

dbt

Kristo Raun

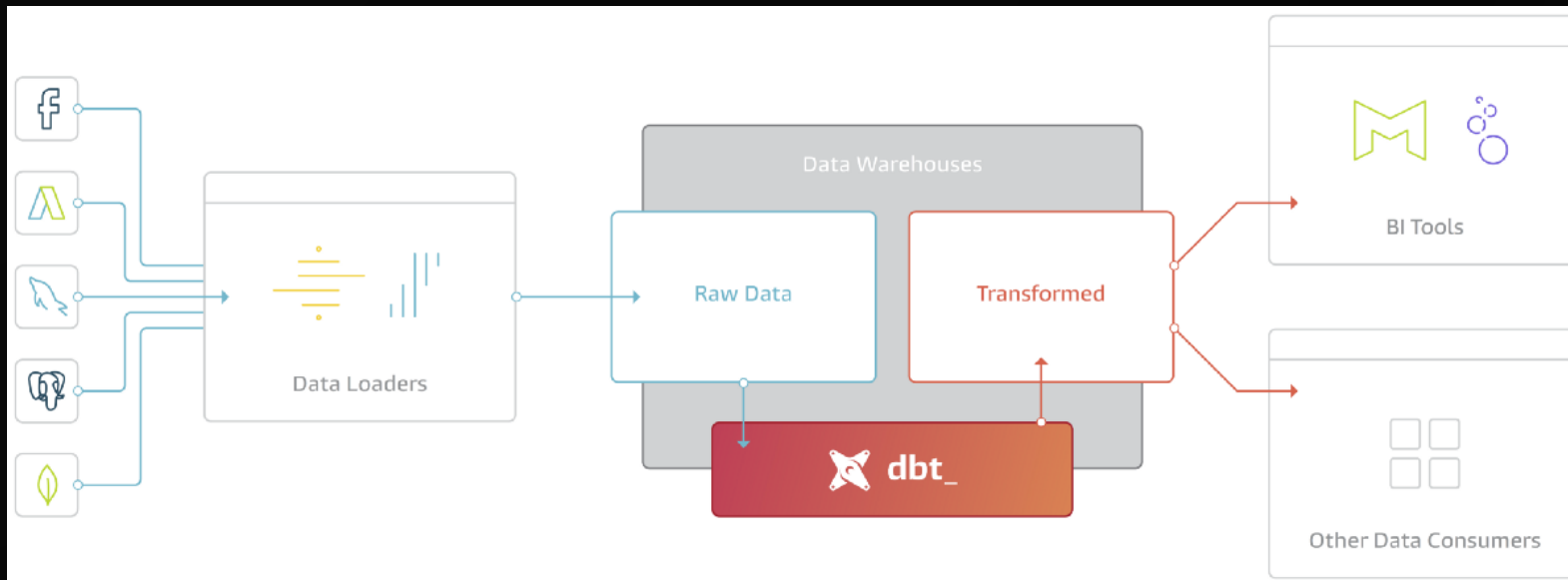
Data Engineering 2024 Fall



UNIVERSITY OF TARTU

What is dbt?

- ❖ dbt (data build tool) enables analytics engineers to transform data in their warehouses by simply writing select statements. dbt handles turning these select statements into tables and views.
- ❖ dbt Core is an open-source tool that enables data teams to transform data using analytics engineering best practices.



Why dbt?

- ◆ From ETL to ELT -> queries are executed on the DWH itself
- ◆ The number of scripts/statements will grow. Fast.
- ◆ Benefits
 - ◆ Parameters
 - ◆ Easier version control
 - ◆ Dependencies
 - ◆ In-built documentation
 - ◆ Testing
 - ◆ Snapshotting
 - ◆ ...

dbt dictionary

- ◆ Project
 - ◆ Sql and yml files
- ◆ Model
 - ◆ Sql select statement
- ◆ Profile
 - ◆ Connector to dwh
- ◆ Macro
 - ◆ Reusable function
- ◆ Test
 - ◆ Executable data validation
- ◆ Snapshot
 - ◆ Type-2 SCD (slowly changing dimension)
- ◆ Package
 - ◆ Modularized code

Final words...

- ❖ dbt is a rich and consistent way of transforming data in your DWH
- ❖ dbt is not made for extracting or pushing data to another system

