

IBPAD | DataScience Team - Exercício OpenSUS Covid

Exercício para testar habilidades básicas de programação em Python ou R. Para este exercício é necessário baixar duas bases de dados do OpenDataSUS (uma via API e outra através de um link de um arquivo comprimido), dar um merge entre as bases para agregar novas informações na base de vacinados acerca dos estabelecimentos do SUS, e por último exportar a base no formato de arquivo requisitado.

- Passo 1 - Download das 10.000 primeiras linhas da base de vacinados do OpenDataSUS via API (Não existe necessidade do download da base inteira. Apenas uma requisição de size igual a 10.000). Segue o link do manual da API: https://opendatasus.saude.gov.br/dataset/b772ee55-07cd-44d8-958f-b12edd004e0b/resource/5916b3a4-81e7-4ad5-adb6-b884ff198dc1/download/manual_api_vacina_covid-19.pdf
- Passo 2 - Verificar a presença de outliers na base de vacinados a partir da variável idade [Utilizar o método Z-score para verificar os outliers, aplicando o índice limite (threshold) de 2], e remoção dessas observações da base
- Passo 3 - Download da base com as informações de todos os estabelecimentos do SUS através do link: https://sage.saude.gov.br/dados/repositorio/cadastro_estabelecimentos_cnes.zip
- [O processo de download, unzip do arquivo CSV e sua leitura deve ser realizado via script. Sem interferência manual.]
- Passo 4 - Merge da base de estabelecimentos com a base de vacinados para agregar mais informações sobre as unidades básicas de saúde. [Nessa etapa é importante identificar a chave de cruzamento e garantir que as duas chaves estejam no mesmo formato, e lembrar que serão adicionados dados na base principal, que é a base de vacinados]
- Passo 5 - Exportação da base de vacinados com os dados agregados da base de estabelecimentos no formato XLSX (Excel) e no formato CSV [Não exportar com a coluna Index ou Rownames]