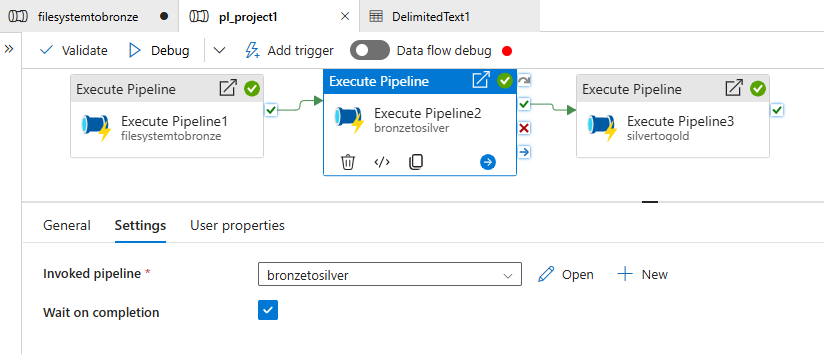
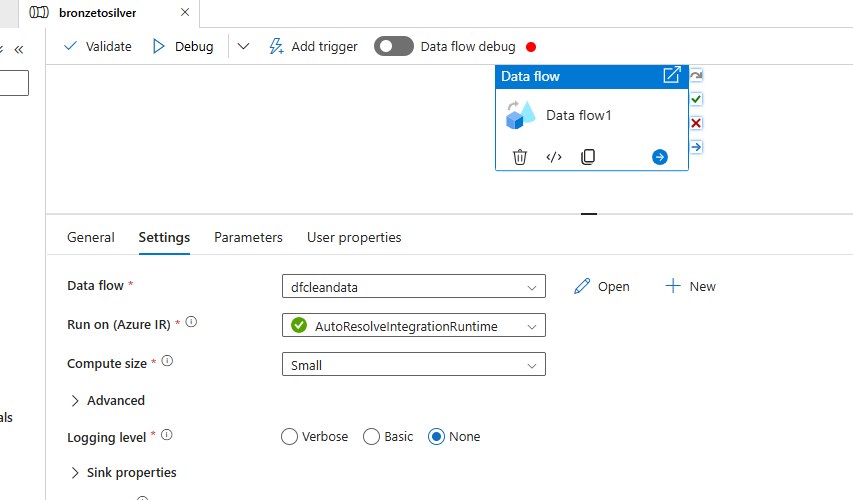
**Step 2:** **Databricks Activity (Incremental/Delta Processing)**

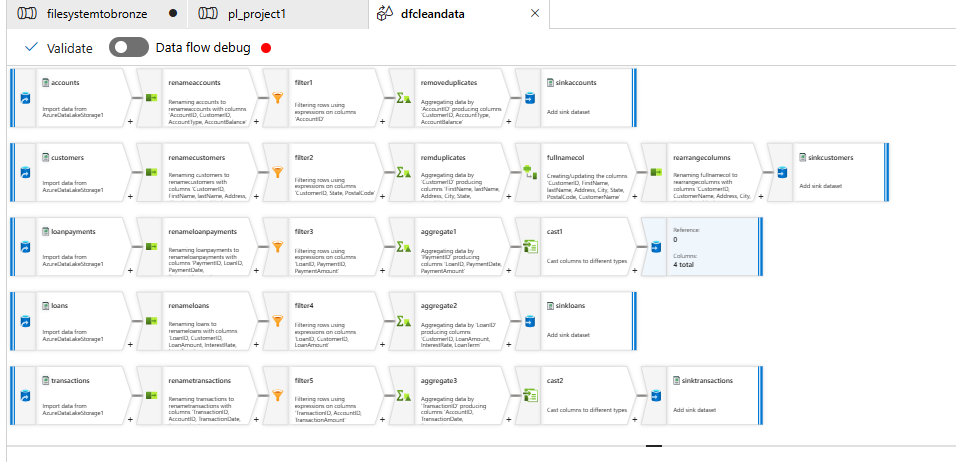
Here we have to take another execute activity in same master pipeline and give another pipeline named as bronze to silver

