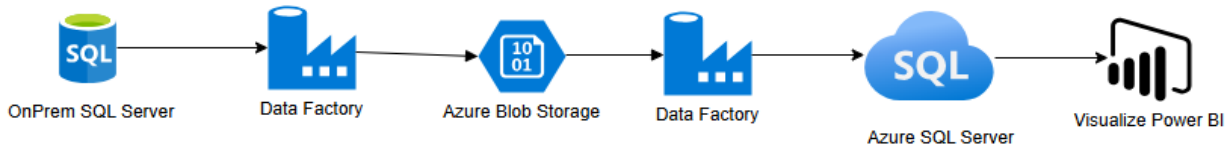
**Data Driven Data Pipeline for Patient Admission Analysis.**

**Objective:**

The Projects aims to design and implement a meta data driven pipeline in Azure Data factory for processing Patient admission data from an on prem SQL Server OLTP system to a data warehouse in Azure SQL Server.

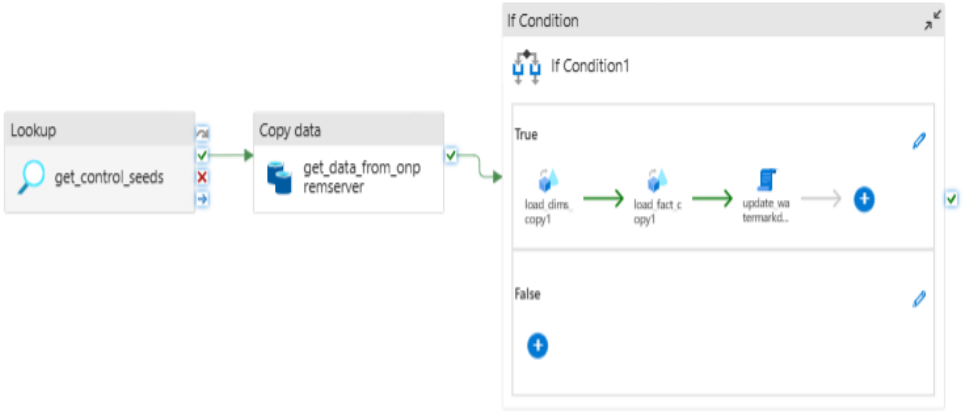
**Architecture:**



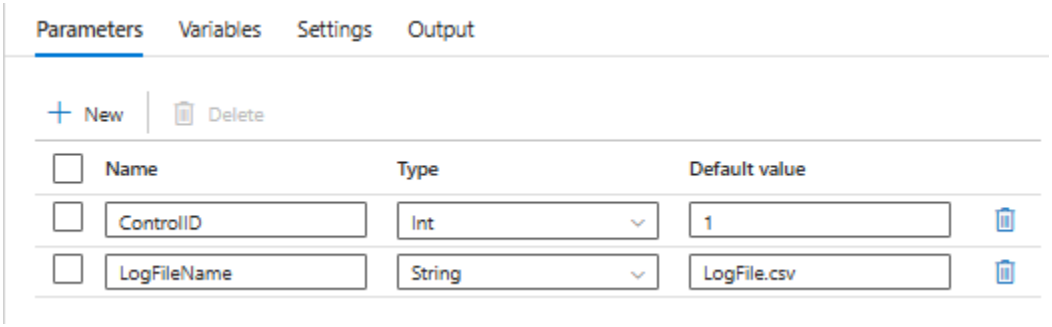

The structure of the data stored in the Patient Admission Table in the OLTP system is as follows:

| PatientAdmissionID | Name | Age | Gender | BloodType | MedicalCondition | DateOfAdmission | Doctor | Hospital | InsuranceProvider | BillingAmount | RoomNumber | AdmissionType | DischargeDate | Medication | TestResults | LastModifiedat |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Bobby JacksOn | 30 | Male | x- | Cancer | 2024-01-31 | Matthew Smith | Sons and Miller | Blue Cross | 18856.28 | 328 | Urgent | 2024-02-02 | Paracetamol | Normal | 2025-04-26 09:37:43.850 |
| 2 | LesLie TErRy | 62 | Male | A+ | Obesity | 2019-08-20 | Samantha Davies | Kim Inc | Medicare | 33643.33 | 265 | Emergency | 2019-08-26 | Ibuprofen | Inconclusive | 2025-04-22 00:00:00.000 |
| 3 | DaNnY sMitH | 76 | Female | A- | Obesity | 2022-09-22 | Tiffany Mitchell | Cook PLC | Aetna | 27955.10 | 205 | Emergency | 2022-10-07 | Aspirin | Normal | 2025-04-22 00:00:00.000 |
| 4 | andrEw waTtS | 28 | Female | O+ | Diabetes | 2020-11-18 | Kevin Wells | Hernandez Rogers and Vang, | Medicare | 37909.78 | 450 | Elective | 2020-12-18 | Ibuprofen | Abnormal | 2025-04-22 00:00:00.000 |

1. An Insert/Update trigger has been created on the Patient Admission table which updates the lastmodifiedat column with current date, every time a record has been updated or inserted in the table. This enables the changes to be propagated to the downstream systems.
2. The Master Pipeline which consists of the following steps below moves the delta changes based on watermark date from the OLTP system to the Data warehouse.



