

# Pipeline de Dados

---

- Ler dados de vendas Desnormalizados da NovaDrive Motors
- Salvar em disco arquivos parquet em padrão “DataLake”:
  - \vendas\estado=rj\ano=2025\mes=10
- Idepotente: Novos dados, novo arquivo
  - Sem dados novos: não cria nada.
  - Com dados novos: cria um novo arquivo (append de arquivos) na partição estado=UF/ano=YYYY/mes=MM/.
- Pasta raiz: /opt/airflow/data
- Watermark: arquivo json que grava ultima atualização

