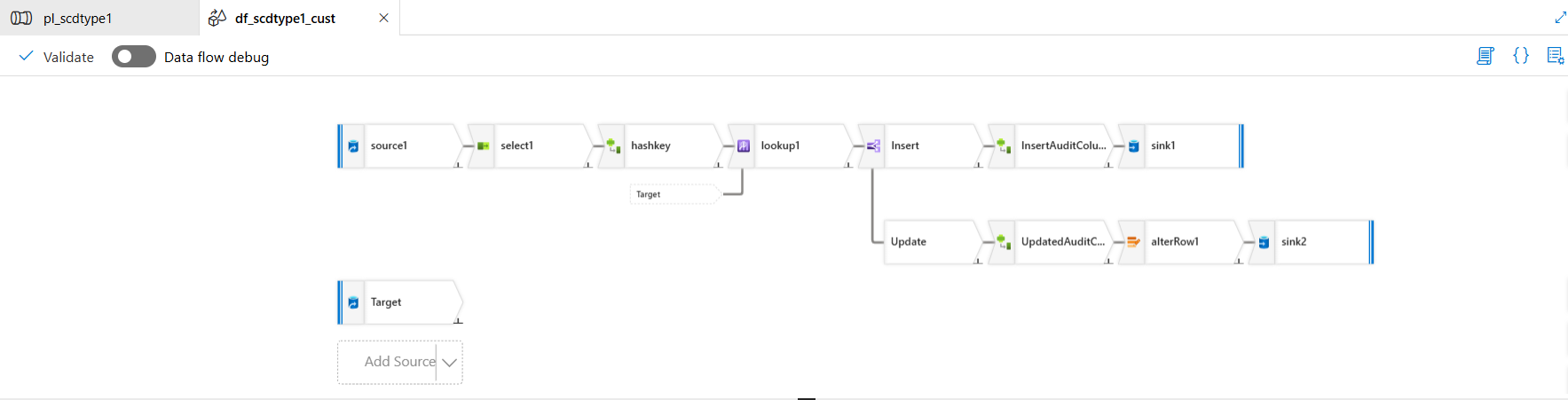
UpdatedDate datetime,HashKey Bigint

)



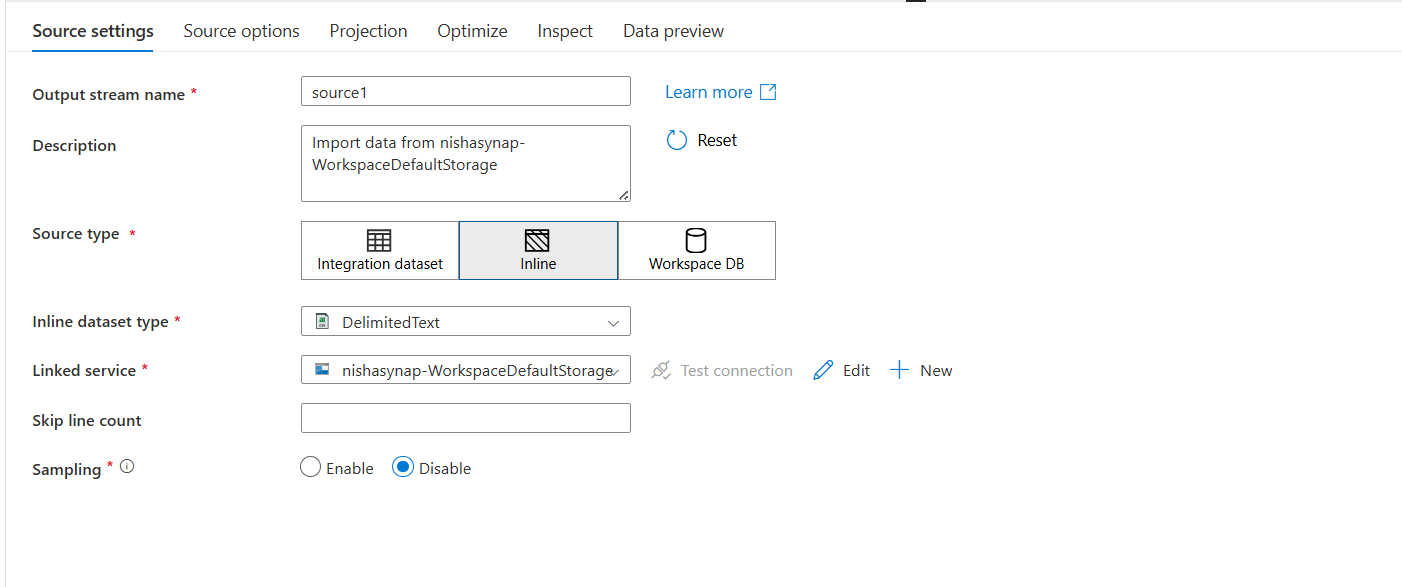
**Pipeline Creation and Optimization**

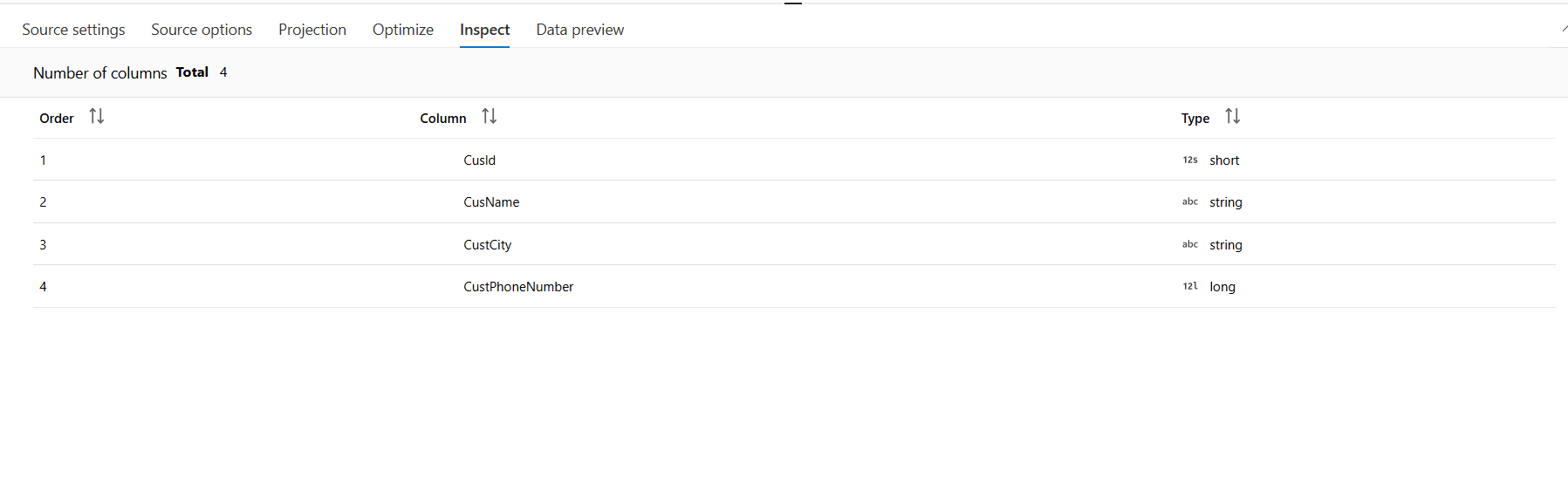
\* Practiced creating pipelines in Azure Synapse Analytics, focusing on:

