

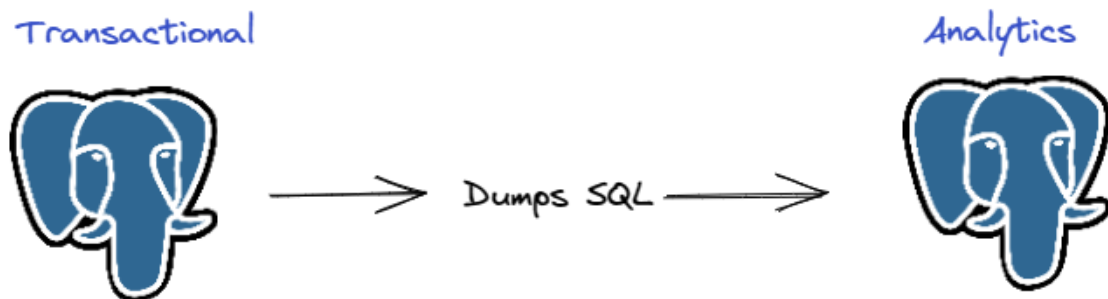
DESAFIO

Luis Felipe Almeida Nogueira

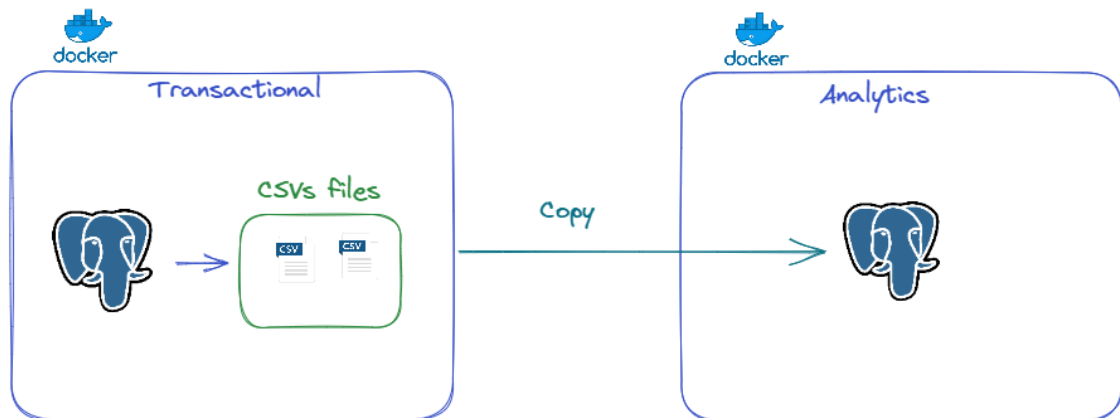
refera

An orange curved line, resembling a stylized smile or a checkmark, positioned below the word "refera".

Minha ideia inicial era criar uma classe chamada Postgres com funções para manipular os bancos de dados, executar essas funções em um código main e gerar dumps da tabela transactional para rodá-los na analytics. Nesse modelo:



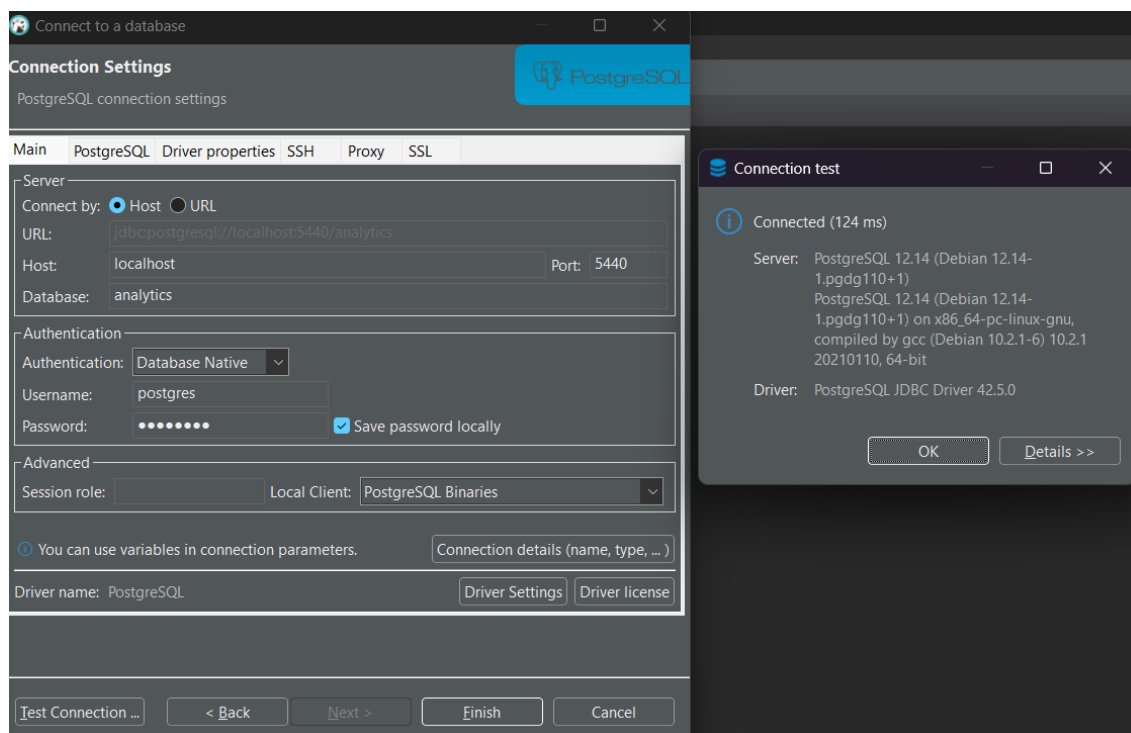
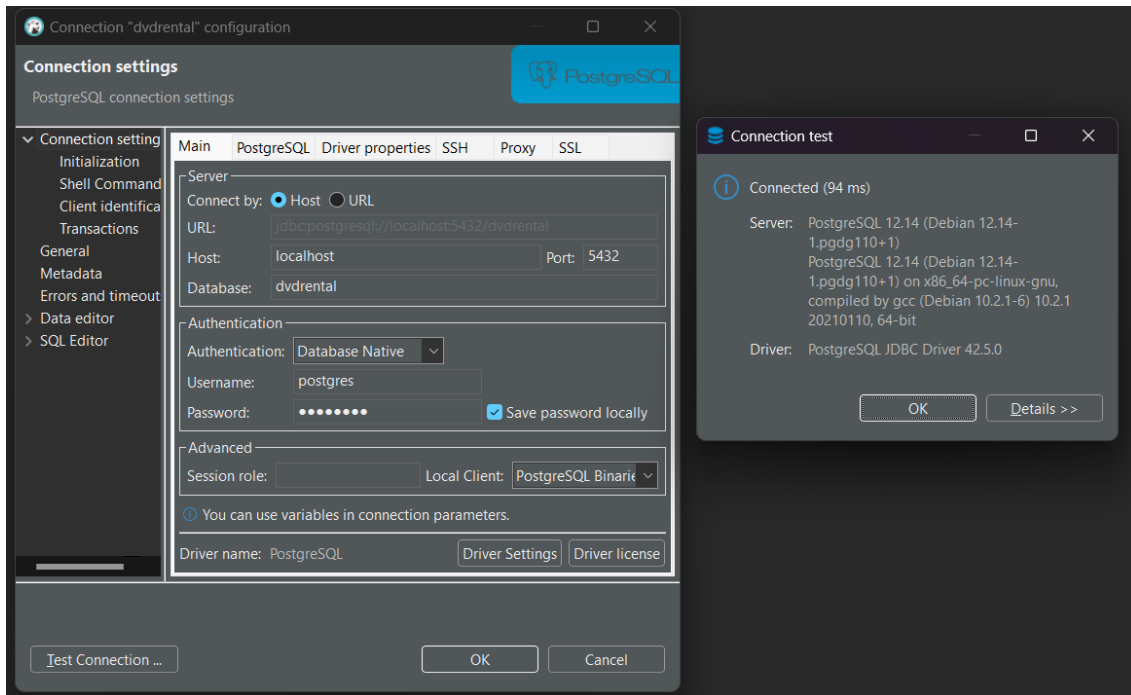
Porém o banco transactional tem muitos triggers, functions etc. e acabei mudando de estratégia: gerar CSVs e copiá-los no banco analytics. A insistência em realizar COPY ao invés de vários inserts é por ser mais performático e gerar menos conexões com o banco.