open the bronze to silver pipeline

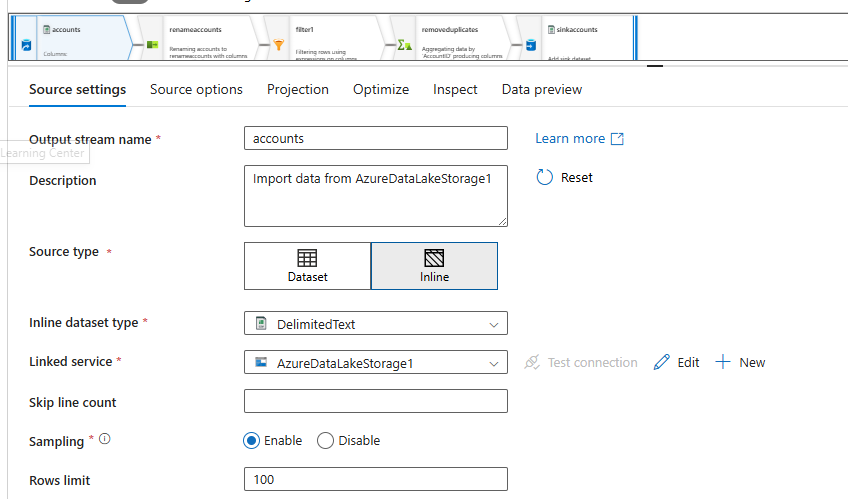
Create a dataflow to clean data in bronze layer and copy to silver layer



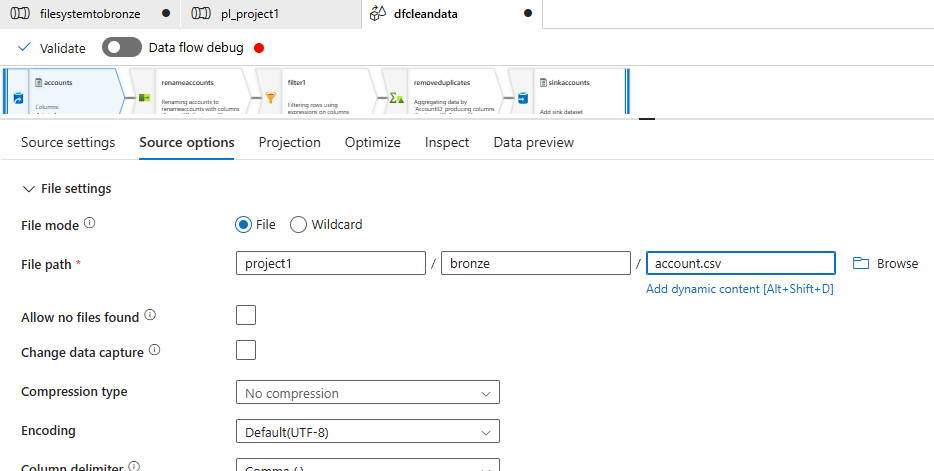
Create five separate sources and use transformations such as select to rename columns, filter to remove nulls, aggregate for distinct values to remove duplicates and add sink for each source in the bronze layer to silver layer as follows



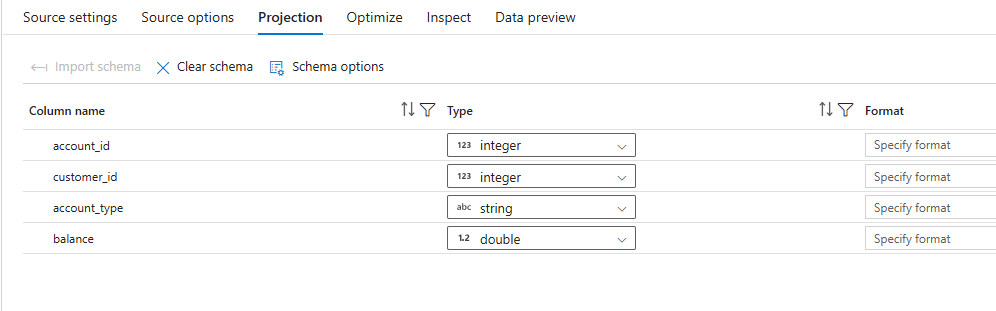
Let us go to one source accounts.csv and see the steps inside it



