

Relatório Projeto Integrador DATASUS

DATASUS é o Departamento de Informática do Sistema Único de Saúde no Brasil. Trata-se de um órgão vinculado ao Ministério da Saúde, responsável por gerenciar e disponibilizar informações sobre saúde no país. O principal objetivo do DATASUS é coletar, processar, analisar e disseminar dados relacionados à saúde para apoiar o planejamento, gestão e avaliação do sistema de saúde brasileiro.

O propósito deste projeto consiste na elaboração de uma base de dados contendo informações sobre as internações hospitalares ocorridas entre os anos de 2019 e 2023, com uma granularidade temporal mensal e uma agregação geográfica por município.

Web Scraping

A coleta dos dados foram feitas com uma biblioteca em Python chamada “Selenium”. Selenium é uma ferramenta muito poderosa e popular no ecossistema Python, especialmente no campo da automação de testes e interações com navegadores web. Ela oferece uma gama de funcionalidades para controlar e interagir com navegadores da web de forma programática.

O código feito utilizou o “Chromium”, que é um Webdriver do navegador Chrome. No código acessamos o site do DATASUS e pegamos as informações de 2019 até 2023, após toda a coleta, foi gerado um arquivo em CSV para a leitura no Python ou SAS Viya.

Carga de Dados

Após a conclusão de todos os estágios de tratamento dos dados, procedemos com a carga manual utilizando a ferramenta DBeaver para realizar a inserção dos dados no banco de dados PostgreSQL.

Gráficos

Para realizarmos os gráficos utilizamos uma biblioteca em Python chamada “matplotlib”, que é uma ferramenta poderosa em Python para criação de gráficos e visualizações de dados. Ela oferece uma ampla variedade de funções e métodos que permitem aos usuários gerar gráficos de alta qualidade em diversos formatos.