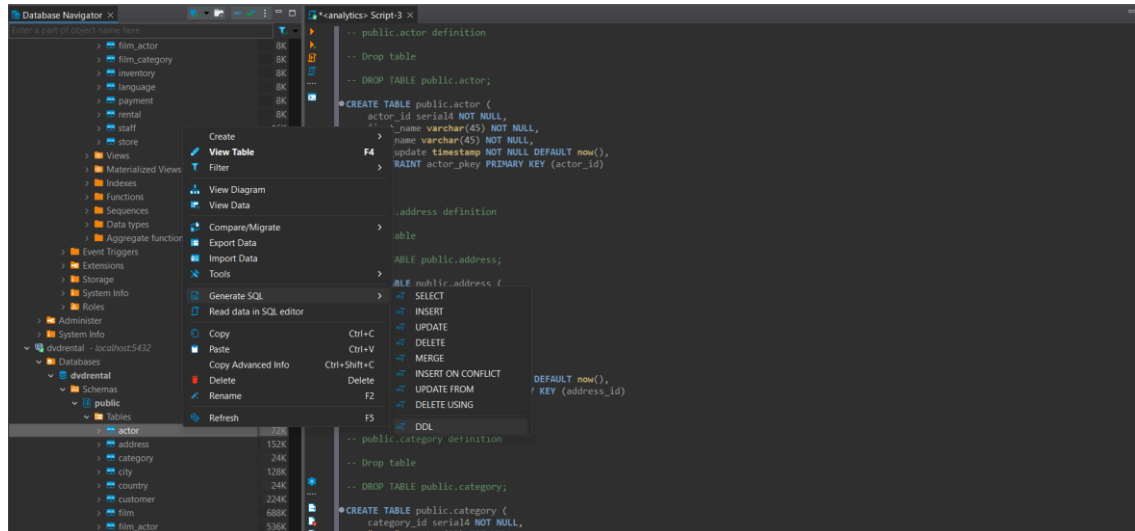
Subi os containers com **docker-compose -d** com uma pequena mudança: criei uma pasta code e criei um bind-mount direcionando para esta pasta. Criei para poder atender o diferencial de implementação e rodar o código no Docker.

```
volumes:
  - postgres:/var/lib/postgresql/data
  - ./data/dvdrental.sql:/docker-entrypoint-initdb.d/dvdrental.sql
  - ./code:/code
```

Me conectei aos bancos com a ferramenta DBeaver para uma visualização melhor:



Para criar as tabelas eu gerei os DDLs das tabelas do **transactional** com o próprio dbeaver. Apesar de entender que uma maneira mais apropriada seria fazer via código, achei que seria mais prático para apenas essa solução.



Após criar as tabelas no banco analytics como mostrado acima, fui para o código Python. Criei um arquivo postgres.py e criei uma classe Postgres. Trabalho com uma classe parecida com Redshift e segui a mesma linha. E então implementei as funções no main.py.

Para executar o código, basta criar os containers com **docker-container -d**, acessar o container e instalar o **Python**, **Pip** e as dependências, pois a imagem não possui Python. Códigos abaixo:

```
apt-get update
apt-get install python3
apt-get install python3-pip python-dev
cd code
mkdir sql_dumps
pip install -r requirements.txt
pip install pycpg2-binary # A instalação dessa lib precisa ser forçada
python3 main.py
```

Gostaria de deixar duas observações: testando algumas vezes houve 4 erros de tipagem de dados e os códigos/comentários estão em português. Não olhei mais a fundo quanto aos erros que aconteceram algumas vezes, mas eu faria a inserção manual deles. Numa escala maior, levantaria quais são esses erros e os corrigiria no banco de origem.

Quanto aos códigos e comentários em português, deixei por acreditar que facilita a interpretação do desafio, mas entendo que em códigos de produção deve-se escrever em inglês.