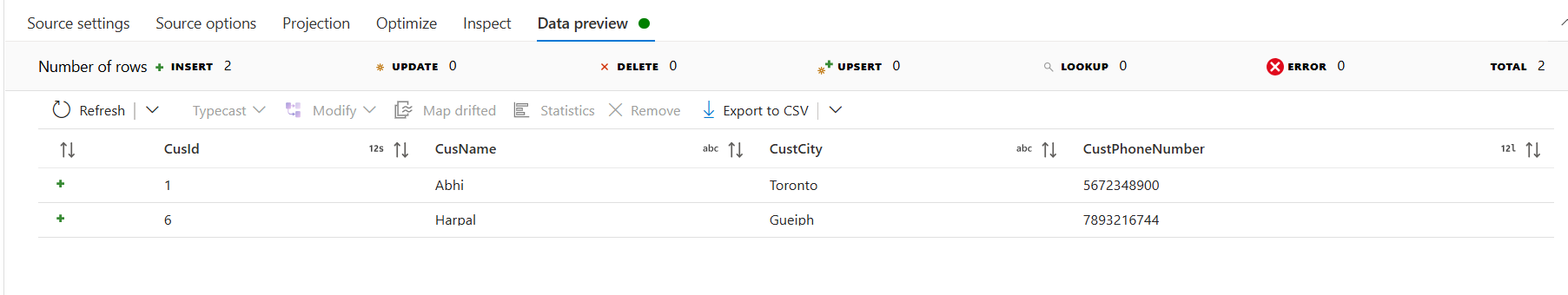


\* Source configuration, column renaming, and hash key generation.

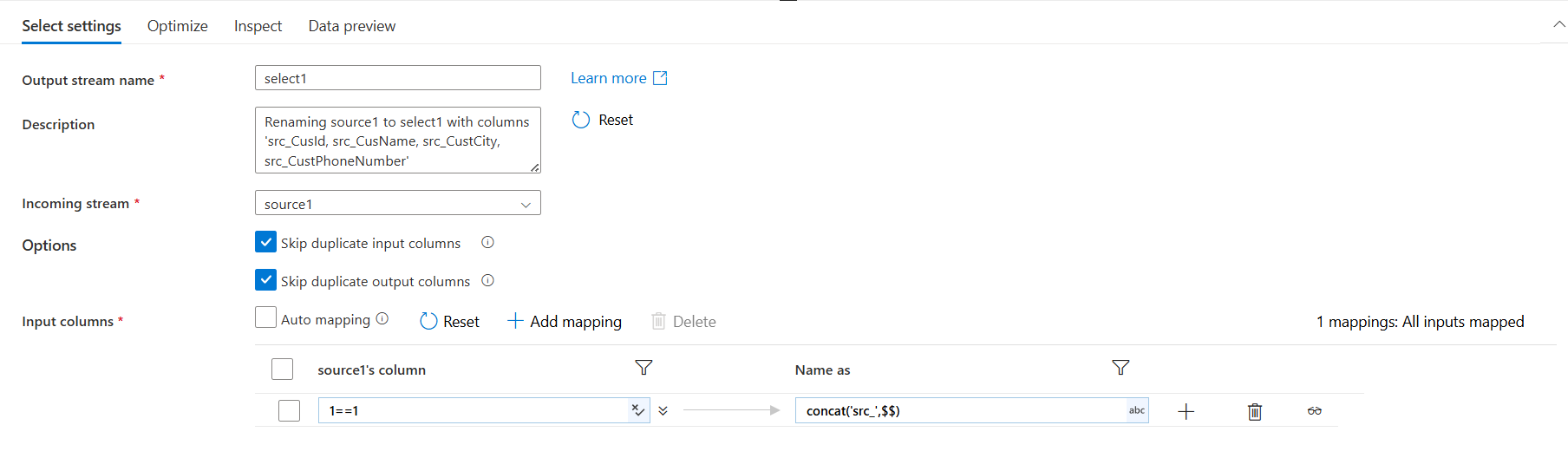
* In Source we are choosing ADLS Gen 2 as inline data type, and in file path we are selecting the CSV file which has data consisting details such as ID, Name, City, Phone Number



