



XXXV SALÃO de INICIAÇÃO CIENTÍFICA

6 a 10 de novembro

| | |
|-------------------|---|
| Evento | Salão UFRGS 2023: SIC - XXXV SALÃO DE INICIAÇÃO CIENTÍFICA DA UFRGS |
| Ano | 2023 |
| Local | Campus Centro - UFRGS |
| Título | Coleta automatizada de dados |
| Autor | GUILHERME DA SILVA FRAINER |
| Orientador | MARCELO SAVINO PORTUGAL |

O processo de coleta de dados é essencial no dia-a-dia de um economista, sendo necessário estar sempre em dia com as principais estatísticas, tanto no mundo profissional quanto no acadêmico. Contudo, esse processo por vezes pode ser repetitivo e tedioso, pois envolvem sempre o acesso aos mesmos sítios e a busca pelas mesmas séries temporais. Desse modo, justificam-se tentativas de automatizar esse processo. Nesse sentido, o objetivo desta pesquisa foi de encontrar meios de automatizar a coleta de dados das principais séries temporais de variáveis macroeconômicas (inflação, PIB, desemprego). Para tanto, escreveram-se programas na linguagem de programação Python que acessam essas bases de dados remotamente de maneira automatizada. Os programas se utilizam de diversas bibliotecas, como a *requests* para a maior parte dos acessos, mas também a biblioteca *beautifulsoup4* quando não há uma API apropriada. Ademais, no caso específico das séries temporais do SIDRA, utilizou-se a biblioteca *sidrapy*. Para a interpretação dos dados devolvidos, utilizou-se a biblioteca oficial *json*. Para a posterior manipulação dos dados, indicaram-se as bibliotecas *matplotlib*, *pandas* e *xlsxwriter*. Como resultado, encontrou-se que o processo de coleta de dados é facilmente automatizado, dessa forma poupando incontáveis horas de trabalho humano. As estatísticas cuja obtenção foi automatizada incluem IPCA, CPI, CAGED, PIB brasileiro, taxa de câmbio dólar/euro.