Dbt snowflake assesments

Set up snowflake

Install dbt in our local (vscode)

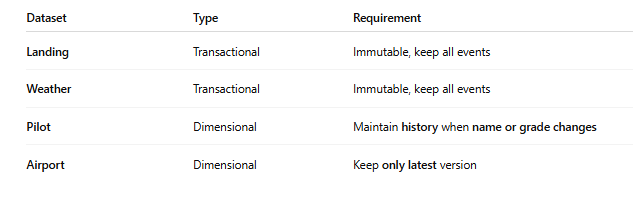
We have given **4 JSON datasets**:

1. **Landing**: Data about each plane landing
2. **Weather**: Temperature data recorded every 5 minutes at airports
3. **Pilot**: Pilot details (can change over time)
4. **Airport**: Airport details (can change, but we only care about the latest)

**Business requirements**

**------------------------------**

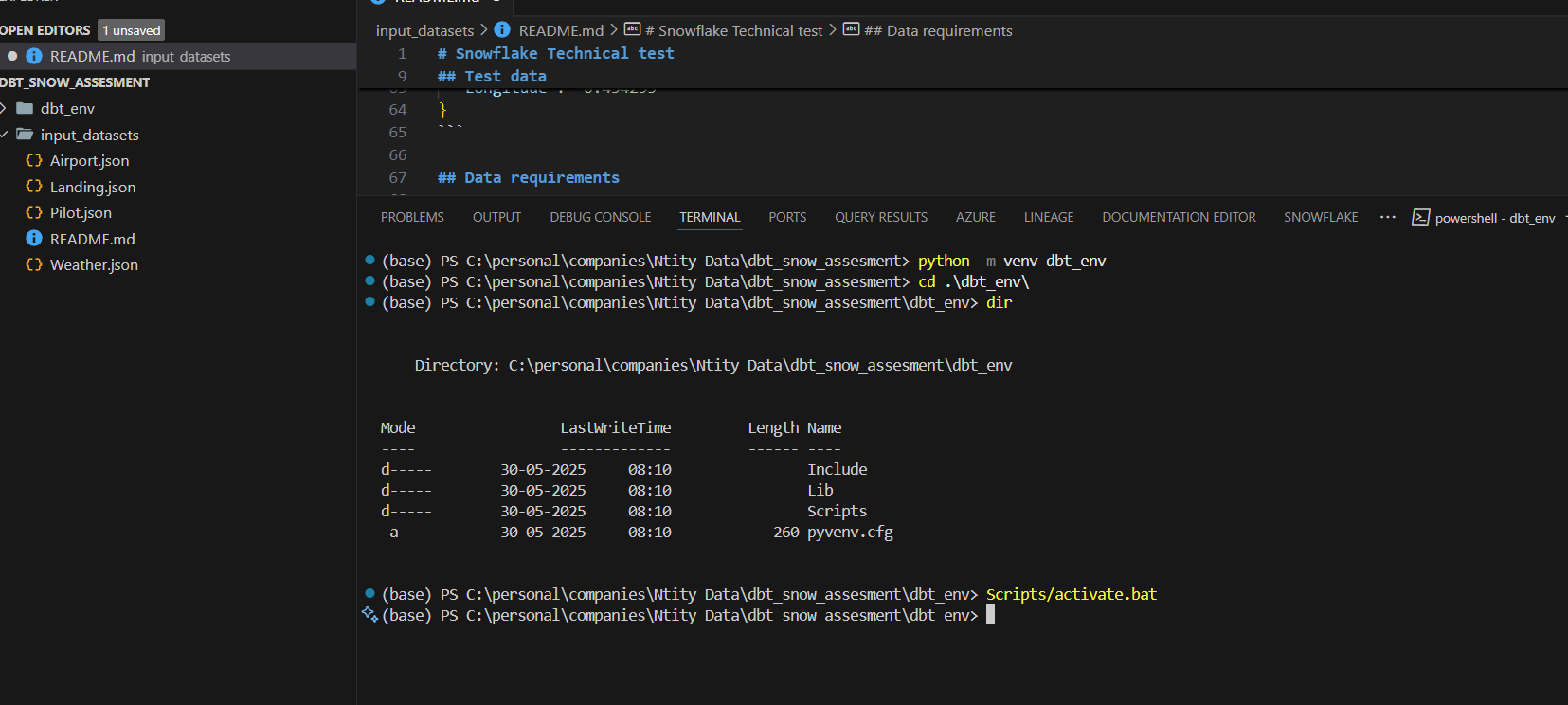
| **Dataset** | **Type** | **Requirement** |
| --- | --- | --- |
| **Landing** | Transactional | Immutable, keep all events |
| **Weather** | Transactional | Immutable, keep all events |
| **Pilot** | Dimensional | Maintain **history** when **name or grade changes** |
| **Airport** | Dimensional | Keep **only latest** version |

