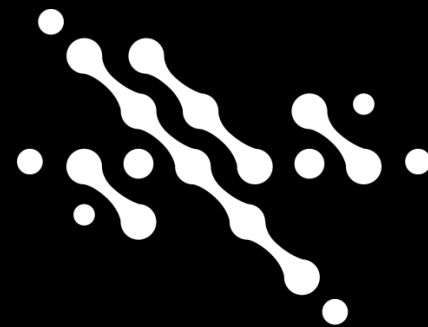




# **IGMAtech**



**CASE TÉCNICO – ENGENHEIRO DE DADOS**

# ESCOPO

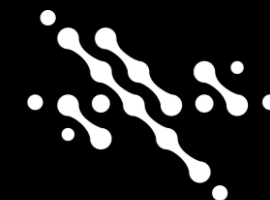
## CONTEXTO:

Você faz parte de um time de tecnologia, atuando em um projeto de estruturação das informações gerenciais de uma empresa de transporte por aplicativo. O objetivo desse projeto é estruturar as informações em um *Data Warehouse* (DW), desenvolver modelos preditivos em *Machine Learning* e disponibilizar resultados das metas e insights em relatórios e painéis Power BI para os gerentes e diretores da empresa. Como Engenheiro de Dados, sua função é entender as regras de negócio, construir *pipelines* de tratamento dos dados e disponibilizá-los para a equipe de Consultores em Gestão e Cientistas de Dados.

## O DESAFIO:

A equipe de Consultores em Gestão e Cientistas de Dados precisará fazer uma análise diagnóstica das viagens realizadas por cada taxi cadastrado em seu aplicativo, permitindo identificar oportunidades e ações ver-e-agir. Para isso, será necessário estruturar os dados em um esquema estrela (*star schema*) para ser carregado no Power BI. A equipe de TI do cliente enviou arquivo contendo todas as viagens do mês de Janeiro/2025.

**Você deverá construir o pipeline de ETL em Python, para estruturar os dados em esquema estrela. O pipeline será rodado mensalmente, gerando tabelas fato e tabelas dimensão, armazenadas na camada *silver* de um *datalak*.**



## O QUE DEVERÁ SER ENTREGUE:

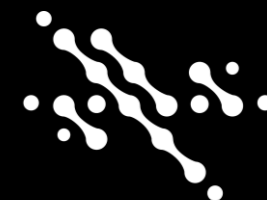
- Script Python com o pipeline de tratamento e exportação das tabelas fato e dimensão em arquivos parquet.
- Script SQL para criação das tabelas fato e dimensão (CREATE TABLE).
- Documentação contendo o dicionário de tabelas e campos gerados.

## O QUE AVALIAREMOS:

- Capacidade em identificar eventuais inconsistências nos dados.
- Clareza e objetividade na lógica do pipeline de dados.
- Coerência na modelagem *star schema* e scriptsSQL.
- Clareza na documentação das tabelas e campos.

## PRAZO PARA DEVOLUÇÃO

O prazo para resolução é de **5 dias** a partir do recebimento do documento.



**IGMAtech**

