

Your task as an Analytics Engineer at Clara is to take data in its rawest form and create a model using dbt as a base for a report. This task contains Github event data for 1 hour.

Your job is to use DBT to build a report consisting on the following indicators:

Top 10 active users sorted by amount of PRs created and commits pushed

Top 10 repositories sorted by amount of commits pushed

Top 10 repositories sorted by amount of watch events

In order to do this you will access to 4 files

1. Events
  - a. ID : Id of event
  - b. Type: Type of event
  - c. Actor\_id: identifier of actor
  - d. repo\_id : identifier of repo
2. Actors
  - a. ID : Id of actor
  - b. Username: username of actor
3. Commits
  - a. Sha: sh key
  - b. Message: commit message
  - c. Event\_id : identifier of event
4. Repos
  - a. Id: id of repo
  - b. Name: name of repo

This assignment must be done using DBT.

We recommend building some intermediate tables and then the tables for the report

Please include your diagrams or thought process for data modeling as well as all the code used

We expect you to do some basic tests in DBT