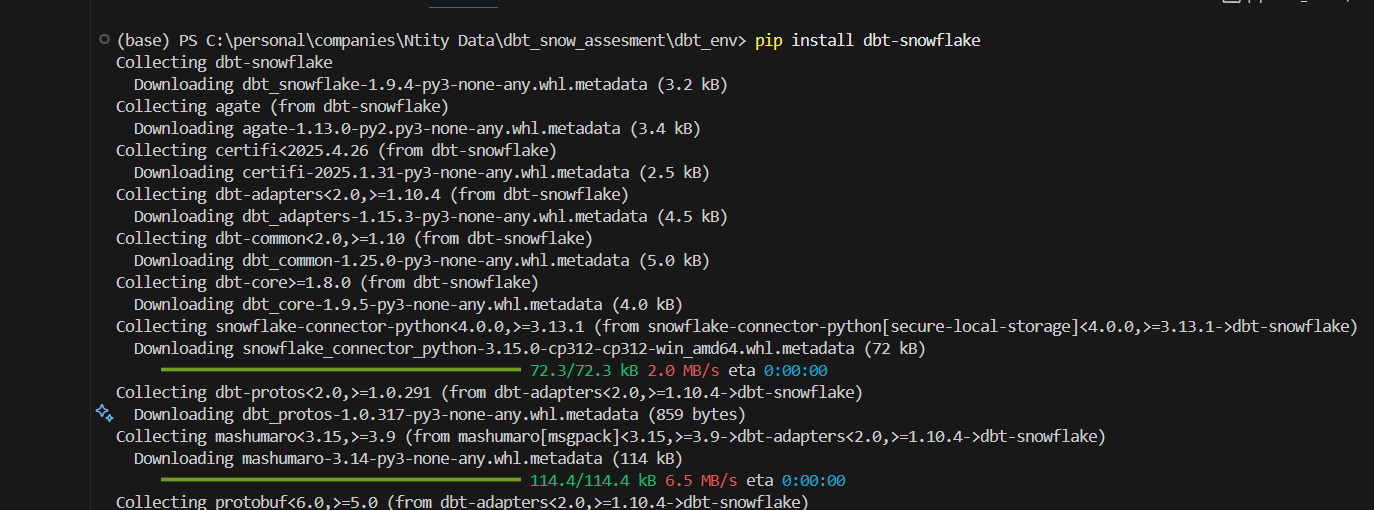


Create snowflake account

Install dbt in vscode

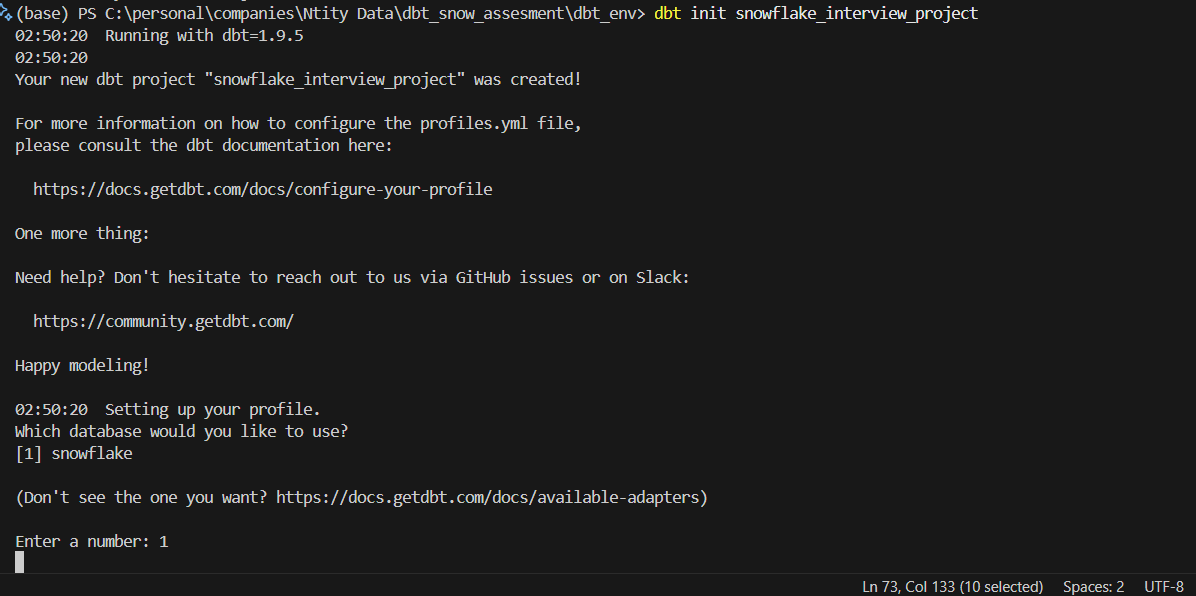
Created a virtual env and activate it

