

Automation

- An orchestrator (e.g., Airflow) will trigger all tasks, including dbt runs.
- dbt runs will be defined by using tags/folder structure for scheduling.
- In case of a failure during a dbt run/build, a notification will be sent to a slack/teams channel alerting the model owners with a summary of failures. The dbt run logs could be accessed through the run metadata, or in tabular format by add the dbt elementary package.

Performance Optimization

- Models would be partitioned by a date since we want to keep a historical snapshot of the calculations, and most models will run incrementally. Clustering could be added later if needed.
- For frequently ran queries, we can enable caching them. This could be done by accessing the query history on a project level and identifying those queries.