

Documentação

Processo de ETL



# Processo de ETL



- GitHub
- Repositório com 6 bases
- Informações de clients, vendas, entregas, cidades, vendedores, review\_score



- Python
- Processo de extração dos arquivos do Github
- Processo de subida dos arquivos para GCP(Storage)



- GCP
- As bases ficam salvas no storage account do GCP



- Python
- Processo de extração dos dados do GCP
- Job para fazer o join de todas as bases
- Job de tratamento dos dados
- Desenvolvimento das análises e respostas do problema a partir da base gerada

# Fluxo de construção da base