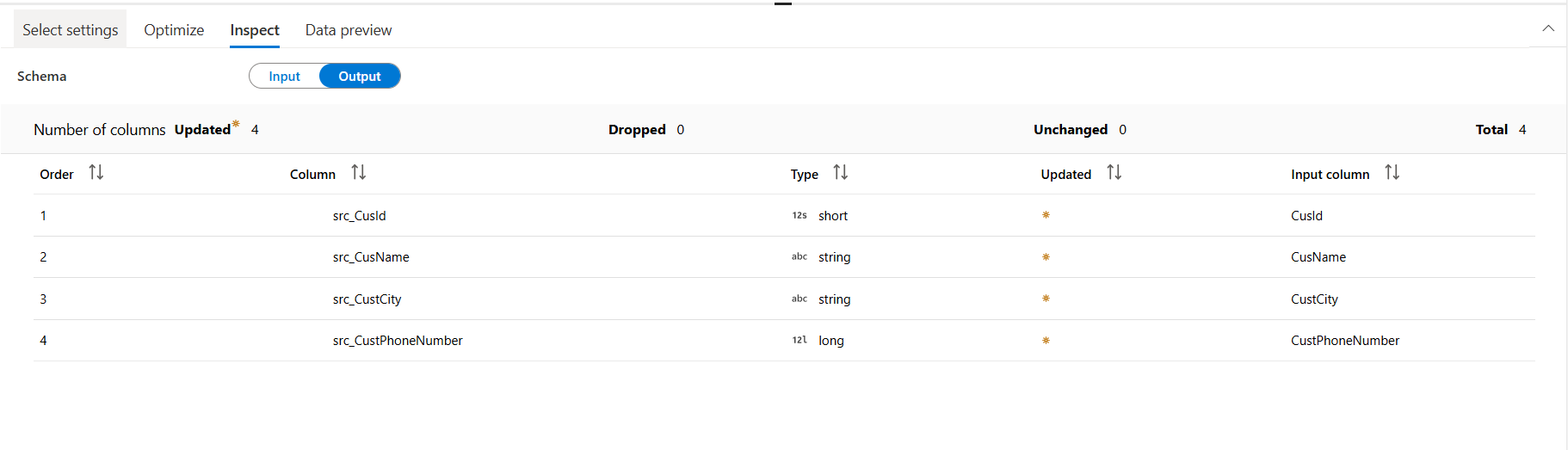
****

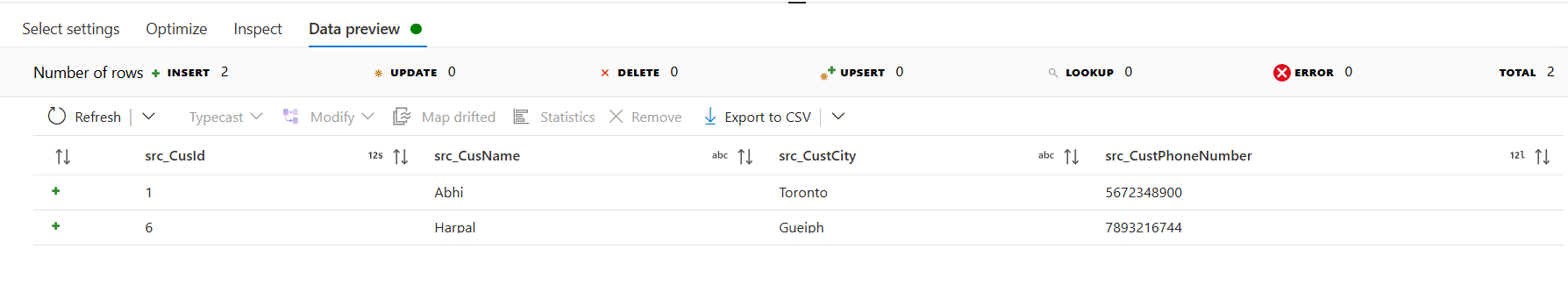
****

* Secondly, we are using a select settings to rename the columns



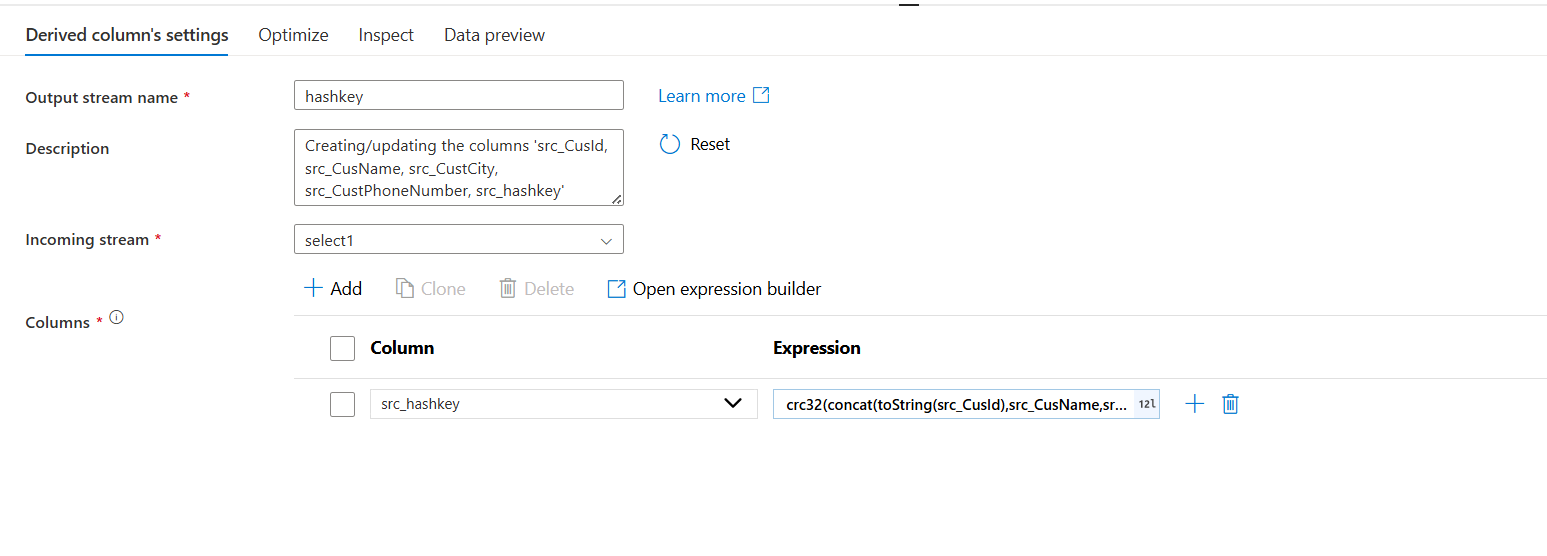
* The renamed column will look like below

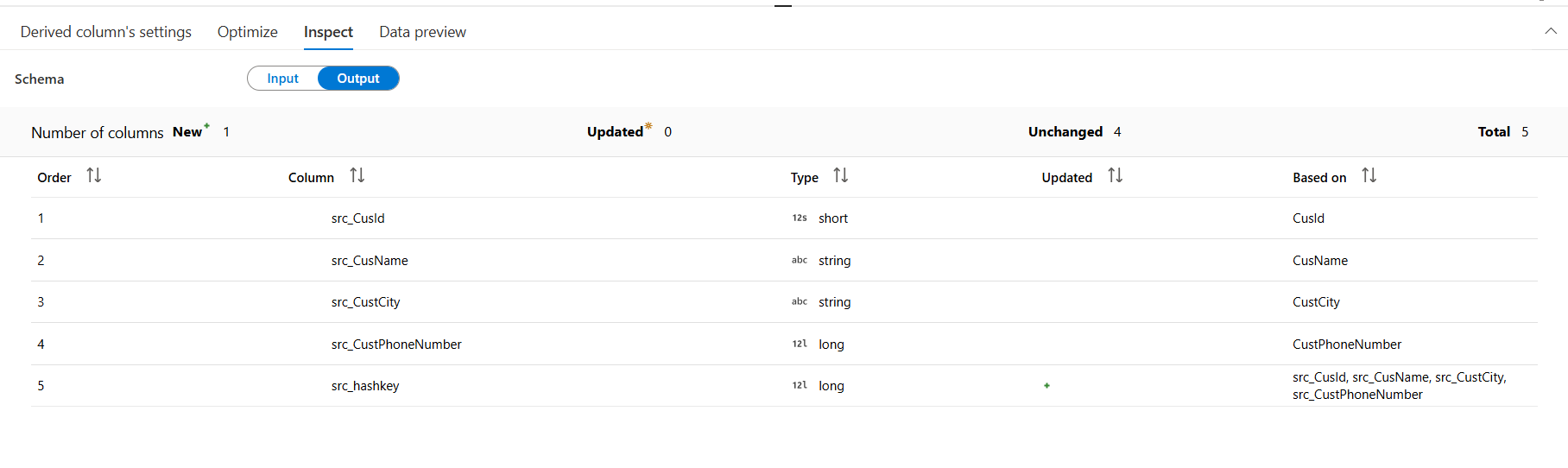


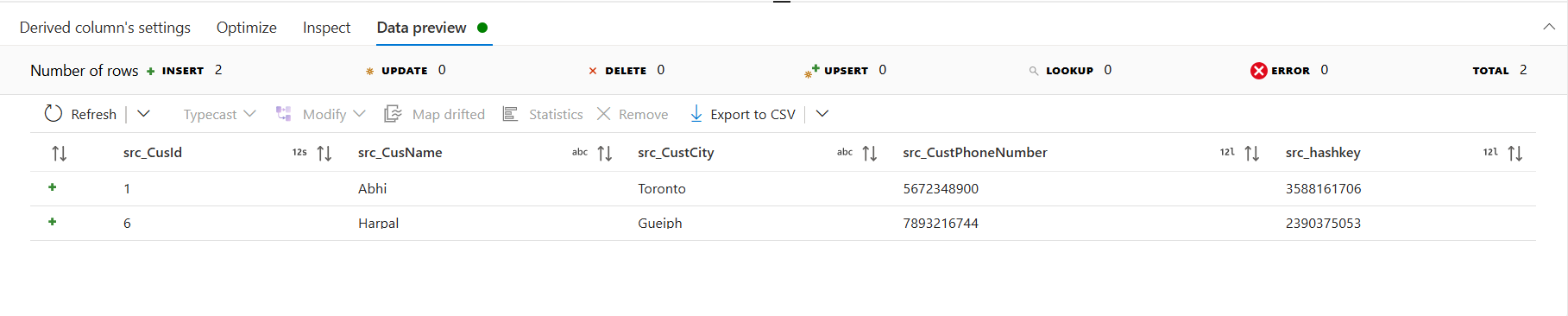


* Next, a derived column setting is used to generate unique hashkey for all the data which are in the CSV file. The hashkey will be updated if any changes is made to the existing data
* The below expression is used to generate the hashkey