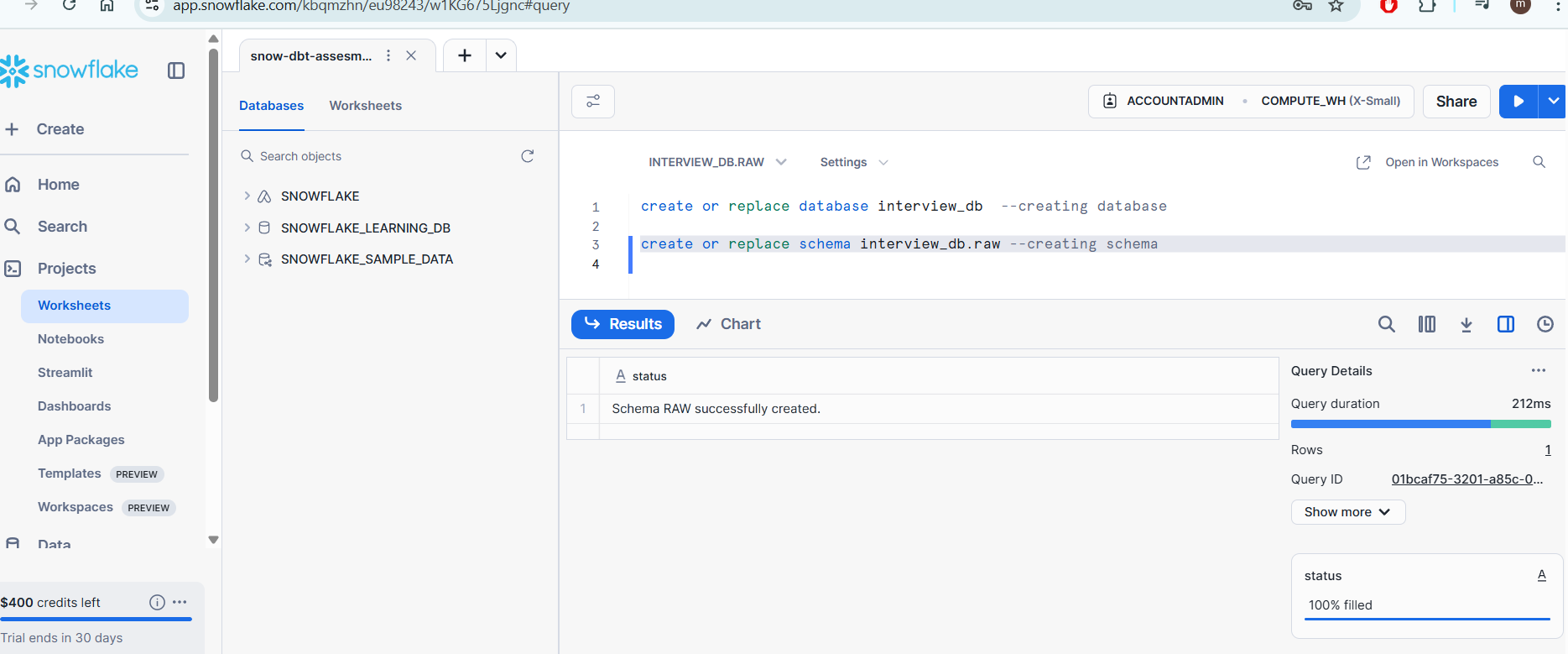




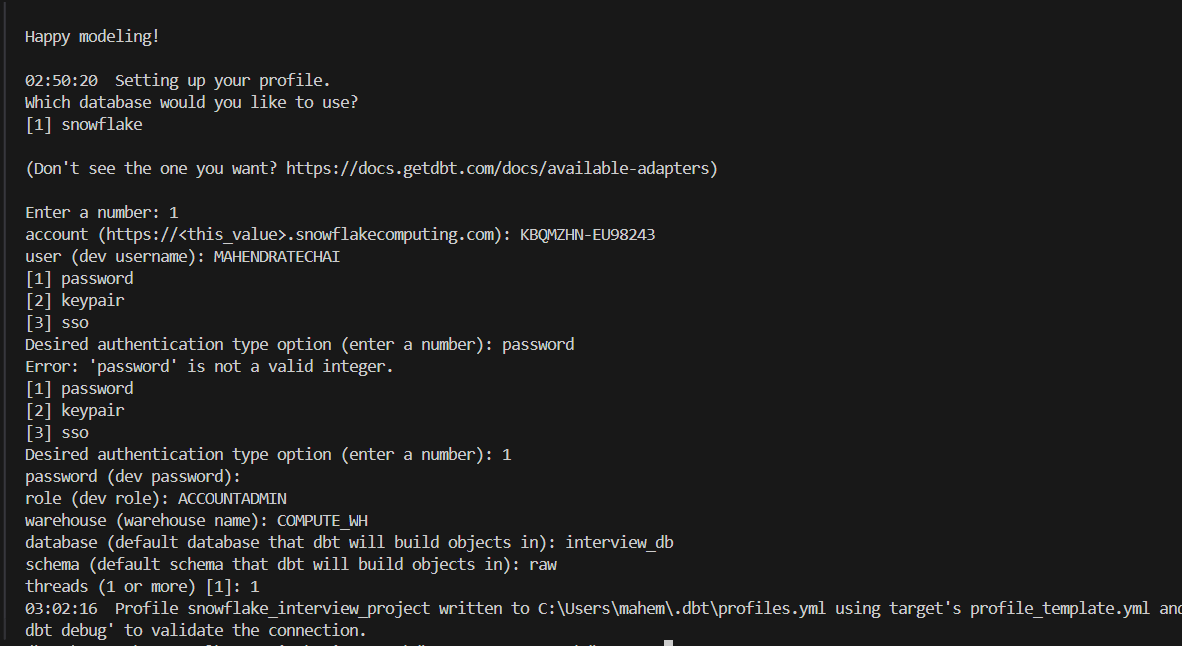
Initialize dbt snow project

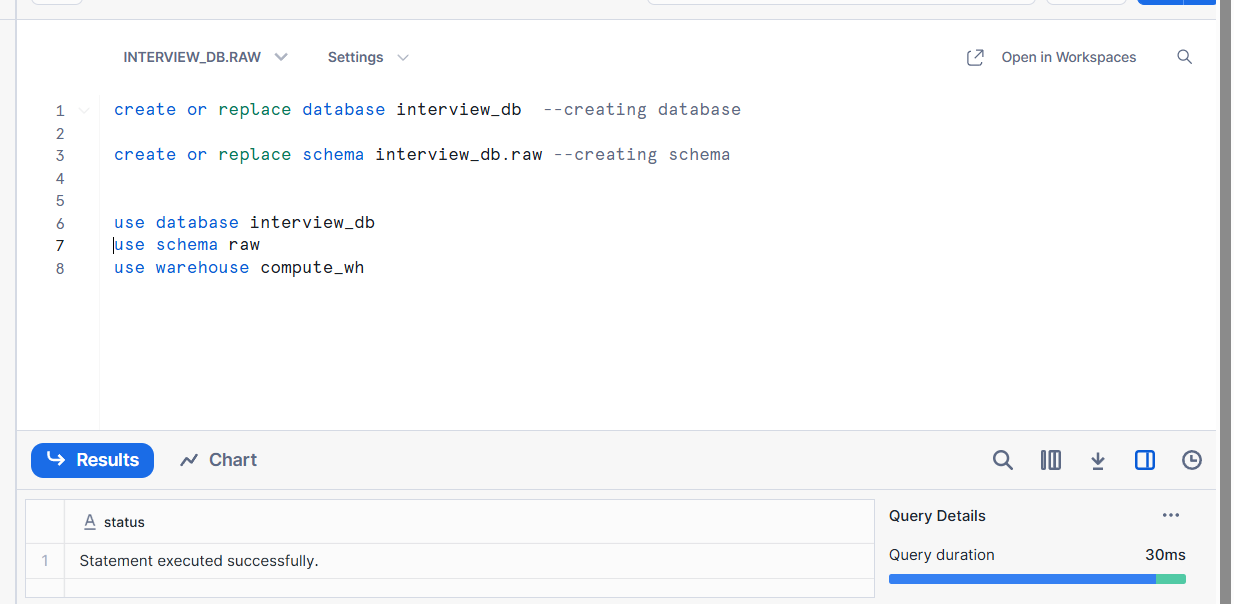
* dbt init snowflake\_interview\_project



