

dbt for Data Transformations

by Fariz Wakan

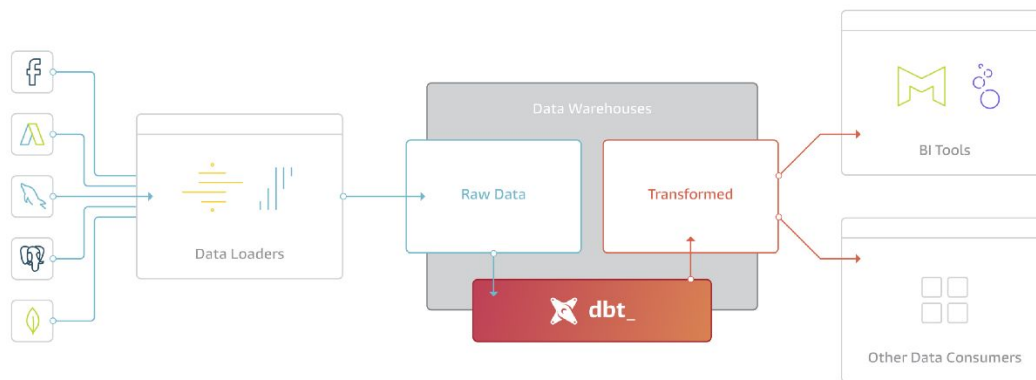
April 23th, 2024



dbt Fundamentals

What is dbt?

- dbt stands for **Data Build Tool**
- It does “**T**” in the **ELT**



The screenshot shows a web browser displaying the dbt documentation website. The URL in the address bar is `docs.getdbt.com/docs/introduction`. The page has a dark theme with a sidebar on the left containing a navigation menu. The main content area is titled "What is dbt?" and includes two paragraphs of introductory text, a link to "The dbt Viewpoint", and a section titled "dbt optimizes your workflow" with a bulleted list of features. A right sidebar contains a "CONTENTS" section with links to various dbt products and related docs, along with an "Edit this page" link.

dbt

What is dbt?

Guides

- Supported data platforms
- About dbt Cloud
- Set up dbt
- Develop with dbt Cloud
- Build dbt projects
- Deploy dbt
- Collaborate with others
- Use the dbt Semantic Layer
- dbt Cloud APIs
- Available dbt versions
- dbt support

Docs Reference v1.7

Resources Community

Search

Create a free account

What is dbt?

What is dbt?

dbt is a transformation workflow that helps you get more work done while producing higher quality results. You can use dbt to modularize and centralize your analytics code, while also providing your data team with guardrails typically found in software engineering workflows. Collaborate on data models, version them, and test and document your queries before safely deploying them to production, with monitoring and visibility.

dbt compiles and runs your analytics code against your data platform, enabling you and your team to collaborate on a single source of truth for metrics, insights, and business definitions. This single source of truth, combined with the ability to define tests for your data, reduces errors when logic changes, and alerts you when issues arise.

Read more about why we want to enable analysts to work more like software engineers in [The dbt Viewpoint](#).

dbt optimizes your workflow

- Avoid writing boilerplate [DML](#), and [DDL](#), by managing transactions, dropping tables, and managing schema changes. Write business logic with just a SQL `select` statement, or a Python DataFrame, that returns the dataset you need, and dbt takes care of [materialization](#).
- Build up reusable, or modular, data models that can be referenced in subsequent work instead of starting at the raw data

CONTENTS

- dbt optimizes your workflow
- dbt provides more reliable analysis
- dbt products
 - dbt Cloud
 - dbt Core
- The power of dbt
- Related docs

Edit this page

dbt Project

The screenshot shows the dbt documentation website. The left sidebar contains a navigation menu with the following items: 'What is dbt?', 'Guides', 'Supported data platforms', 'About dbt Cloud', 'Set up dbt', 'Develop with dbt Cloud', 'Build dbt projects' (expanded), 'About dbt projects' (selected), 'dbt tips and tricks', 'Build your DAG', 'Build your metrics', and 'Enhance your models'. The main content area is titled 'About dbt projects' and includes a breadcrumb 'Build dbt projects'. The text explains that a dbt project informs the context of the project and how to transform data, enforcing a top-level structure with files like `dbt_project.yml` and directories like `models`, `snapshots`, and `seeds`. It also mentions that a project needs at least a `dbt_project.yml` file and can include various resources. A table at the bottom lists these resources and their descriptions.

Contents:

- Project configuration
- Project subdirectories
- New projects
- Sample projects
- Related docs

[Edit this page](#)

Resource	Description
models	Each model lives in a single file and contains logic that either transforms raw data into a dataset that is ready for analytics or, more often, is an intermediate step in such a transformation.
snapshots	A way to capture the state of your mutable tables so you can refer to it later.
seeds	CSV files with static data that you can load into your data platform with dbt.
data tests	SQL queries that you can write to test the models and resources in your project.

Play with dbt BigQuery

Demo Session

References

- <https://docs.getdbt.com/docs/introduction>



Thank you!