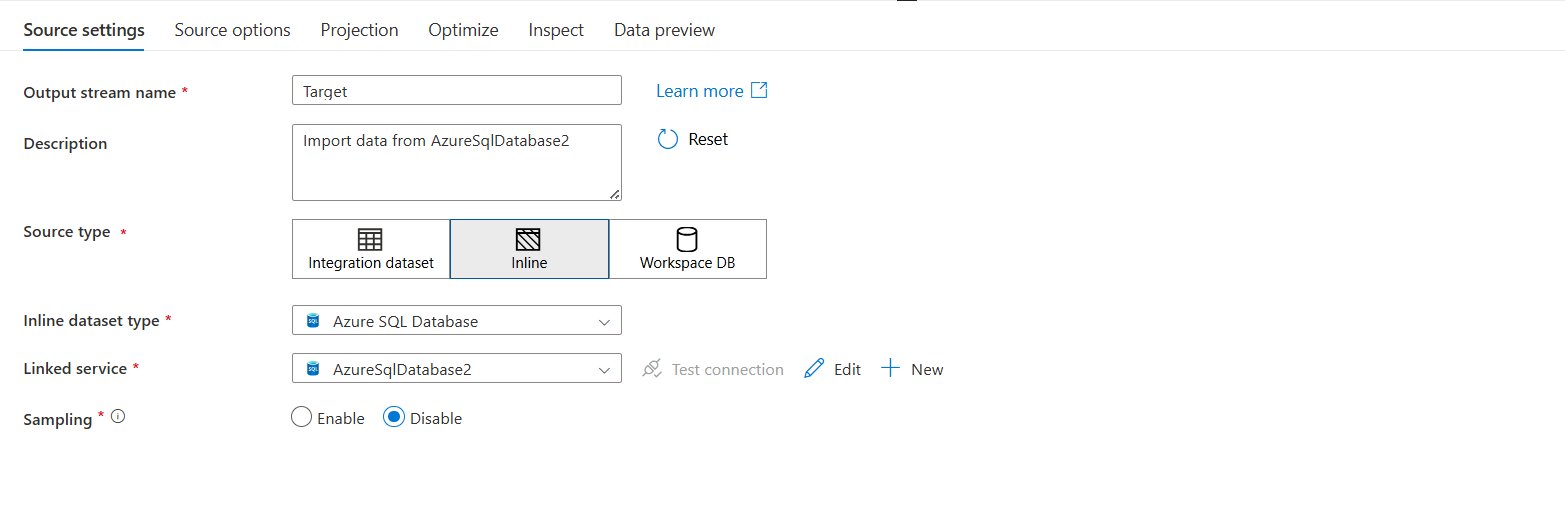
crc32(concat(toString(src\_CusId),src\_CusName,src\_CustCity,toString(src\_CustPhoneNumber)))





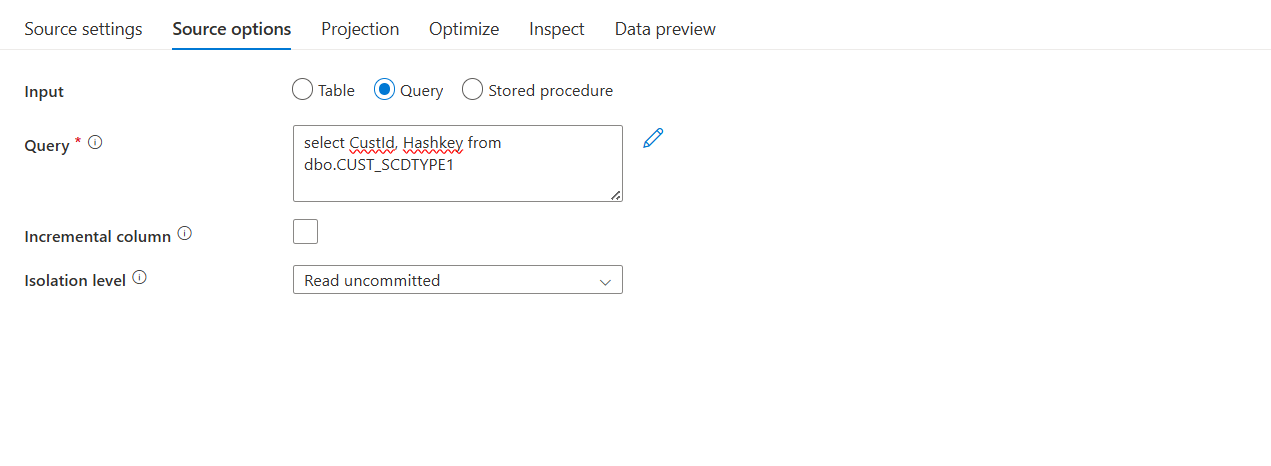


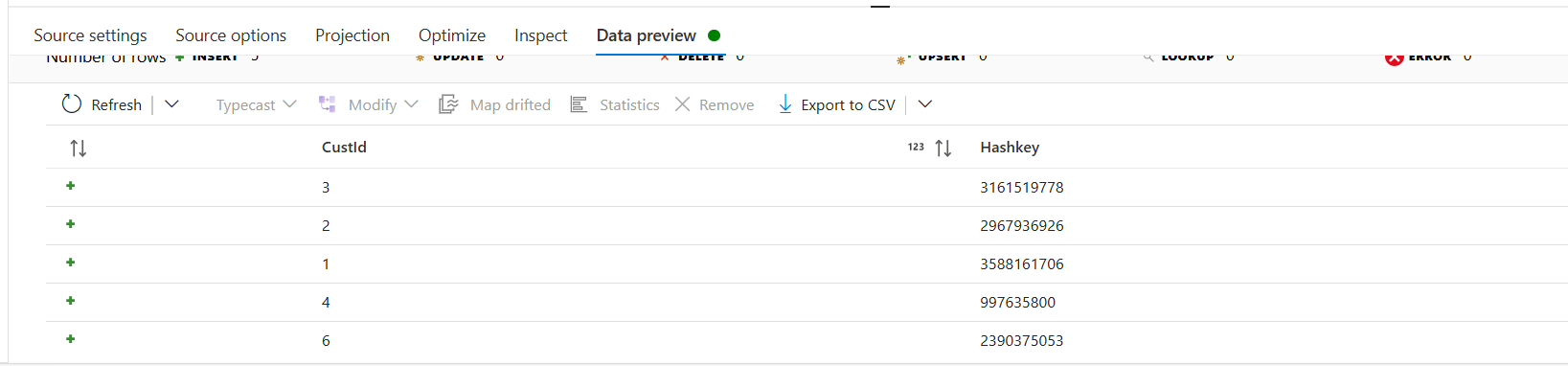
* Now we are using a source settings to take two columns from Database to compare what values are there in the database and CSV file