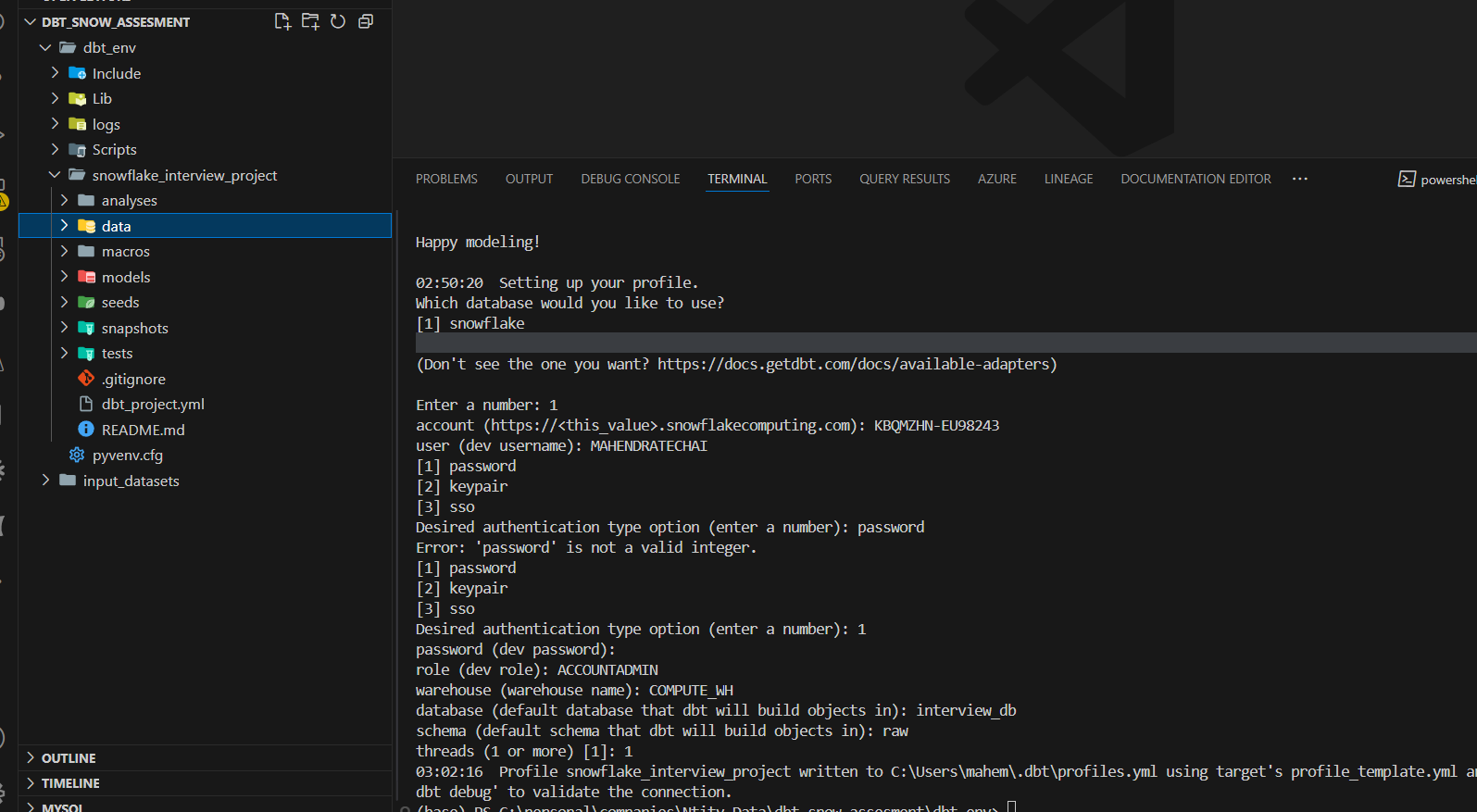


And then enter valid databse and schema name

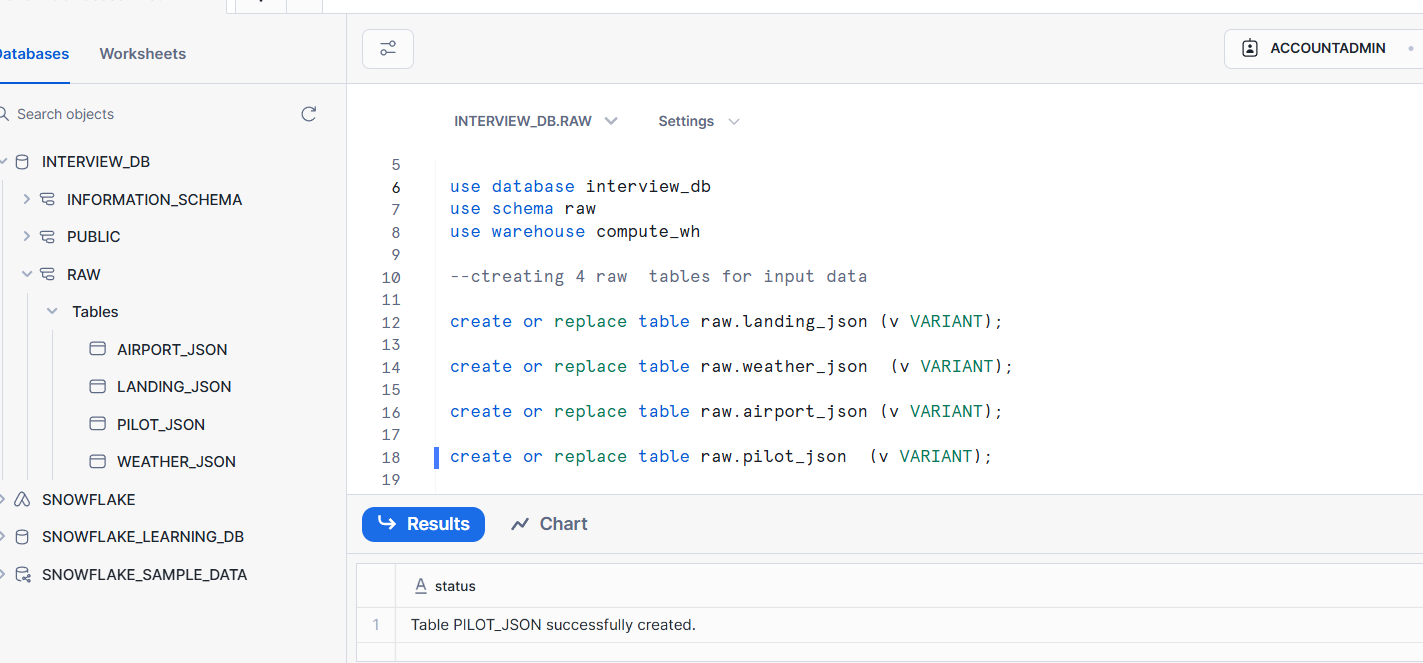




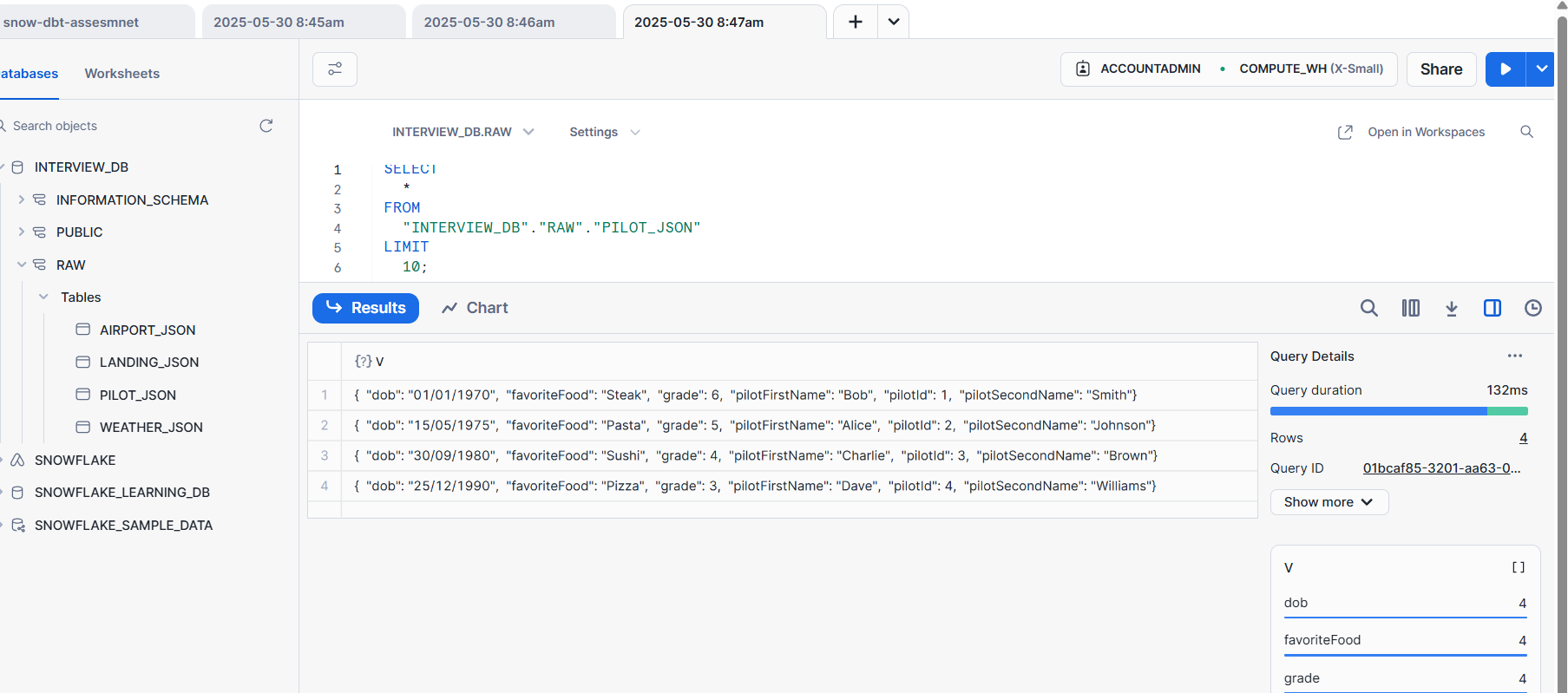
Copied input json files to data folder



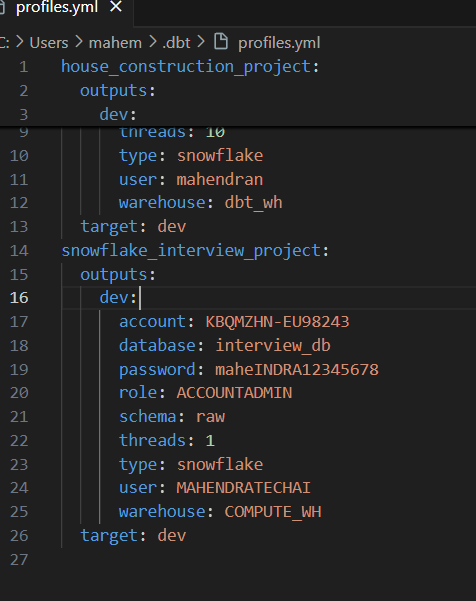
Creating raw tables in snowflake for the above data