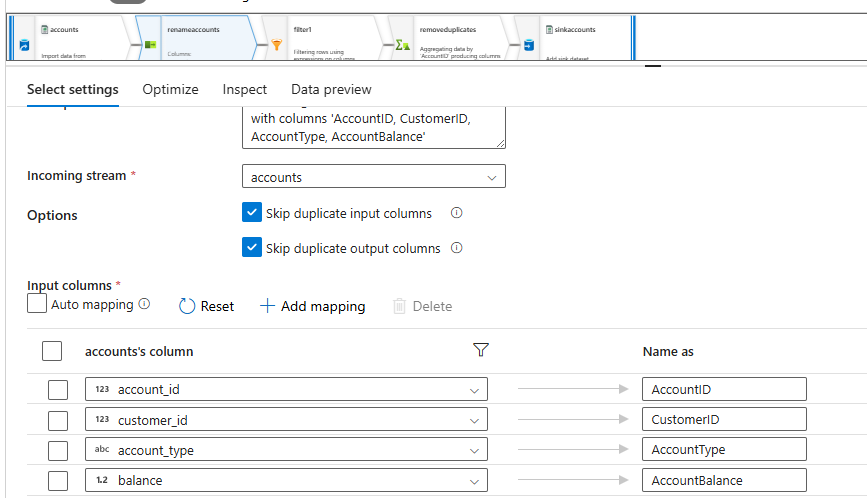
Go to source options for accounts file



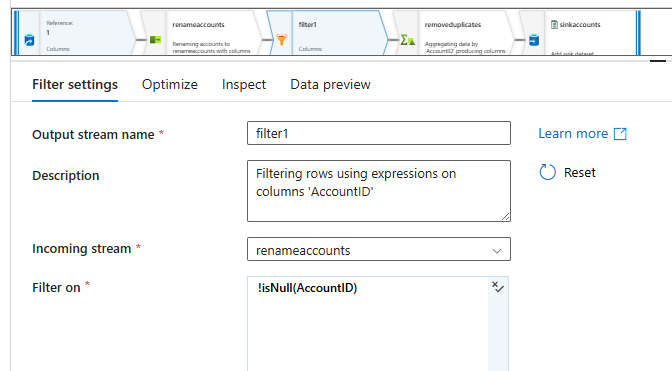
After that go to projections and import schema



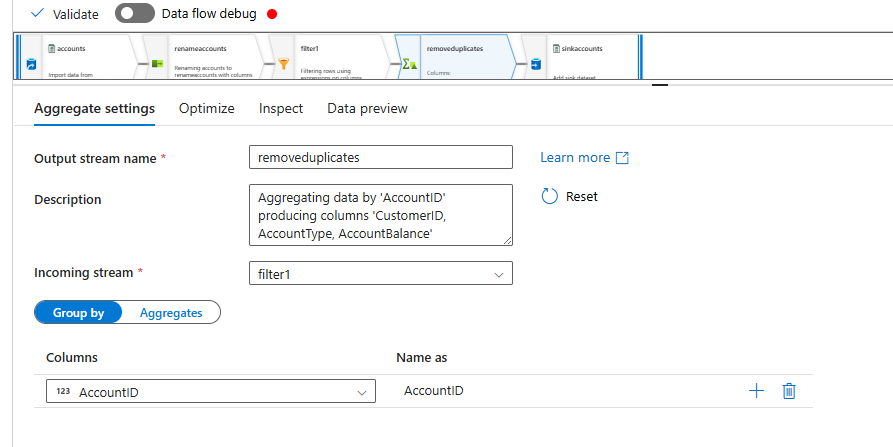
Add a select transformation to rename columns

