



Projeto Final

Pipeline de Cotações Cambiais com Python + LLM

Python Programming for Data Engineers

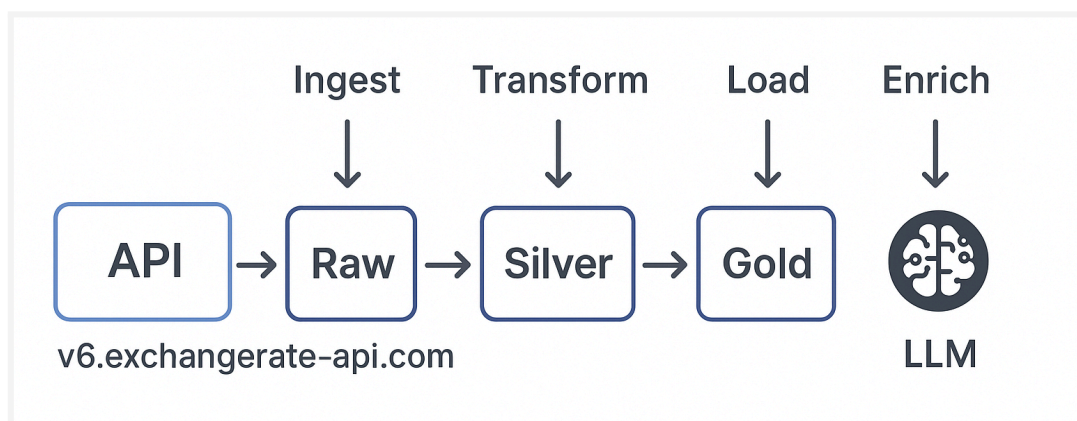
MBA em Data Engineering

Professor: Eduardo Miranda

Definição e objetivos do projeto

Este projeto implementa um pipeline de dados para coleta, processamento e análise de cotações cambiais, utilizando Python. O fluxo segue o conceito de camadas (raw, staging, gold) e inclui a integração com uma LLM (Large Language Model) para gerar insights automáticos e interativos e enviados por email. A entrega Final inclui a parte técnica (Código, dados e automação) como também a artefatos executivos (Documentação e apresentação) que irão auxiliar no entendimento do projeto como um todo.

Fluxo do Projeto



Estrutura do Projeto:

- .env # Variaveis de ambiente
- requirements.txt # Dependencias
- gold # Pasta com arquivos parquet particionado pelo nome do Cambio e data
- silver # Camada intermediária de dados, que armazena os dados normalizados em CSV
- raw # Pasta de dados bruta
- src #Pasta com codigos fonte
 - Ingestor.py # ETL De Injestao de dados
 - Llm.py # ETL de geração de Insights
 - Load.py # ETL de Carga de Dados
 - SendEmail.py # Envio de email
 - Transform.py # ETL de Transformação os dados brutos
 - [main.py](#) # Execução principal

Implementação Técnicas de Ingestão de dados

- Consumo da API via requests da lista de cambios listados no .ENV
- Salvamento da resposta em JSON em “raw/” com nome padronizado em cambio_YYYY-MM-DD-DD.
- Logging estruturado (INFO/ERRO) da operação de requisição
- Garantir qualidade (nenhuma taxa negativa ou nula).
- Armazenar em /silver/.

Implementação Técnicas de Transformação de dados

- Normalização das colunas e tipos
- Normalização com adição das colunas obrigatórias (moeda, taxa, base_currency, timestamp)
- Adição das colunas dos valores das outras moedas para fins de análise
- Remoção de dados com valor nulo ou inferior a 0.
- Logging dos erros e passos
- Persistência dos dados em um unico arquivo CSV em “silver/”

Implementação Técnicas de Insights e Envio de email

- Logging dos erros e passos
- Leitura dos arquivos Parquet
- Garante que dados de mesmo nome de câmbio e data não sejam duplicados.
- Geração de prompt pré-definido para relatório
- Uso de servidor SMTP para disparo de email com o relatório gerado pelo Gemini

Resumo financeiro diário

Resumo Executivo

A análise dos dados de câmbio entre 27 e 30 de setembro de 2025 revela um período de flutuações cambiais geralmente contidas, com a maioria das variações percentuais abaixo de 1%.

Insights Relevantes:

- **Iene Japonês (JPY):** Destacou-se como a moeda de maior valorização. Fortaleceu-se notavelmente contra várias moedas, incluindo:
 - +0.67% contra o Dólar de Hong Kong (HKD) e a Rúpia Indiana (INR).
 - +0.61% contra o Dólar Americano (USD).
 - +0.51% contra o Franco Suíço (CHF).
 - +0.48% contra o Dólar Canadense (CAD).
- **Rúpia Indiana (INR):** Apresentou a maior desvalorização geral. As quedas mais significativas foram:
 - -0.66% contra o Iene Japonês (JPY).
 - -0.51% contra o Real Brasileiro (BRL).
 - -0.47% contra o Dólar Australiano (AUD).
- **Dólar de Hong Kong (HKD):** Também demonstrou enfraquecimento generalizado, em especial:
 - -0.67% contra o Iene Japonês (JPY).
 - -0.54% contra o Real Brasileiro (BRL).
 - -0.51% contra o Dólar Australiano (AUD).
- **Dólar Americano (USD):** Mostrou uma tendência de leve enfraquecimento geral, particularmente contra:
 - -0.62% contra o Iene Japonês (JPY).
 - -0.51% contra o Real Brasileiro (BRL).
 - -0.50% contra o Dólar Australiano (AUD).
- **Real Brasileiro (BRL):** Fortaleceu-se notavelmente, especialmente contra:
 - +0.60% contra o Dólar Americano (USD).
 - +0.58% contra a Rúpia Indiana (INR).
 - +0.54% contra o Dólar de Hong Kong (HKD).

Integrantes:

Caio Macedo Santos - 2502444
Gabriel Gonçalves Cruvinel - 2502454