Delta Report

1. Update the Constellation Schema

(DimTaxi: Change into Type 1)

(DimAirbnb\_Property: Change into Type 2 – Adding end\_date and is\_current\_flag)

(FactTrip: Merge or Drop the duplication made by Location\_id and Date\_id, remove pickup/dropoff datetime\_key as well)

A picture containing text, handwriting, diagram, parallel

Description automatically generated

SQL Query:

A screenshot of a computer

Description automatically generated with medium confidence

A screenshot of a computer

Description automatically generated with medium confidence

Drop Price\_prev: (ALTER TABLE DimAirbnb\_Property

DROP COLUMN Price\_prev;

Perform a SCD Type 2 on DimAirbnb\_Property:

Redesign FactTrip into FactTaxiTrip:

SELECT Taxi\_id, MAX(Location\_id) AS Location\_id, MAX(Date\_id) AS Date\_id, SUM(Passenger\_count) AS Passenger\_count,

SUM(Trip\_distance) AS Trip\_distance, SUM(Trip\_duration) AS Trip\_duration, SUM(Fare\_amount) AS Fare\_amount,

SUM(Total\_amount) AS Total\_amount, MAX(Vendor\_id) AS Vendor\_id

INTO FactTaxiTrip

FROM FactTrip

GROUP BY Taxi\_id;

ALTER TABLE FactTaxiTrip

ADD TaxiTrip\_id INT IDENTITY(1,1) PRIMARY KEY;

A screenshot of a computer

Description automatically generated

Then drop the FactTrip table.

Improvement:

1. Adding more realistic metrics in DimTaxi and FactTaxiMonthlySummary to calculate the real distance between Airbnb spot and Taxi pickup or dropoff location. Also, need to record the street name near the Airbnb location to make it easier for Taxi to know whether can pickup or dropoff location for passengers.
2. In BQ 1, we should change the units of Airbnb rental rate to rental price to have the same units to compare with the taxi fare.
3. Create a Dim table to store the month column in FactTaxiMonthlySummary. In this way, we can increase the efficiency of loading the query. Also, the move can help us reduce the total time complexity in this FactTable.
4. The challenge is that it’s hard to find the perfect match dataset which are highly correlated to each other. Some columns in Airbnb tables are not enough to provide a linkage with taxi taking rate.