Project 6

Elements of Data Integration (CS 329E)
Second Edition

March 14, 2025

dbt setup

Person 1

VS Code

SSH tunnel

Person 2

VS Code

SSH tunnel

dbt VM

dbt project 1

dbt project 2

Compute Engine

SQL Models

SQL Models

Python Models

Python Models

Warehouse tables

BigQuery

PySpark runtime

Dataproc
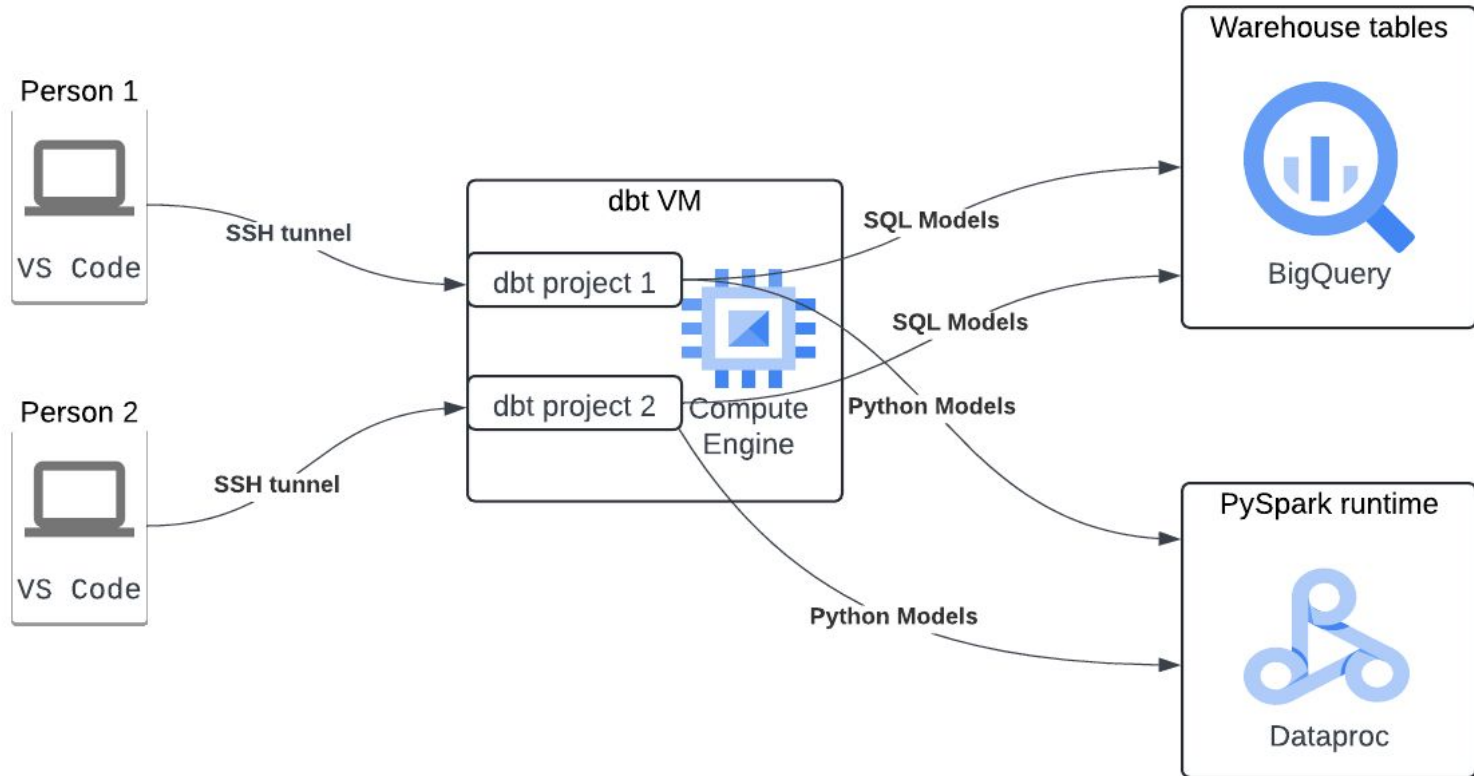
dbt setup with GitHub

Push and pull updates to/from GitHub to stay in sync with your partner