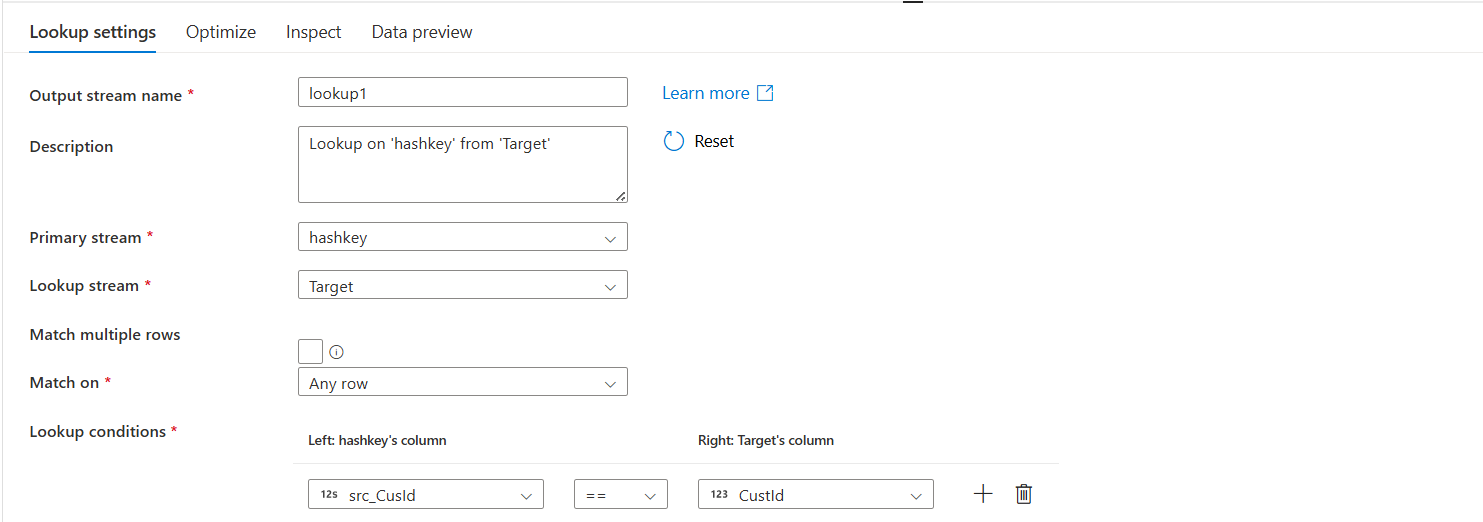
* Under source operation we need to select query and write the below query

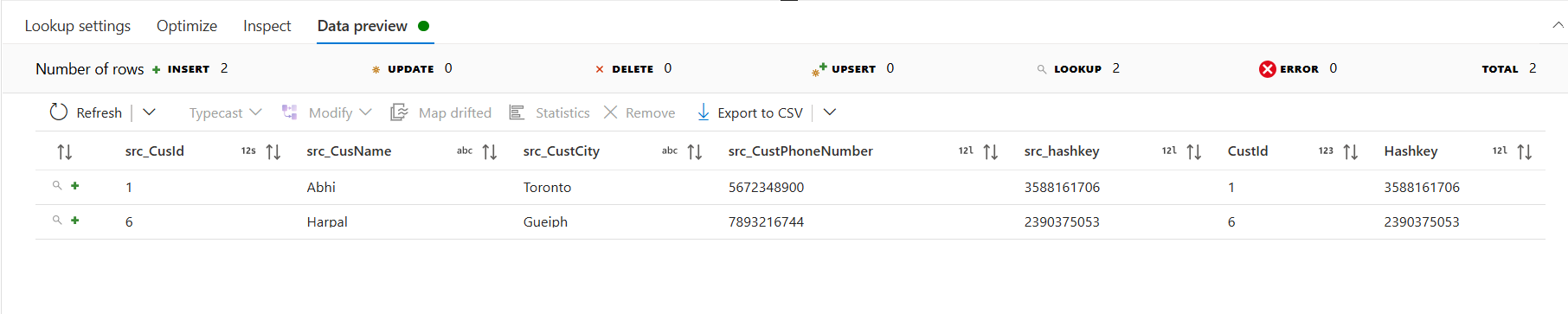
select CustId, Hashkey from dbo.CUST\_SCDTYPE1





* Following to that a lookup transformation is used in which the primary stream will point to Hashkey and lookup stream should connect to Target
* In Lookup condition the Src\_custid should be equal to CustId



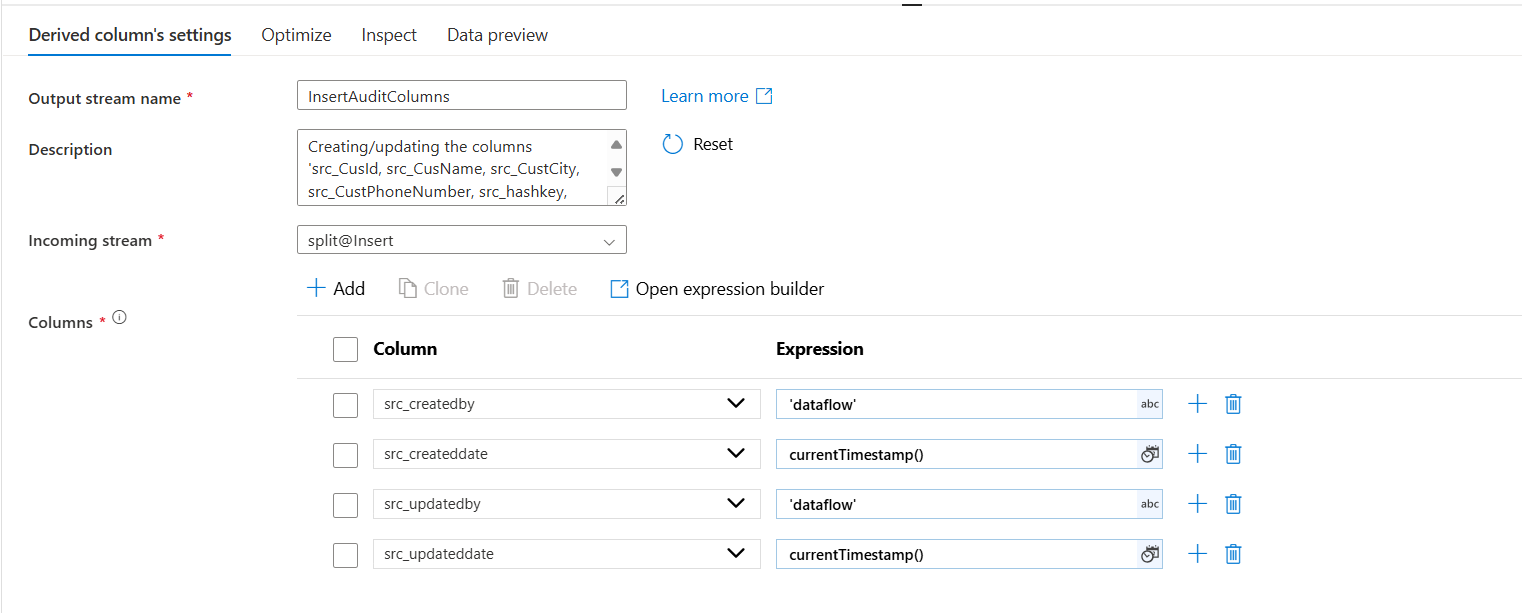


* A conditional split is used to split Insert and Update details, the below expression is used

