

DBT

- diferencias entre dbt core, dbt cloud free y dbt cloud de pago
- Paso a paso de un proyecto dbt + postgre
- Qué es dbt y para qué nos sirve (demostrado con el ejemplo anterior)?

INTRODUCCIÓN

Data Build Tool(dbt) es una herramienta de análisis de datos open source que permite a los equipos de análisis y datos transformar, documentar y orquestar datos de una manera eficiente.

Se enfoca principalmente en la transformación de datos, toma datos en crudo de diferentes fuentes, los limpia y los transforma. Posteriormente los entrega en un formato estructurado para que los analistas o científicos de datos puedan trabajar con ellos.

DBT permite trabajar con una amplia variedad de sistemas de almacenamiento de datos, como bases de datos relacionales, cloud, data warehouse, etc.

DBT se escribe en SQL y permite a los equipos de análisis colaborar en el código, documentar y automatizar la transformación de datos.

TARIFAS

A continuación, se muestran las principales diferencias entre DBT core, DBT Cloud Free y DBT Cloud de pago:

1. DBT Core: Versión gratuita, puede ser instalador localmente en un ordenador o servidor. Incluye las funciones esenciales de DBT como transformación de datos, testeo y orquestación de tareas. También dispone de una comunidad activa para proporcionar soporte.

2. DBT Cloud Free: Versión gratuita de DBT Cloud, plataforma que ofrece herramientas de orquestación, programación, ejecución y monitoreo de flujos de trabajo de DBT. Al ser gratuita tiene una limitación en cuanto a la cantidad de proyectos y ejecuciones que pueden realizar. Es la ideal para probar y familiarizar.
3. DBT Cloud de pago: Versión de pago, ofrece más funcionalidades en comparación a la gratuita y más capacidad de procesamiento de datos. Incluye funciones avanzadas como la integración con otras herramientas de análisis de datos, soporte a múltiples data warehouse, programación avanzada de flujos y el monitoreo y análisis en tiempo real. Es ideal para usuarios que necesitan trabajar con grandes volúmenes de datos y equipos de análisis de datos más grandes y complejos.

AIRBYTE

Plataforma de integración de datos open source, que permite a las empresas mover datos de diversas fuentes a una amplia variedad de destinos de datos en tiempo real. Se enfoca en que la integración de datos sea fácil, rápida y confiable.

La instalación de Airbyte se puede realizar ya sea en máquinas locales, virtuales como en Docker o Kubernetes, etc. Airbyte ofrece instrucciones detalladas para mucha variedad de instalaciones.

A diferencia de otras herramientas de integración de datos, Airbyte ofrece:

- Código abierto
- Facilidad de uso
- Conectores preconstruidos
- Escalabilidad
- Comunidad

La diferencias de los planes esta