

سرفصل دوره‌ی آموزشی

مدل‌سازی داده با dbt

dbt for Data Engineers

آموزش مهندسی تحلیل با dbt در محیط محلی؛ شامل نصب dbt Core، راه‌اندازی انبار داده‌ی Postgres در Docker و تبدیل داده از لایه‌ی Staging تا آزمون و ساخت Mart.

سطح دوره	مدت پیشنهادی	بخش و درس	شیوه‌ی برگزاری
متوسط تا پیشرفته	۴ روز (۳۲ ساعت)	۱۲ بخش، ۵۶ درس	حضوری یا آنلاین

— آنچه شرکت‌کنندگان می‌آموزند

- ◆ نصب و اجرای dbt Core روی یک انبار داده‌ی واقعی
Postgres در Docker
- ◆ ساخت مدل‌های Staging، میانی و Mart که با ref() و source() به هم پیوند خورده‌اند
- ◆ آزمون و مستندسازی داده با آزمون‌های عمومی،
اختصاصی و واحد dbt
- ◆ ساخت مدل‌های افزایشی و Snapshot و زمان‌بندی
Airflow با dbt

پیش‌نیازها

- نیازی به تجربه‌ی قبلی dbt نیست؛ دوره از اولین مدل شروع می‌شود
- تسلط بر SQL (SELECT، JOIN و GROUP BY)
- آشنایی با خط فرمان
- Docker Desktop (نصب آن در بخش ۱ آموزش داده می‌شود)

ابزارها و فناوری‌ها

Postgres Docker dbt Core
Airflow Python

پروژه‌ی عملی

یک DAG کامل، آزموده و مستند در dbt برای
تبدیل داده‌های سفر CityFlow روی
Postgres و DuckDB، با استقرار و زمان‌بندی
در Airflow

۵ درس	۱ آشنایی با مهندسی تحلیل
۱. مهندسی تحلیل در واقع چیست	۴. اولین مدل
۲. dbt Core، dbt Cloud و گذار به Fusion	۵. پروژه‌ی راهنمادار: بارگذاری داده‌های سفر
۳. راه‌اندازی انبار داده	CityFlow در انبار داده
۵ درس	۲ مدل‌ها، DAG و Materialization
۱. ref() و source(): ساخت گراف وابستگی	۴. انبار داده‌ی دوم با همان SQL
۲. انواع Materialization: View، Table و Ephemeral	۵. پروژه‌ی راهنمادار: بازسازی اولین مدل CityFlow
۳. DAG، فرمان dbt build و اجرای انتخابی	به‌صورت یک DAG واقعی
۵ درس	۳ منابع داده و مدل‌های Staging
۱. تعریف منابع داده	۴. Staging داده‌های سفر CityFlow
۲. تازگی داده‌های منبع	۵. پروژه‌ی راهنمادار: یک لایه‌ی Staging کامل
۳. لایه‌ی Staging در عمل	
۵ درس	۴ آزمون داده
۱. چرا یک کوئری SQL را آزمون کنیم؟	۴. dbt_utils و مدیریت بسته‌ها
۲. چهار آزمون عمومی داخلی	۵. پروژه‌ی راهنمادار: لایه‌ی Staging آزموده
۳. آزمون‌های اختصاصی	
۵ درس	۵ مدل‌های میانی و Martها
۱. لایه‌ی میانی	۴. ساخت جدول واقعیت سفرهای CityFlow
۲. مدل‌سازی جدول‌های واقعیت و بُعد در dbt	۵. پروژه‌ی راهنمادار: گزارش منطقه-ساعت در dbt
۳. ساخت بُعد منطقه‌ی CityFlow	
۵ درس	۶ Jinja و ماکروها
۱. Jinja در SQL	۴. ماکروهای dbt_utils
۲. کنترل جریان با For و If	۵. پروژه‌ی راهنمادار: حذف تکرار از مدل‌های CityFlow
۳. نوشتن ماکروی اختصاصی	

۵ درس

۷ مدل‌های افزایشی

۱. چرا بازسازی کامل در مقیاس بزرگ جواب نمی‌دهد
۲. Materialization افزایشی
۳. راهبردهای افزایشی
۴. unique_key و داده‌های دیررس
۵. پروژه‌ی راهنمادار: جدول واقعیت سفرهای CityFlow به صورت افزایشی

۵ درس

۸ Snapshot و ابعاد با تغییر کند

۱. مسئله‌ای که Snapshot حل می‌کند
۲. Snapshot در dbt با نگارش جدید
۳. راهبرد زمان‌محور در برابر بررسی‌محور
۴. کوئری روی وضعیت تاریخی
۵. پروژه‌ی راهنمادار: Snapshot از بُعد منطقه‌ی CityFlow

۵ درس

۹ مستندسازی و سایت مستندات dbt

۱. مستندسازی مدل‌ها و ستون‌ها
۲. فرمان‌های dbt docs generate و dbt docs serve
۳. خواندن گراف تبارنامه‌ی داده
۴. Exposureها
۵. پروژه‌ی راهنمادار: پوشش کامل مستندات

۵ درس

۱۰ آزمون واحد و آزمون پیشرفته

۱. آزمون داده در برابر آزمون واحد
۲. نوشتن آزمون واحد در dbt
۳. آزمون حالت‌های مرزی بدون داده‌ی واقعی
۴. ترتیب کامل اجرای dbt build
۵. پروژه‌ی راهنمادار: آزمون واحد یک ماکروی واقعی CityFlow

۵ درس

۱۱ استقرار، CI و الگوهای عملیاتی

۱. محیط‌ها و Targetها
۲. انتخاب مبتنی بر وضعیت
۳. CI در dbt
۴. زمان‌بندی dbt با Airflow
۵. پروژه‌ی راهنمادار: به‌روزرسانی زمان‌بندی‌شده‌ی CityFlow

۱ درس

۱۲ پروژه‌ی پایانی

۱. پروژه‌ی پایانی: انبار داده‌ی تحلیلی شرکت CityFlow

درخواست برگزاری اختصاصی این دوره برای سازمان شما

ثبت درخواست این دوره

lotfinejad.ir/contact

سرفصل، مدت، سطح و تمرین‌های دوره بر اساس نیاز تیم شما تنظیم می‌شود و آموزش به زبان فارسی، به صورت حضوری یا آنلاین برگزار می‌شود. برای ثبت درخواست، فرم مربوط به همین دوره را تکمیل کنید.