

### 1. Jelaskan apa itu DBT?

DBT(Data Build Tool) adalah Open-source tool yang membantu data engineer dalam melakukan transformasi data di warehouse mereka secara efektif. DBT dirancang khusus untuk bekerja dengan data warehousing modern seperti Bigquery, Snowflake, Redshift dll. DBT memiliki beberapa fungsi diantaranya:

- Transformasi data: DBT adalah alat yang memungkinkan data engineer melakukan transformasi data, seperti menggabungkan data, membersihkan data, menormalisasi data dan agregasi data
- Modeling data: DBT memungkinkan untuk data engineer membuat model data yang terstruktur, sehingga data dapat dengan mudah mengeksekusi query SQL pada data yang telah diproses
- Orkestrasi: DBT dapat membantu data engineer mengatur transformasi dan modeling data secara otomatis, sehingga lebih mudah dalam memperbarui dan memelihara alur kerja data

### 2. Apa keuntungan menggunakan DBT?

Berikut beberapa menggunakan DBT:

- Skalabilitas: DBT dirancang untuk menangani volume data yang besar dan kompleks, dengan menggunakan DBT, data engineer dapat mengelola transformasi dan pemodelan data dengan efisien
- Orkestrasi Otomatis: DBT menyediakan kemampuan untuk mengatur dan menjalankan proses transformasi dan pemodelan data secara otomatis
- Konsistensi dan Keterbacaan Kode: DBT memungkinkan data engineer untuk mendefinisikan transformasi dan pemodelan data dalam bentuk SQL yang terstruktur
- Integrasi dengan Data Warehousing Modern: DBT dirancang untuk berintegrasi dengan data warehousing modern seperti BigQuery, Snowflake, dan Redshift

### 3. Jelaskan Dependency Tree dan Versioning pada DBT?

Dependency tree dalam DBT menggambarkan hubungan antara setiap model dan transformasi data dalam DBT. Setiap model pada DBT saling berhubungan satu sama lain sesuai dalam urutan ketika mereka dieksekusi. Misalnya, jika suatu model mengandalkan pada hasil dari transformasi data lain, maka model tersebut harus dieksekusi setelah transformasi data yang menjadi inputnya. Dependency Tree membantu pengguna untuk memahami dan mengelola ketergantungan antara berbagai bagian dari alur kerja pengolahan data mereka.

Versioning dalam DBT mengacu pada praktik menyimpan versi dari kode transformasi data dan pemodelan data dalam repositori version control, seperti Git. Ini memungkinkan pengguna untuk melacak perubahan kode seiring waktu dan mengelola versi kode dengan lebih terstruktur. Dengan menggunakan versioning, tim dapat bekerja sama dalam pengembangan dan pemeliharaan kode DBT dengan lebih efisien. Mereka dapat melacak siapa yang membuat perubahan, apa yang telah diubah, dan kapan perubahan tersebut

terjadi. Hal ini membantu dalam mencegah konflik kode, memfasilitasi kolaborasi tim, dan memastikan konsistensi dalam alur kerja pengolahan data.