

Process Design Document (PDD)

Projeto: Raspagem de dados sobre População por Estados e Regiões

1. Objetivo

Desenvolver um robô para coletar informações sobre estados brasileiros, capitais, populações e regiões. As informações serão estruturadas, armazenadas em um banco de dados SQLite e exportadas para análises estatísticas. Este processo automatizado tem como objetivo consolidar e analisar os dados para identificar regiões mais populosas e estados com capitais de maior população.

2. Contexto Operacional

Executor: Gustavo Rambaldi Gonçalves Vidal Magalhães

Cargo: Analista de Inteligência de Mercado

Tempo médio manual: 30 minutos

Carga mensal de execução: 30

SLA: N/A

3. Descrição do Processo

3.1 Fluxo Geral do Processo

1. Extração de Dados de Estados, Capitais e Regiões:

- Navegar em um [site](#) para coletar informações de estados brasileiros, suas capitais e regiões utilizando Selenium.

Sigla	Estado	Capital	Região	Region
AC	ACRE	RIO BRANCO	NORTE	NORTH
PA	PARÁ	BELÉM	NORTE	NORTH
RO	RONDÔNIA	PORTO VELHO	NORTE	NORTH
RR	RORAIMA	BOA VISTA	NORTE	NORTH
TO	TOCANTINS	PALMAS	NORTE	NORTH
MA	MARANHÃO	SÃO LUIZ	NORDESTE	NORTHEAST
PB	PARAÍBA	JOÃO PESSOA	NORDESTE	NORTHEAST
PE	PERNAMBUCO	RECIFE	NORDESTE	NORTHEAST
PI	PIAUÍ	TEREZINA	NORDESTE	NORTHEAST
RN	RIO GRANDE DO NORTE	NATAL	NORDESTE	NORTHEAST
SE	SERGIPE	ARACAJÚ	NORDESTE	NORTHEAST
GO	GOIÁS	GOIÂNIA	CENTRO-OESTE	MIDWEST
MS	MATO GROSSO DO SUL	CAMPO GRANDE	CENTRO-OESTE	MIDWEST
MT	MATO GROSSO	CUIABÁ	CENTRO-OESTE	MIDWEST
ES	ESPÍRITO SANTO	VITÓRIA	SUDESTE	SOUTHEAST
MG	MINAS GERAIS	BELO HORIZONTE	SUDESTE	SOUTHEAST
RJ	RIO DE JANEIRO	RIO DE JANEIRO	SUDESTE	SOUTHEAST
SP	SÃO PAULO	SÃO PAULO	SUDESTE	SOUTHEAST
RS	RIO GRANDE DO SUL	PORTO ALEGRE	SUL	SOUTH
SC	SANTA CATARINA	FLORIANÓPOLIS	SUL	SOUTH
AM	AMAZONAS	MANAUS	NORTE	NORTH
AP	AMAPÁ	MACAPÁ	NORTE	NORTH
AL	ALAGOAS	MACEIÓ	NORDESTE	NORTHEAST
BA	BAHIA	SALVADOR	NORDESTE	NORTHEAST
CE	CEARÁ	FORTALEZA	NORDESTE	NORTHEAST
PR	PARANÁ	CURITIBA	SUL	SOUTH

2. Leitura de População das Capitais:

- Conteúdo presente no arquivo PopulaçãoCapital.xlsx.

3. Estruturação de Dados:

- Organizar as informações extraídas.

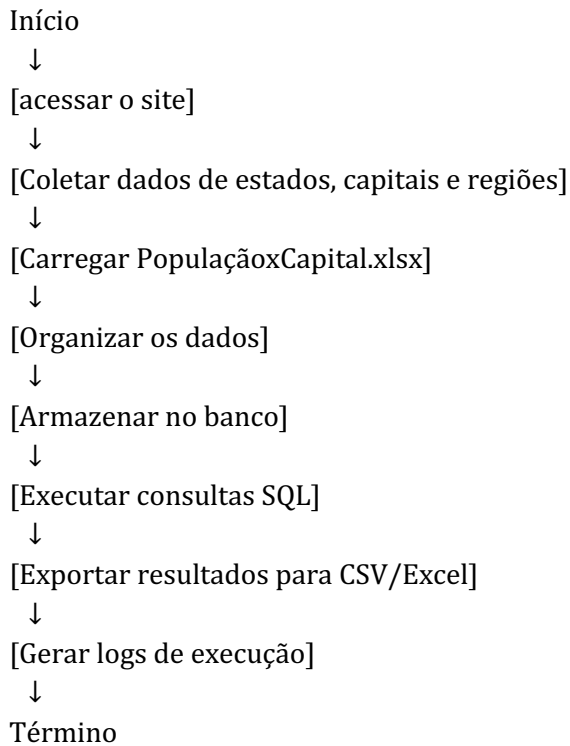
4. Armazenamento em Banco de Dados SQLite:

- Criar ou atualizar uma tabela chamada estados.

5. Análise de Dados e Exportação:

- Possibilidade de obter um arquivo gerado a partir dos dados contidos na base em formato de planilha para a área interessada.

4. Fluxograma do Processo



5. Entrada e Saída do Processo

Entradas:

1. Site com informações de estados, capitais e regiões.
2. Arquivo PopulaçãoCapital.xlsx.

Saídas:

1. Arquivo top3_regioes_populosas.csv.
2. Arquivo regioes_n_capitais.xls.
3. Arquivo estados_mais_populosos.xls.
4. Banco de dados SQLite consolidado.

6. Referências e Anexos

- PopulaçãoCapital.xlsx (anexado).
- Link do site utilizado para extração de dados.