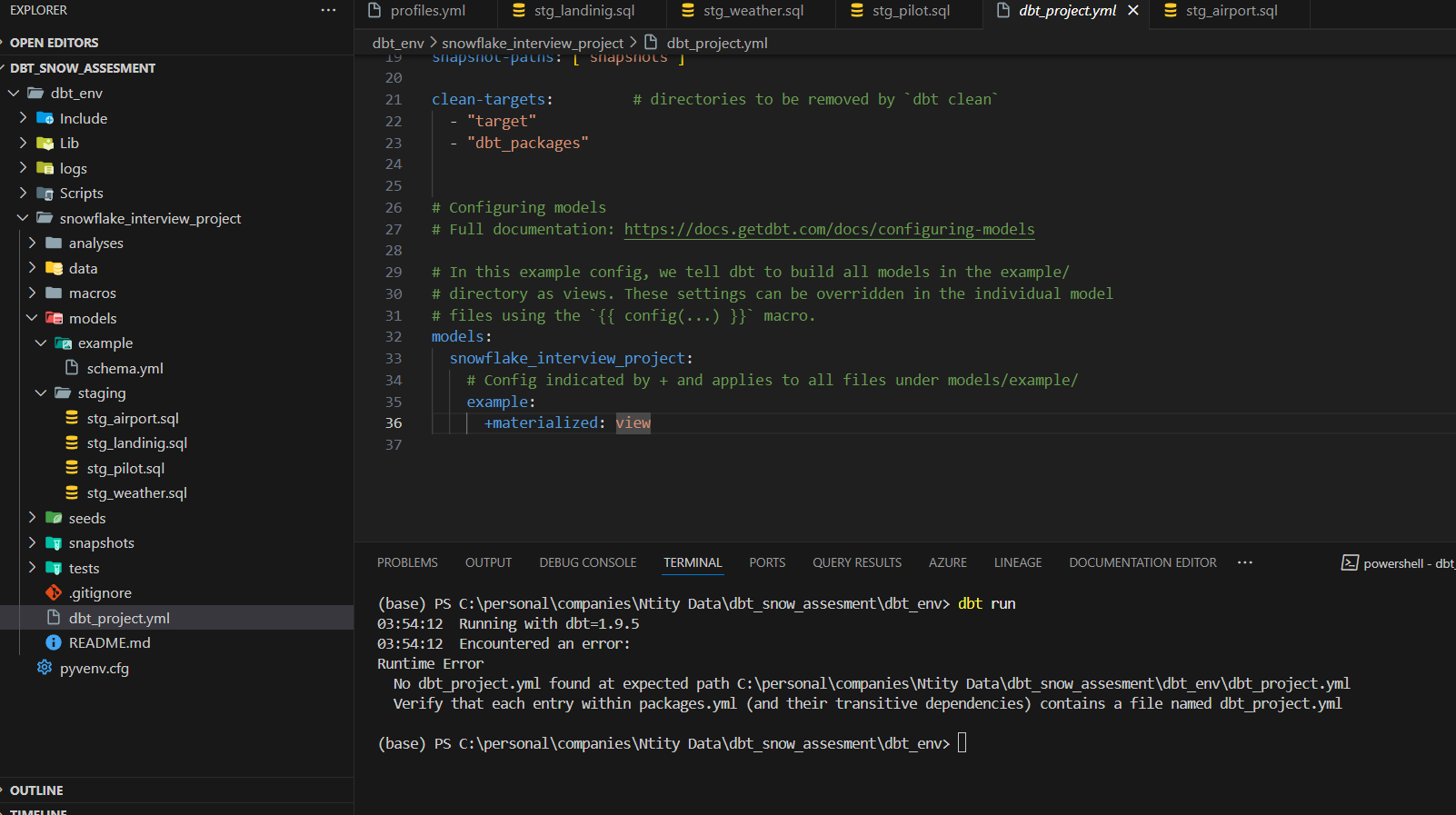
#loading the data into the tables

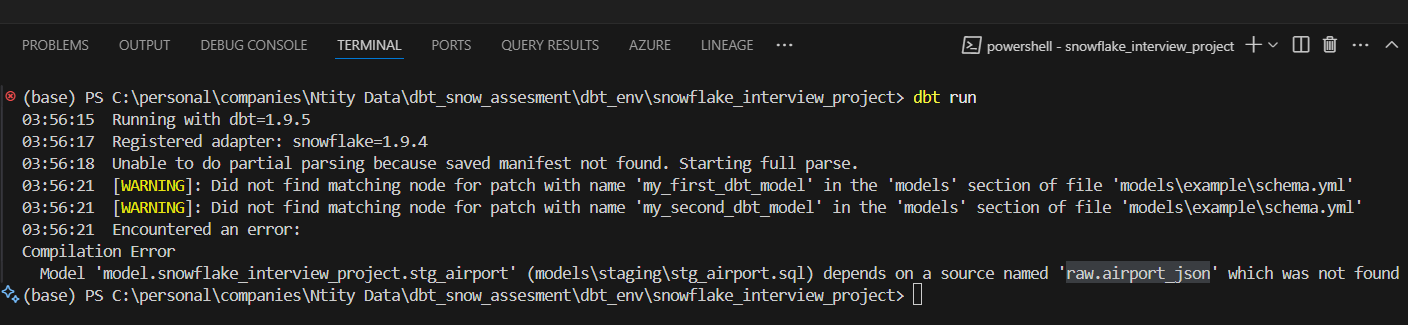


Test:

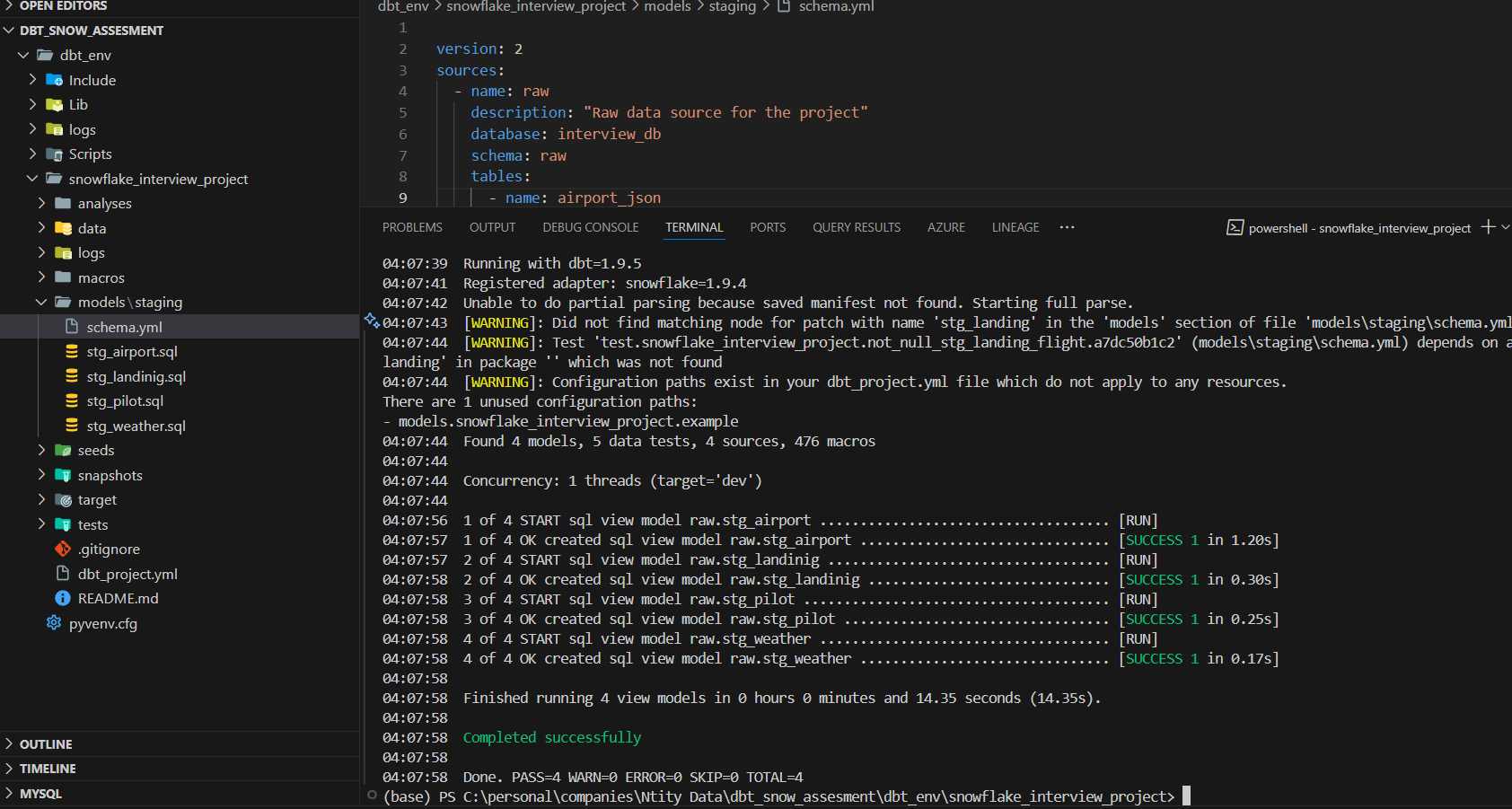


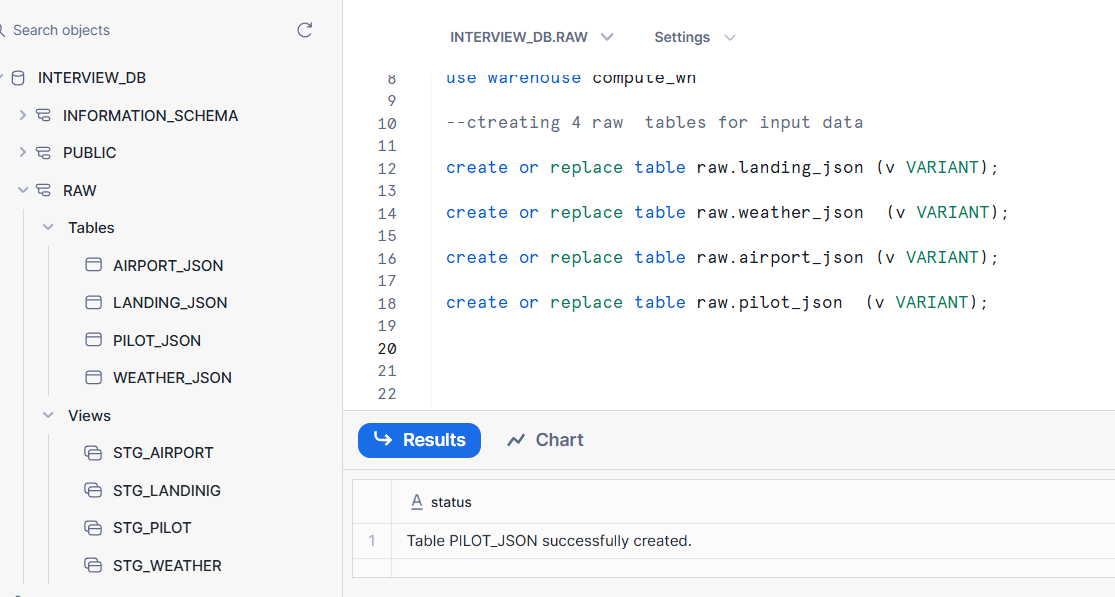
Created models for all the 4 tables and try to run with dbt run