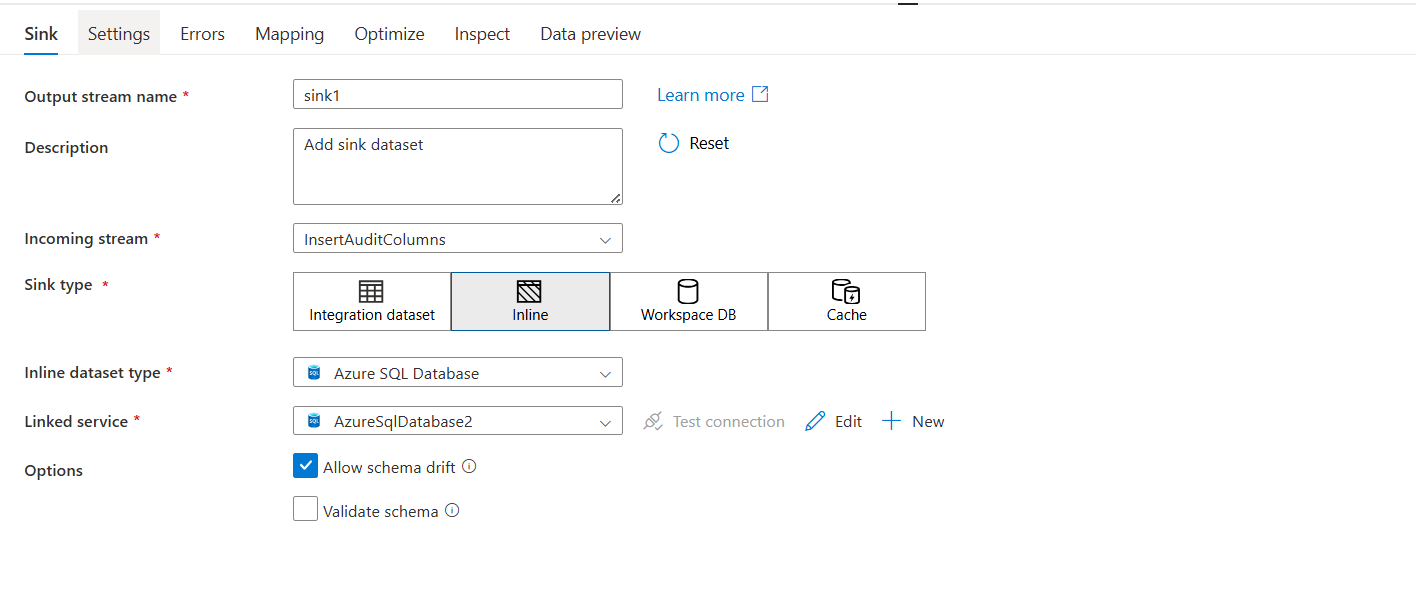
Insert – isNull(CustId)

Update -- src\_CusId==CustId && src\_hashkey !=Hashkey

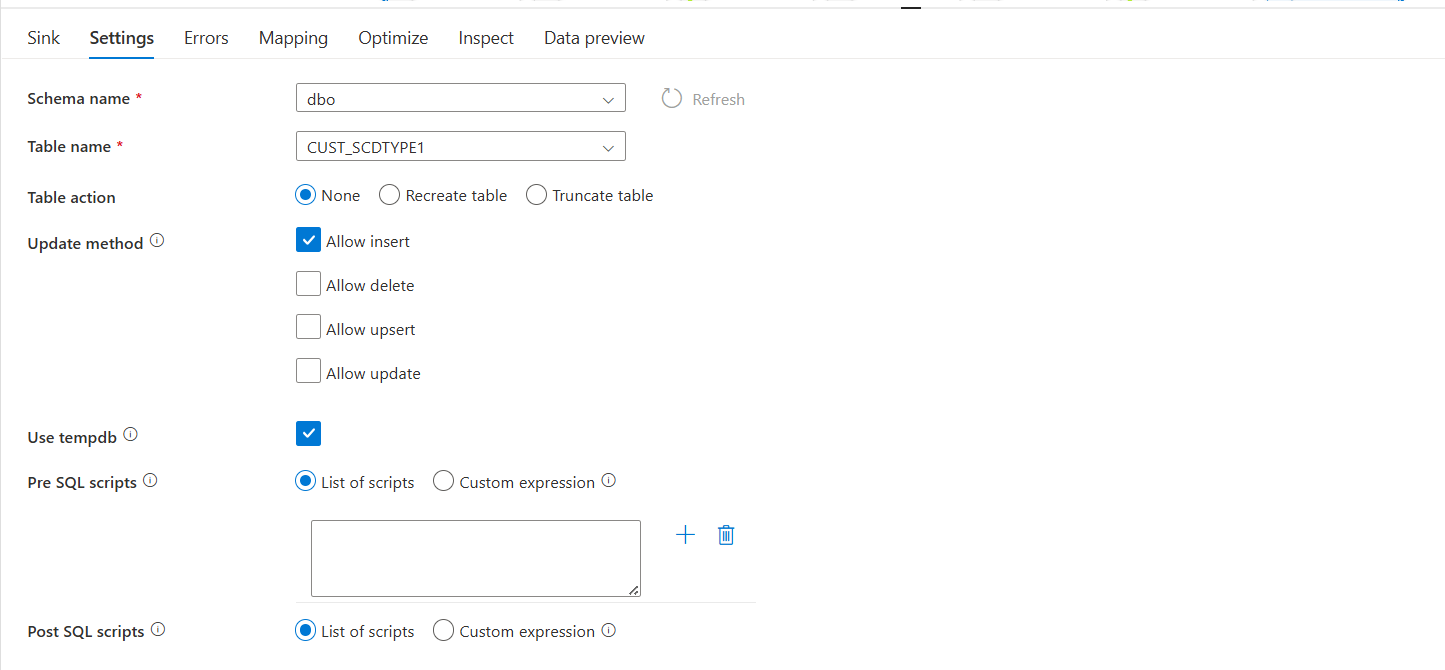
* Next, a audit column has been added to Insert to add createdby, createddate, updatedby, updateddate



