

Data Engineer – Data Fellowship 12 IYKRA

Assignment 3

Alya Mutiara Firdausyi

Instruction

Task

Hey, data rangers! We got a data project from the business team. They need us to analyze NYC Taxi Trip data. Based on the requirements they want to understand the below insight, (at least)

- Monthly total passengers
- Monthly transactions per payment type
- Monthly trip distance per rate code

Based on organization architecture, you have to transform data using dbt. The raw data should be under `raw.nyc_taxi_trip` table. Generate multiple models based on needs. Based on governance, 60% of columns must have tests on each model in the dbt project. The description of the table and columns is mandatory.

Note:

- Custom at least 1 macros
- Custom at least 1 test

Output

GitHub or GitLab repository link (containing a dbt project).

Challenge

Ubah currency menjadi IDR: <https://www.investing.com/currencies/usd-idr-historical-data>

Goal:

1. **Transform NYC Taxi Trip Data:** Use dbt, a data transformation tool, to clean, manipulate, and aggregate data from the raw `nyc_taxi_trip` table.
2. **Generate Business Insights:** Based on the transformed data, answer specific questions requested by the business team. These questions involve understanding:
 - a. Monthly total number of passengers
 - b. Monthly breakdown of transactions by payment type
 - c. Monthly average trip distance grouped by rate code

GitHub Repository:

Please see the repository below for my dbt project result for this week's assignment.

https://github.com/alyamutiara/dbt-transformation-nyx-taxi-data/blob/main/nyc_taxi_data/models/sources/nycTaxi.yml