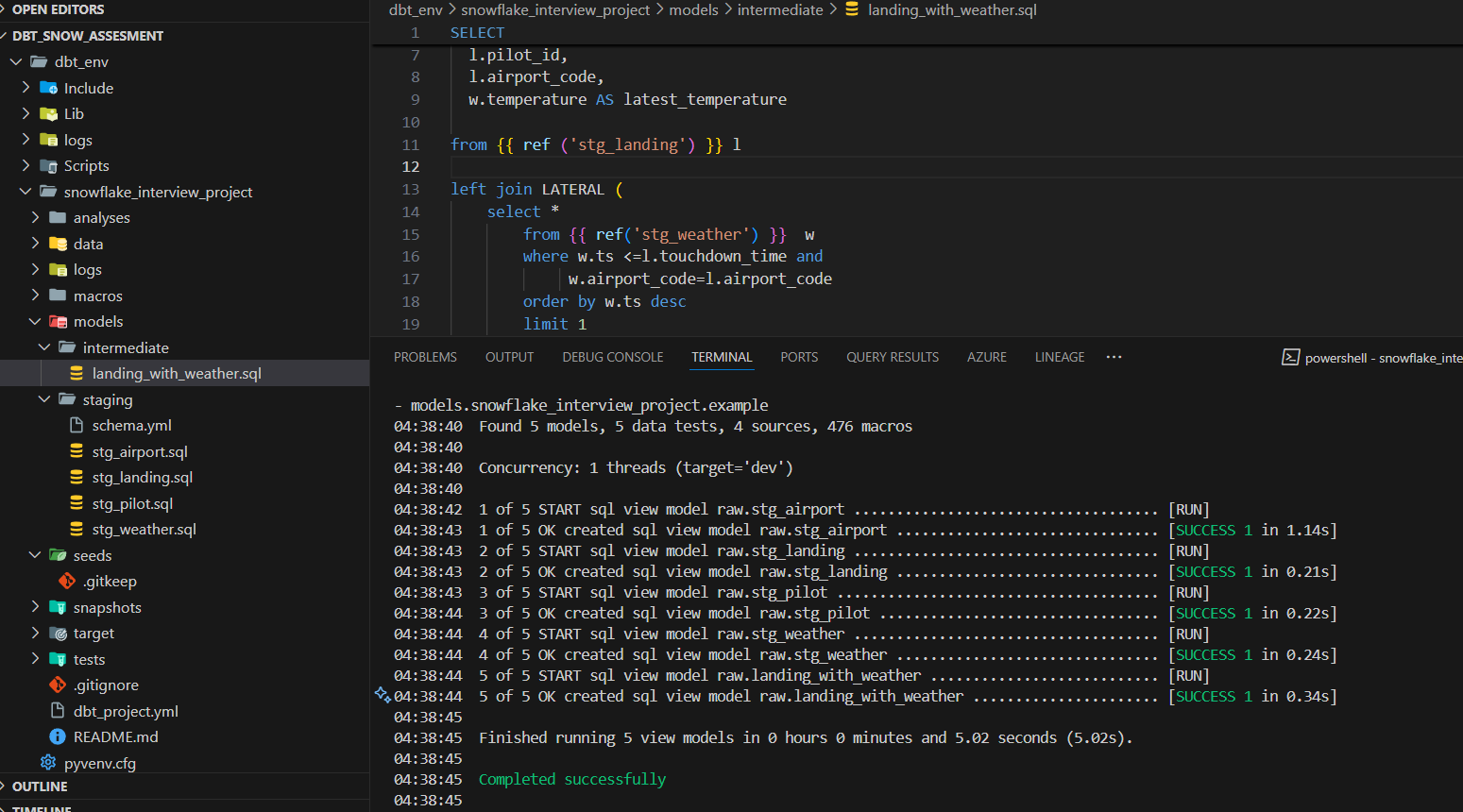
Need to add models in schema.yaml to clear this error

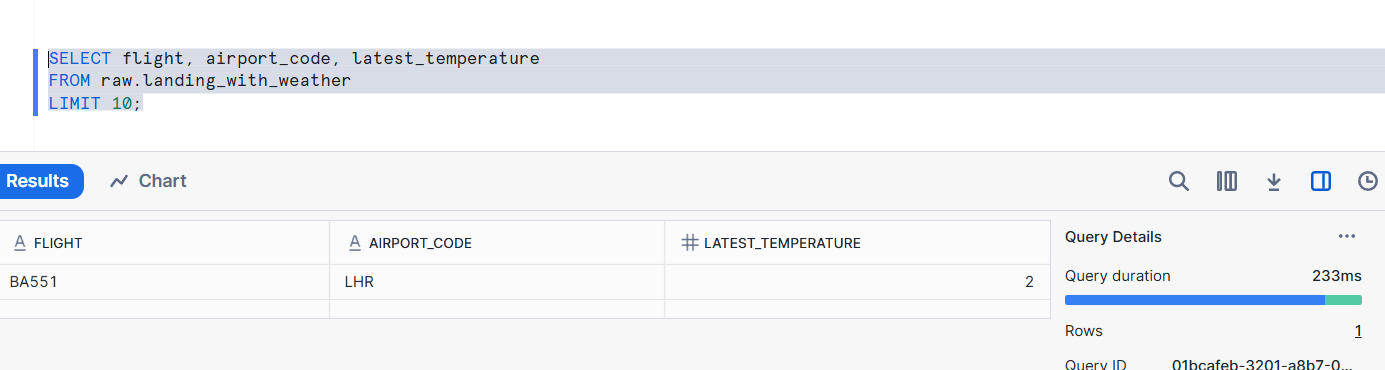




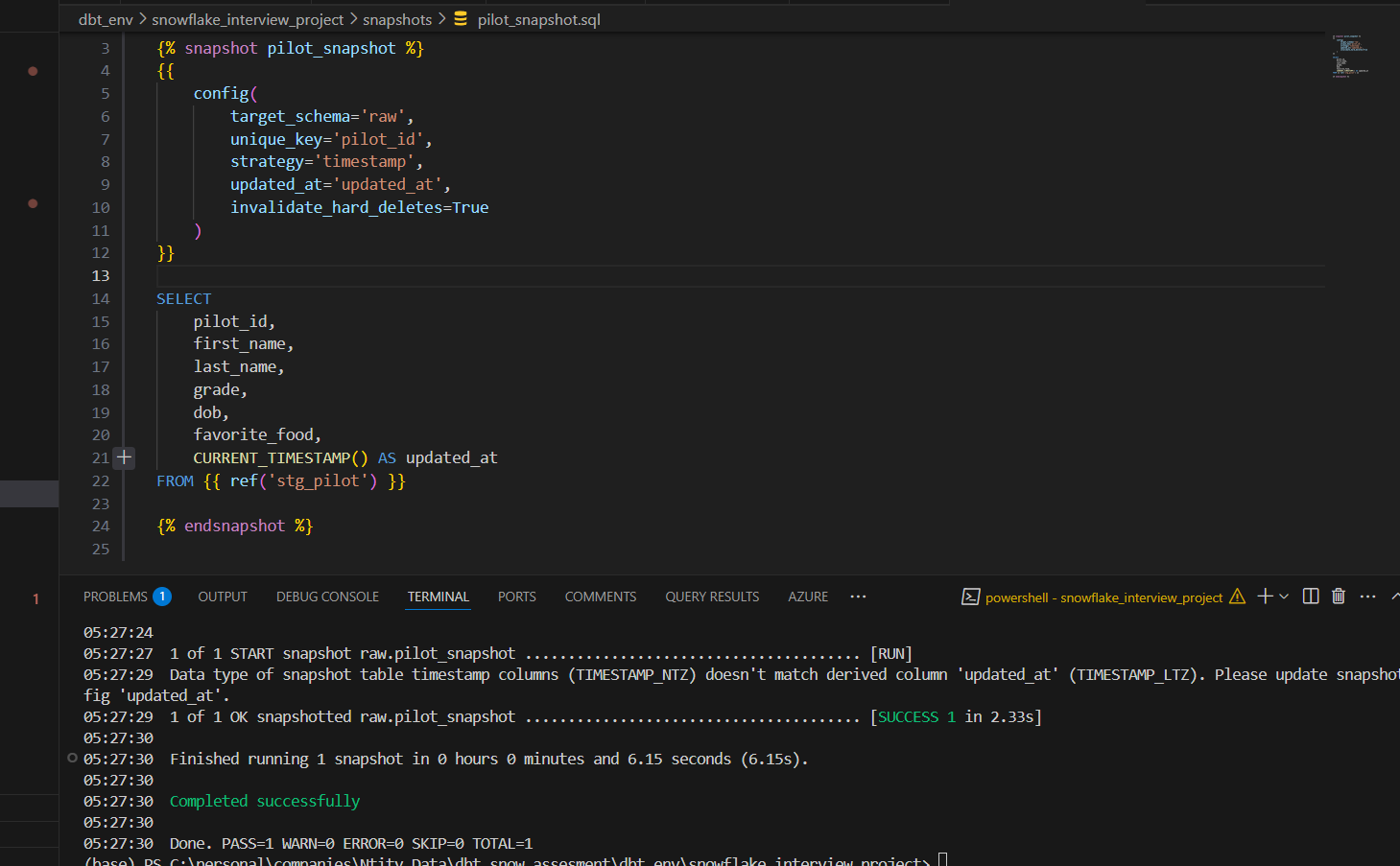
We will create an intermediate layer (models) process(transformation)