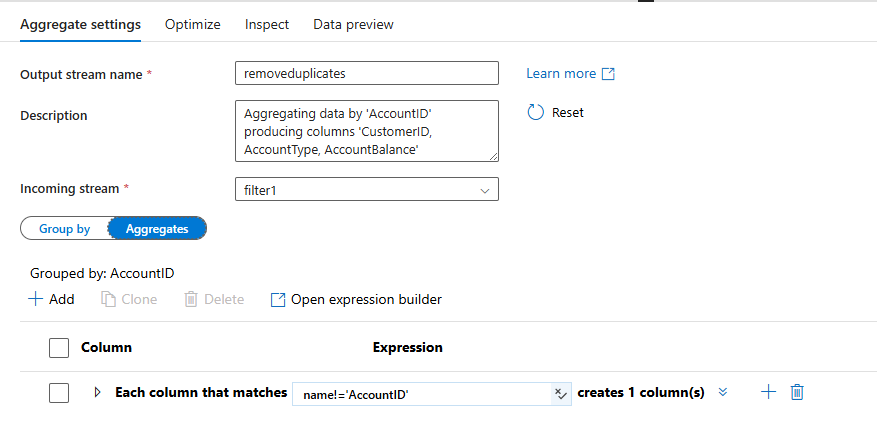


Now add a filter to remove nulls

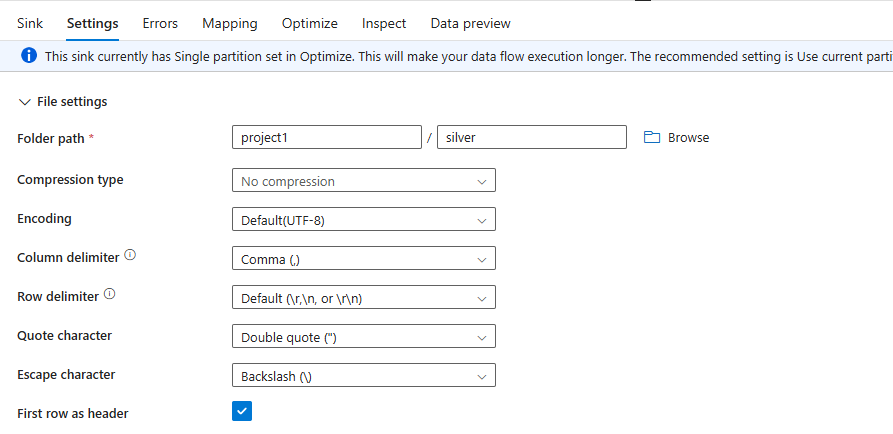


Now add aggregate to remove duplicates

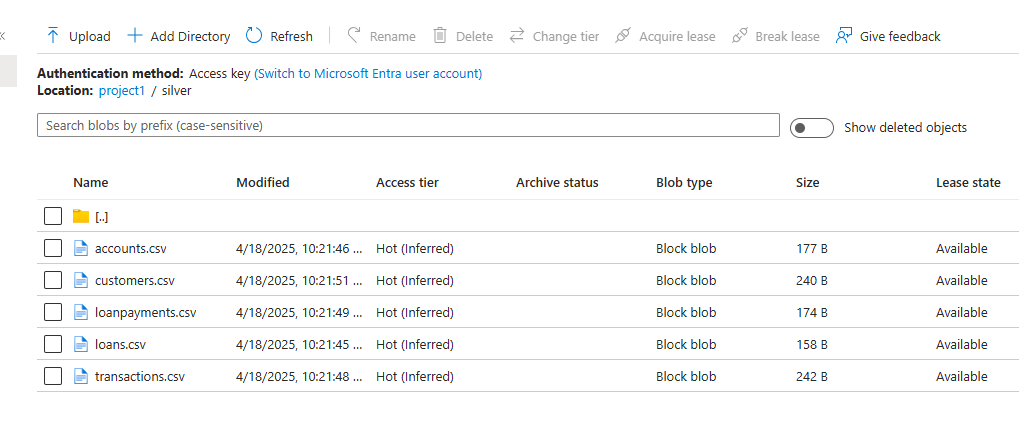




Add sink finally



Repeat the same steps for every file and then the sink silver layer as follows



Note: for day 2 files change source file location for every file to second day and repeat the steps as above and debug the entire pipeline from start so that all data will be copied to bronze and cleaned in silver layer