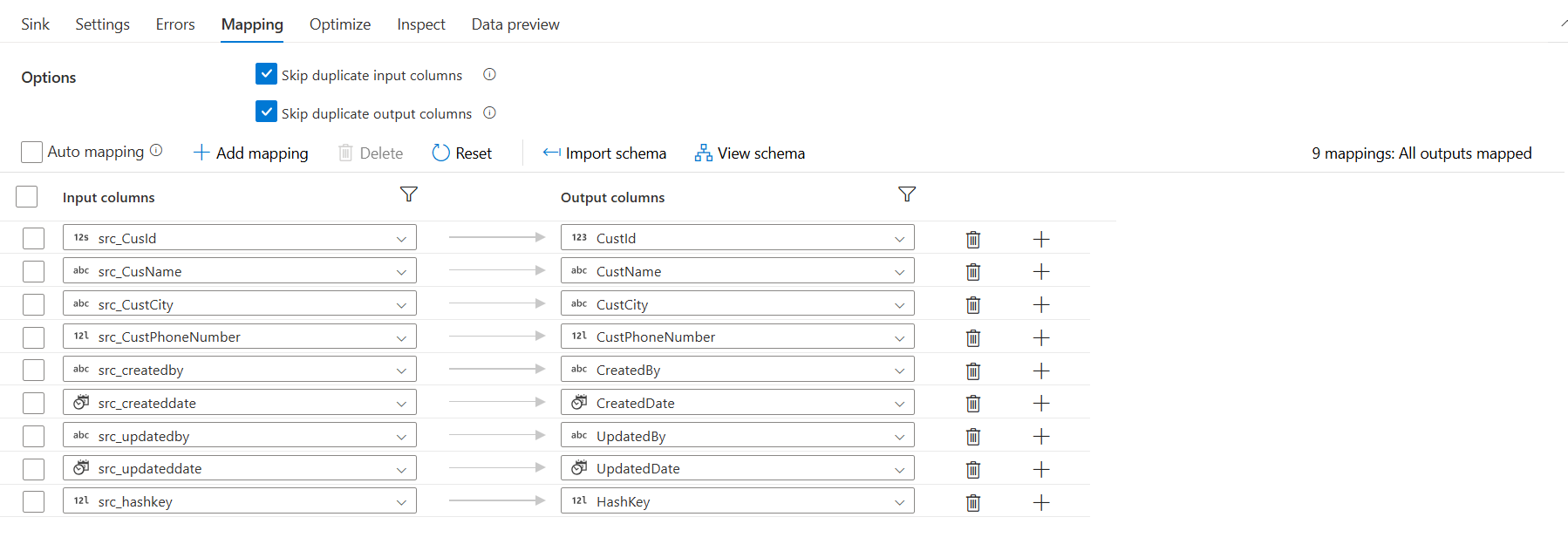
* Use a sink at end to push all the data to the database



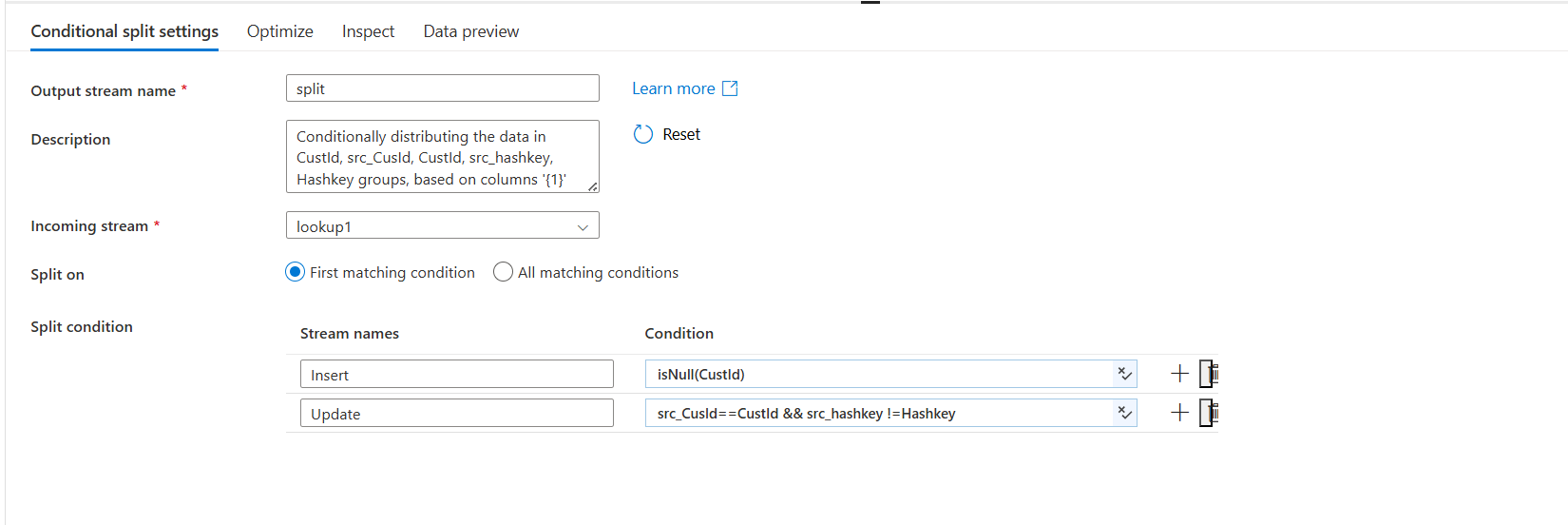
* In source setting we need to select allow insert in update method