3. The master pipeline accepts 2 parameters as below.

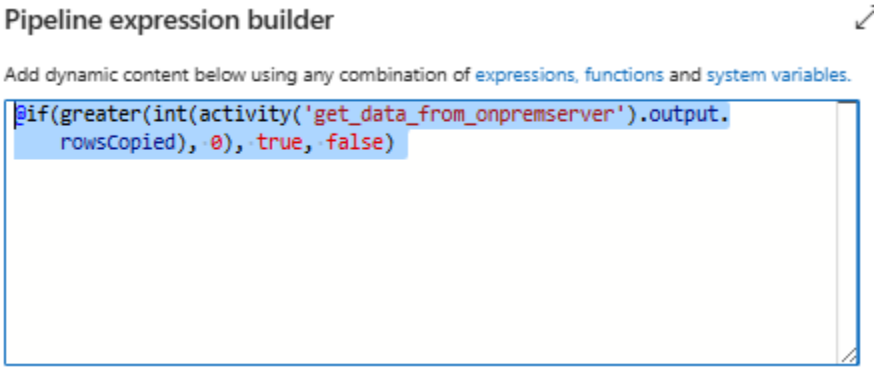



4. **Lookup Activity : get_control_seeds**

   The ControlID passed fetches information such as secrets, table names and other related information which is used in the subsequent steps. This approach makes the pipeline data driven.

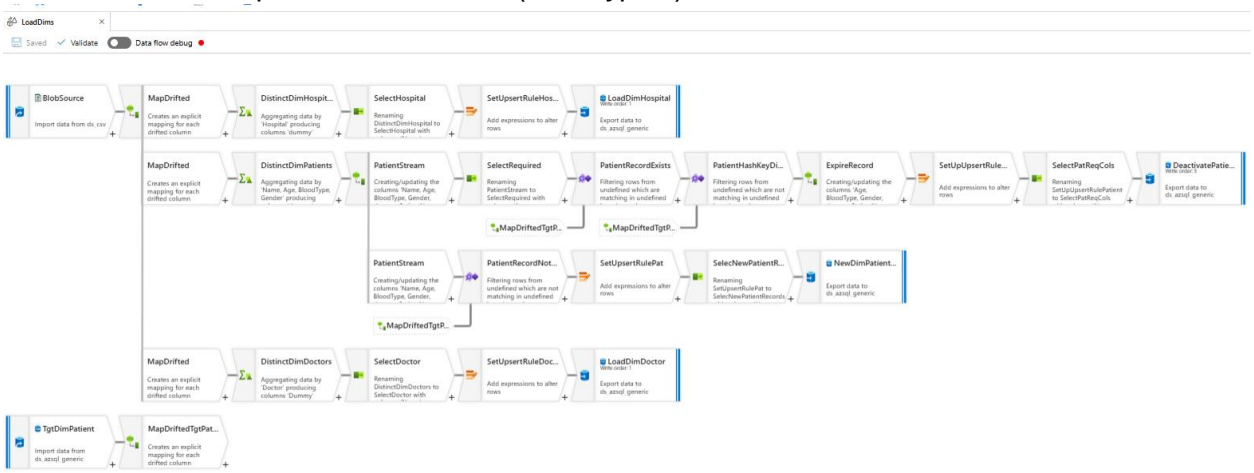
5. **Copy Activity : get_data_from_onpremserver**
   The copy activity fetches only deltas based on watermark date and exports them in CSV format into a blob storage.

6. IfCondition : The IF Condition checks if any rows have been copied.

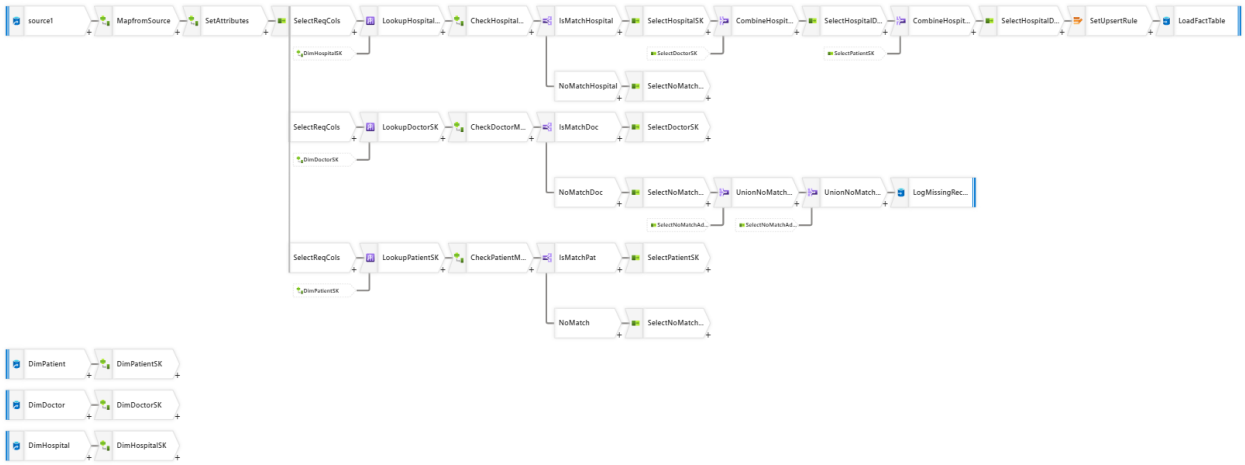



If any rows have been copied then only the subsequent data flows would be executed.


7. Load_Dims data flow splits the source data into dimension streams for loading DimDoctor,DimHospital and DimPatient(SCD Type 2)



8. Load_Facts Data Flow loads the facts.




9. Finally a view is created in the azure sql datawarehouse, this view is further used to create a dashboard in power bi.