

# **Documentação: Integração de APIs e Geração de Banco de Dados**

**Autor: Hugo Lima**

**Data: 06/09/2024**

**Projeto: Integração e tratamento de dados usando APIs públicas do Brasil**

***CODERHOUSE***

## 1. Introdução

Este código tem como objetivo coletar dados de três APIs públicas brasileiras, realizar o tratamento necessário, e consolidar os dados em um banco de dados SQLite chamado **projeto\_final**. O código é voltado para analistas de dados que buscam estudar informações sobre estados brasileiros, participantes do sistema Pix, e corretoras de valores.

### Saída:

A saída do código é um banco de dados SQLite contendo três tabelas com dados tratados, cada uma referente a uma API consultada:

- **df\_estados\_agg**: Informações agregadas dos estados brasileiros (ID, sigla, nome e região).
- **df\_pix**: Transações do sistema Pix.
- **df\_corretora**: Informações sobre corretoras registradas na CVM.

### Nível de Privacidade:

Os dados utilizados são públicos e não contêm informações sensíveis ou pessoais.

## 2. Pré-requisitos

### Ambientes:

- Python 3.9+

### Bibliotecas:

- `plyer==2.1.0`
- `requests==2.31.0`
- `pandas==2.1.1`
- `SQLAlchemy==2.0.20`
- `sqlite3` (módulo embutido no Python)

### Arquivos:

- O projeto utiliza um arquivo `requirements.txt` com as bibliotecas listadas acima, gerado com `pip freeze`.

## 3. APIs Utilizadas

- **API Brasil - Estados (IBGE):**  
URL: <https://brasilapi.com.br/api/ibge/uf/v1>  
Dados sobre os estados brasileiros, como ID, sigla, nome e região.

- **API Brasil - Participantes Pix:**

URL: <https://brasilapi.com.br/api/pix/v1/participants>  
Informações sobre os participantes do sistema Pix, como nome, tipo de participação, e data de início de operação.

- **API Brasil - Corretoras (CVM):**

URL: <https://brasilapi.com.br/api/cvm/corretoras/v1>  
Dados sobre corretoras registradas na CVM, como nome, CNPJ, e patrimônio líquido.

---

## 4. Funções Criadas

- **fetch\_data(url):**

- **Parâmetros:** URL da API.
- **Retorno:** Dados JSON da API.
- **Descrição:** Faz uma requisição HTTP GET para a URL fornecida e retorna os dados no formato JSON.

- **process\_dates(df, columns):**

- **Parâmetros:** DataFrame e lista de colunas de data a serem tratadas.
  - **Retorno:** DataFrame com as colunas de data tratadas (separação de data e hora).
  - **Descrição:** Converte campos de data no formato ISO, separando as informações de data e hora.
- 

## 5. Tratamentos Aplicados

- **Limpeza de Dados:**

- Os campos de data das tabelas df\_pix e df\_corretora foram tratados para separar a data da hora, eliminando a coluna original com o timestamp.
- Na tabela df\_pix, a coluna inicio\_operacao foi separada em duas novas colunas: data\_inicio\_operacao (data) e hora\_inicio\_operacao (hora).

- **Conversão de valores:**

- No df\_corretora, a coluna valor\_patrimonio\_liquido foi convertida para um formato numérico com o símbolo \$ para representar o valor em dólares.
- 

## 6. Método de Saída

**Formato:**

Os dados tratados foram armazenados em um banco de dados SQLite chamado `projeto_final.db`, contendo três tabelas: `df_estados_agg`, `df_pix`, e `df_corretora`.

---

## 7. Exemplo de Consulta

**Consulta:**

sql

```
SELECT * FROM df_pix LIMIT 5;
```

**Resultado Esperado:**

Uma tabela com as primeiras cinco linhas dos participantes do sistema Pix, contendo informações como nome, tipo de participação, data de início de operação, etc.

---

## 8. Versionamento

`requirements.txt`:

- `plyer==2.1.0`
  - `requests==2.31.0`
  - `pandas==2.1.1`
  - `SQLAlchemy==2.0.20`
  - `sqlite3==*`
- 

## 9. Referências

- API Brasil - Estados (IBGE): <https://brasilapi.com.br/api/ibge/uf/v1>
- API Brasil - Participantes Pix: <https://brasilapi.com.br/api/pix/v1/participants>
- API Brasil - Corretoras CVM: <https://brasilapi.com.br/api/cvm/corretoras/v1>