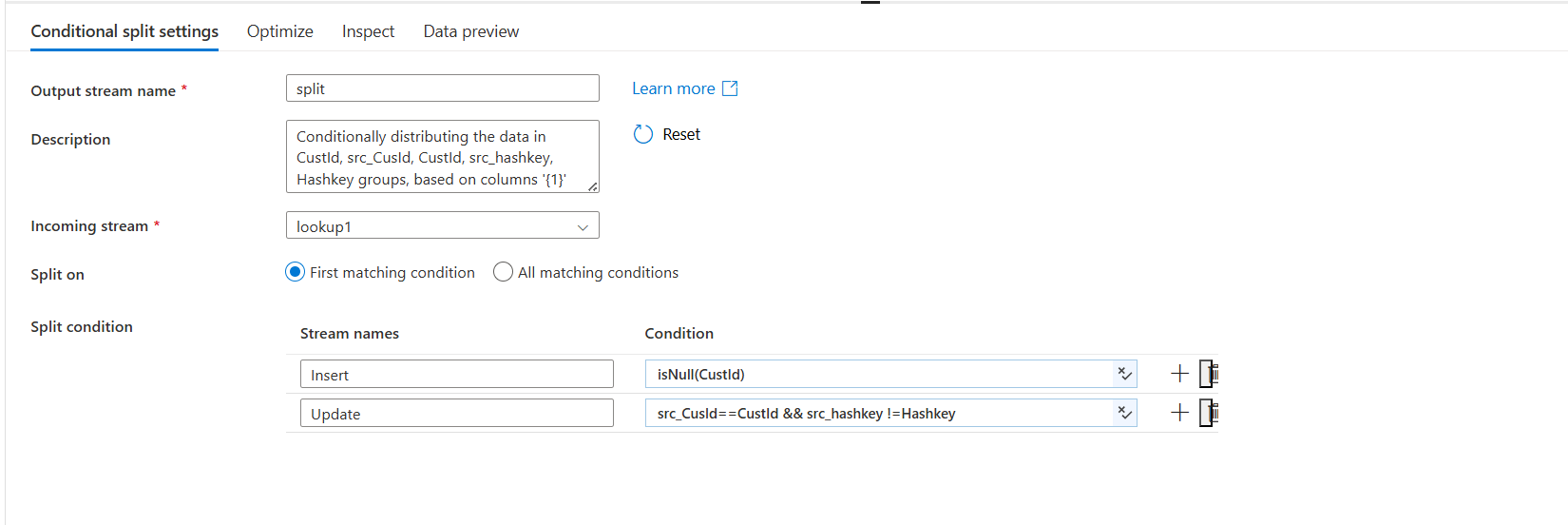


* In mapping we need to map all the columns with their respective columns

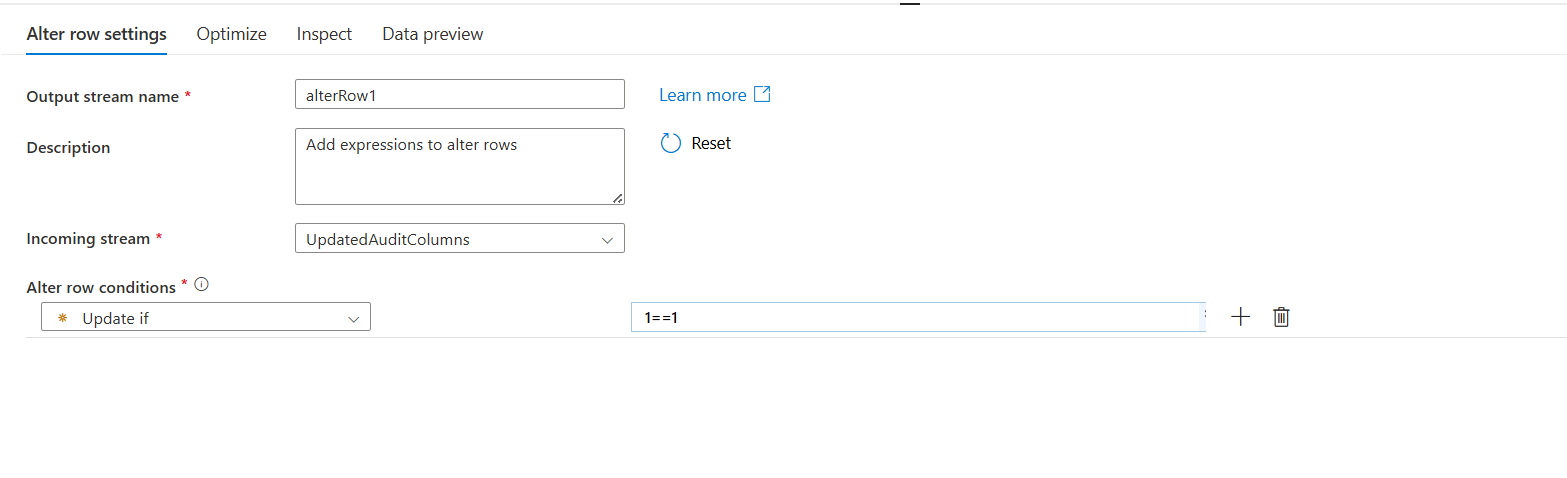


* Next, a audit column has been added to Update to add updatedby, updateddate
* In updatedby we need to give “dataflow” and in updated date we need to give currentTimestamp()

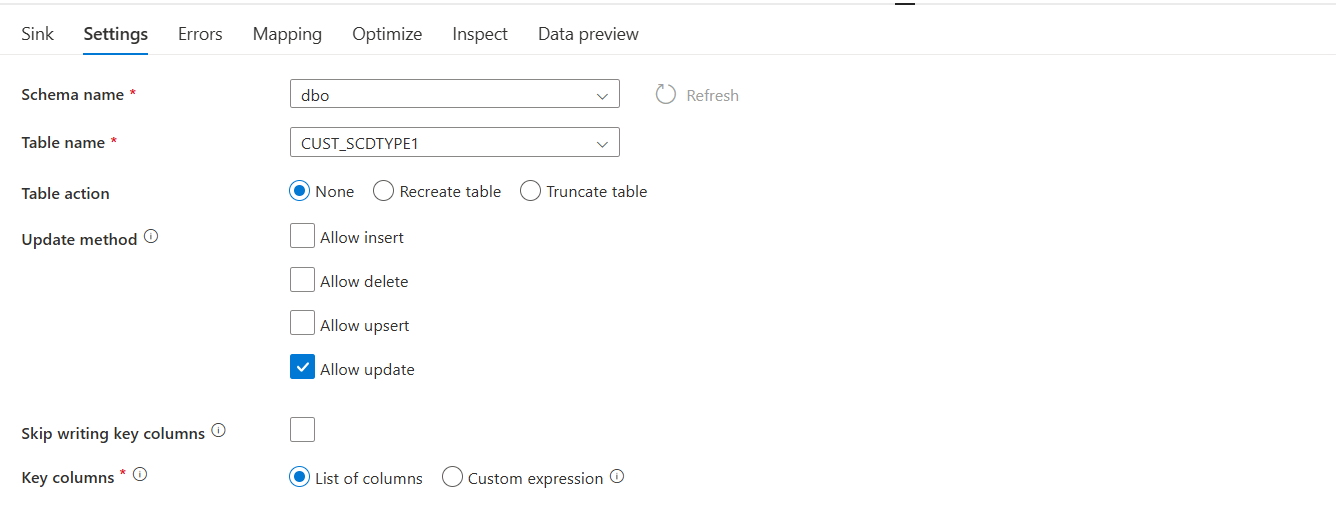




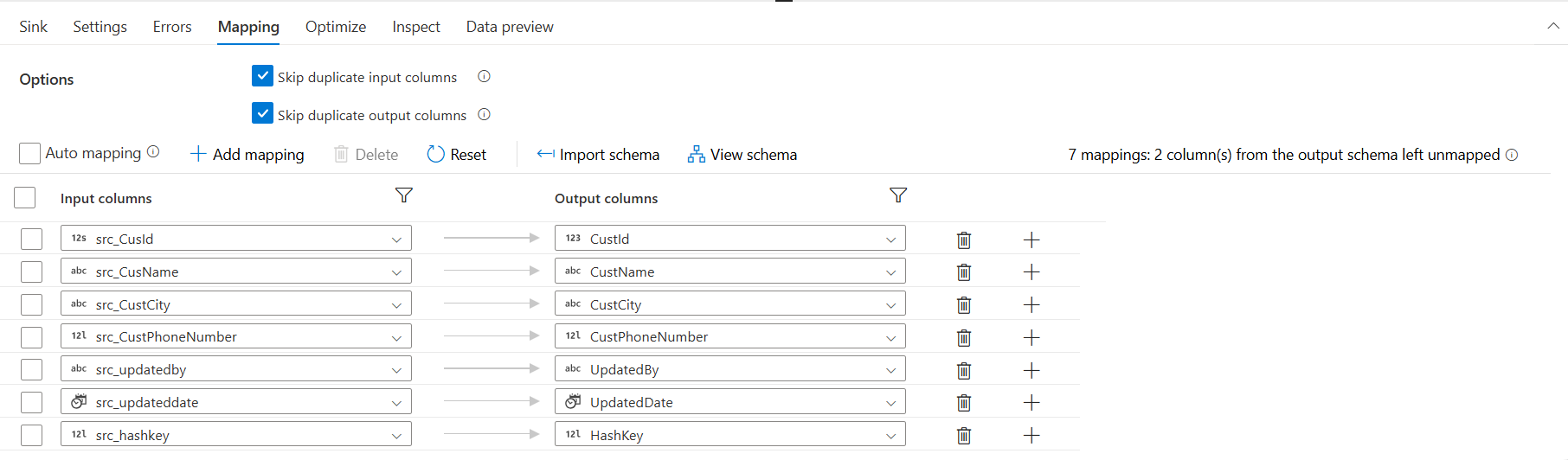
* An alter row has been used to pass a condition



* Use a sink at end to push all the data to the database



* In source setting we need to select allow insert in update method
* In mapping we need to map all the columns with their respective columns



* The final output will look like below in Azure SQL Database