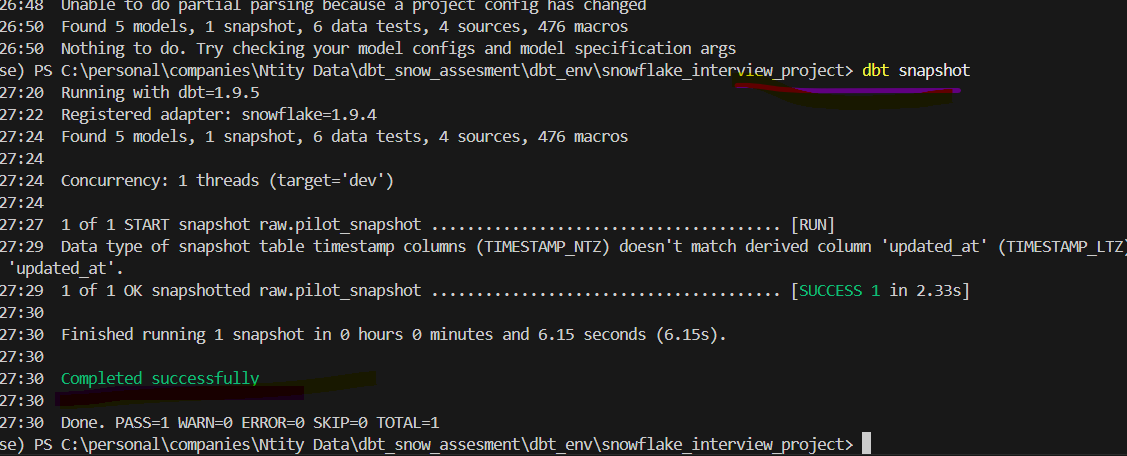
* We do join stg\_landing with the **latest weather** for that airport at/before touchdown time

