



V Mostra de Extensão, Ciência e Tecnologia

XXX Seminário de Iniciação Científica
XV Salão de Ensino e Extensão
V Mostra da Pós-Graduação Stricto Sensu
IV Seminário de Inovação Tecnológica

De 28 de outubro a
01 de novembro de 2024

INSCRIÇÕES ABERTAS

 UNISC

Título:	AUTOMAÇÃO PARA A COLETA E O TRATAMENTO DE DADOS NO TABNET		
Autores:	Murilo Antonio Wollmann Leonardo Antonio Patatt Ana Paula Goetze Daniela Duarte da Silva Bagatini Priscila Schmidt Lora Gabriela de Andrade Monteiro Ana Paula Müller Débora Oliveira da Silva Cristiane Drebes Pedron		
Área	<input type="checkbox"/> Humanas <input type="checkbox"/> Sociais Aplicadas <input type="checkbox"/> Biológicas e da Saúde <input checked="" type="checkbox"/> Exatas, da Terra e Engenharias	Dimensão:	<input type="checkbox"/> Ensino <input checked="" type="checkbox"/> Pesquisa <input type="checkbox"/> Extensão <input type="checkbox"/> Inovação
Resumo: <p>A coleta e análise de dados são essenciais para gerar informações da saúde pública, pois permitem monitorar a ocorrência de doenças e desenvolver estratégias de intervenção. No entanto, esse processo envolve etapas manuais e demoradas, que começam com a seleção de um conjunto de informações, definição de abrangência, escolha de variáveis, seleção de períodos e terminam com a visualização dos resultados. Plataformas como o TABNET (ferramenta de tabulação <i>on-line</i> do DATASUS) não oferecem opções automatizadas de concatenação de dados, o que torna a atualização contínua dos dados um desafio significativo. O objetivo deste trabalho foi desenvolver uma aplicação que automatiza a busca e o tratamento de dados no TABNET, utilizando a linguagem de programação Python e bibliotecas específicas. A aplicação facilita o <i>download</i> de tabelas, automatiza a concatenação dos dados e corrige inconsistências na formatação dos dados (como a falta de padronização do número de colunas das tabelas), permitindo uma coleta mais eficiente e precisa. Para atingir esse objetivo, foram necessários os seguintes passos metodológicos: (1) compreensão do sistema de tabulação do DATASUS, o TABNET; (2) estudo da biblioteca Selenium, versão 1.0.14, que simula as ações do usuário e automatiza tarefas repetitivas; (3) estudo da biblioteca Pandas, versão 2.1.4 para ciência de dados, que permite a manipulação, limpeza e análise de grandes volumes de dados; (4) identificação dos dados de interesse</p>			



V Mostra de Extensão, Ciência e Tecnologia

XXX Seminário de Iniciação Científica
XV Salão de Ensino e Extensão
V Mostra da Pós-Graduação Stricto Sensu
IV Seminário de Inovação Tecnológica

De 28 de outubro a
01 de novembro de 2024

INSCRIÇÕES ABERTAS

 UNISC

para pesquisa na linha de cuidado do câncer de mama, definindo as variáveis e indicadores para a análise junto com especialista médica mastologista; (5) desenvolvimento da automação para navegar no *site* do TABNET e realizar o *download* dos arquivos em formato CSV; (6) desenvolvimento da automação para cruzar e tratar os dados, garantindo a uniformidade e correção das informações coletadas; (7) realização de testes e validações para assegurar que a aplicação execute as tarefas corretamente e de forma eficiente, com a verificação da exatidão dos dados coletados. Portanto, a automação acessa o TABNET, realiza o *download* dos arquivos de interesse, faz a concatenação das tabelas e corrige problemas de apresentação dos dados. Como resultado, a aplicação demonstrou ser eficiente e confiável, eliminando a necessidade de intervenções manuais repetitivas, demoradas e sujeitas a erros. Além disso, permitiu uma redução do tempo necessário para a coleta e o tratamento dos dados epidemiológicos, agilizando a atualização contínua desses dados. A solução desenvolvida provou ser uma ferramenta valiosa para a obtenção de dados para análise e pode ser aplicada em diferentes contextos de coleta de dados. Como contribuição científica, o projeto oferece uma solução computacional para a automação da coleta e tratamento de dados epidemiológicos, facilitando a obtenção de dados para pesquisas na área da saúde pública. Para trabalhos futuros, a aplicação poderá abranger outros tipos de dados disponíveis no TABNET e integrar funcionalidades adicionais para aprimorar a eficiência e a precisão no tratamento dos dados. Este trabalho faz parte do projeto de pesquisa financiado pelo Edital Chamada Nº 21/2023 - Estudos Transdisciplinares em Saúde Coletiva.

Link do Vídeo: <https://drive.google.com/file/d/1dT4W4M-xXHIwo5d3kGp0cltGV25-iZJh/view?usp=sharing>