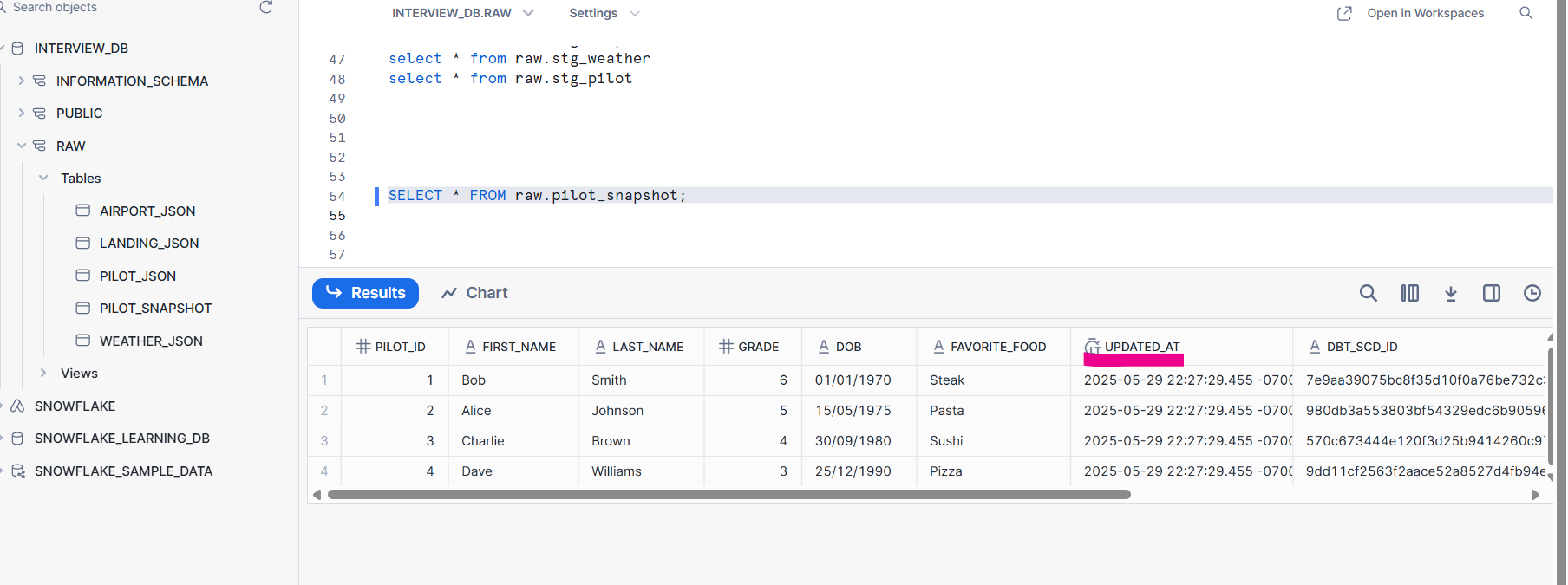


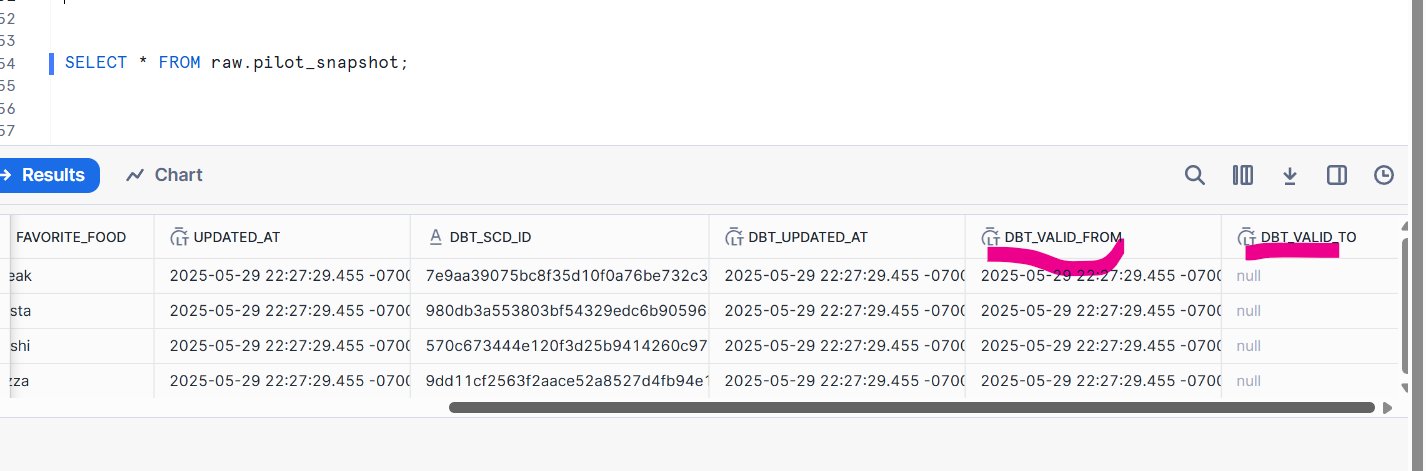


Pilot\_snapshot









models/

└── staging/

└── stg\_\*.sql

└── intermediate/

└── landing\_with\_weather.sql

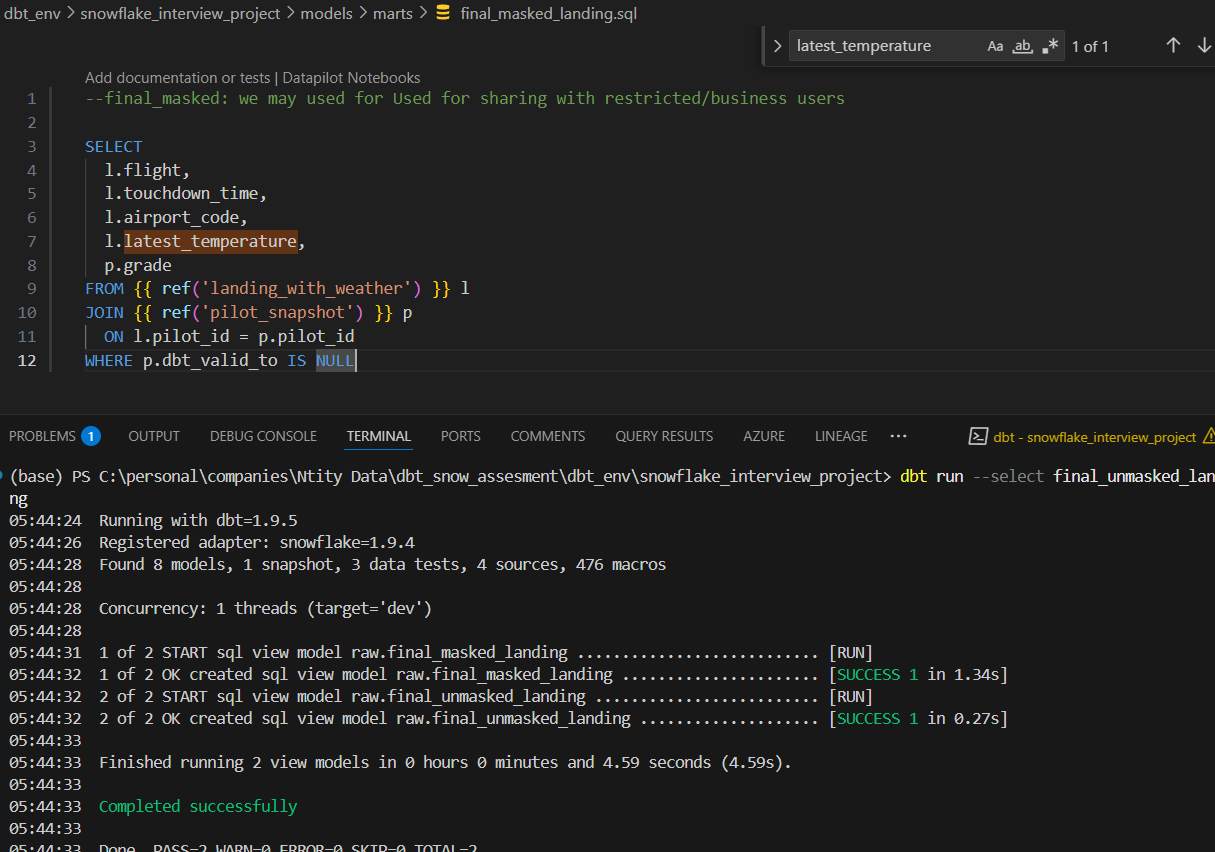
└── marts/

├── final\_unmasked\_landing.sql

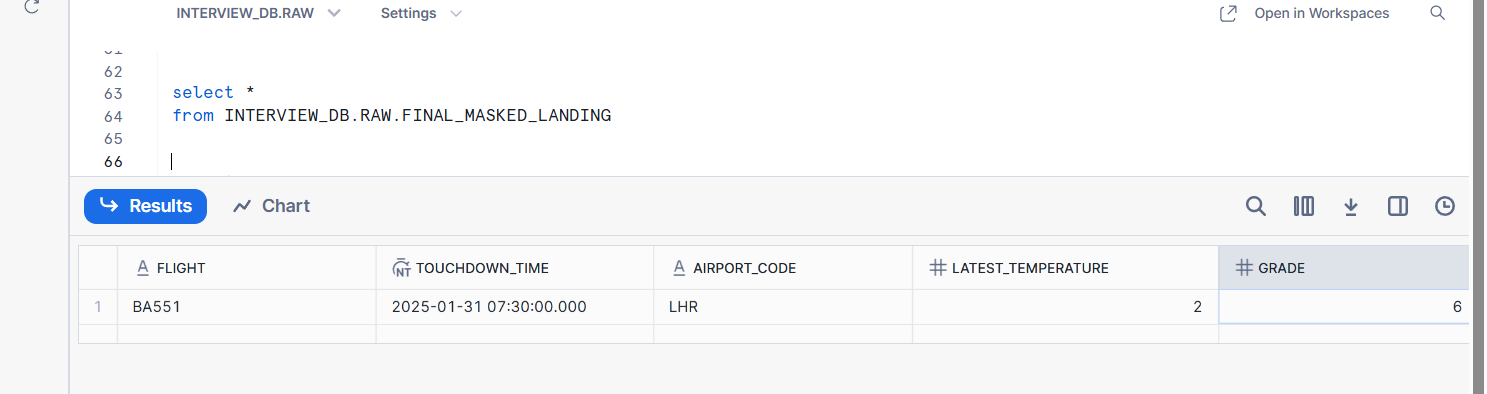
└── final\_masked\_landing.sql

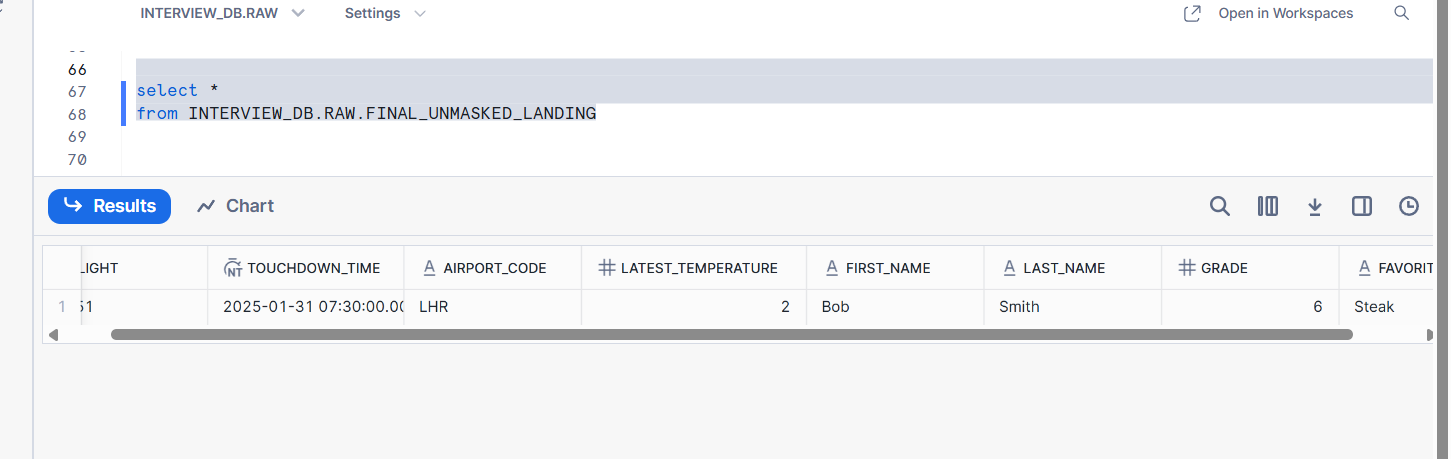
snapshots/

└── pilot\_snapshot.sql



For masked landing





Unmasked flight(include firstname ,lastname etc …sensitive data)

