

## **RELATÓRIO DE ANÁLISES: Equipe 6 – NutriSquad**

Esse relatório tem o propósito de apresentar o passo a passo (algoritmo) do tratamento dos dados durante o *datathon*. É uma versão guiada para o *github* da equipe, seguindo o fluxograma de decisões e facilitando a ordem de acesso aos arquivos e avaliação final.

A equipe escolheu o desafio 2 para investigação: **“Qual o impacto do Suplemento Nutricional na prevenção de quedas na população idosa?”**

Para responder ao desafio, foram utilizadas duas das bases de dados disponibilizadas. O acesso ao microdados (bases cruas) pode ser feito por meio dos links abaixo:

- POF 2017/2018: <http://datathon.polodados.com/#ibge/#pesquisa-de-orcamentos-familiares-2017-2018>
- CIHA – DataSUS: <http://datathon.polodados.com/#datasus/#comunicacao-de-informacao-hospitalar-e-ambulatorial>

A base da POF foi selecionada por conter informações acerca da ingestão alimentar da população brasileira, variável de interesse do estudo. A base do CIHA foi utilizada para detecção de hospitalizações (por meio do CID) no período de interesse. Inicialmente, a base CIHA foi também filtrada apenas para os anos 2017/2018, de modo a equiparar ao mesmo período de avaliação da POF.

Para a base da CIHA, além do filtro por ano, foi também criado a planilha de CID referente aos desfechos de queda por UF. A seleção dos CID a serem incluídos considerou, além do desfecho principal (QUEDAS – CID 10 - W19), outros desfechos secundários indicativos de quedas em idosos, como fratura de fêmur, fragilidade e sarcopenia. A escolha dos CID foi realizada com base em artigos na literatura que também avaliaram o desfecho de quedas e mentorias com a área médica durante o processo de análises.

O tratamento inicial por ano, filtro de idade (60 anos ou mais) e desfechos de interesse podem ser acessados nos arquivos do *github* com caminho abaixo. Na planilha de CID, casos de desfecho de quedas foram categorizados como “1” e demais desfechos e causas de hospitalizações como “0”. Casos que não possuíam nenhum CID de registro (primário ou secundário) ou idade (essencial para categorização posterior) foram excluídos:

- **Base de casos de desfecho de quedas:** Bases de dados → CIHA → CID 10\_casos

As bases cruas e primeiros tratamentos e filtros não foram inseridas no repositório por conta do tamanho das mesmas e espaço disponível no recurso do github.

Após o tratamento e filtro por idade e casos, a base CIHA foi agrupada por UF de acordo com o número total de casos de hospitalizações no período de interesse (2017/2018) e do número total de casos decorrentes de desfechos de quedas.

Filtros e unificações CIHA:

- **Algoritmos → ciha\_concat\_tabelas.ipynb**
- **Algoritmos → ciha\_filtro\_idade.ipynb**

Para a base da POF 2017/2018, a limpeza e tratamento inicial, considerando o agrupamento de alimentos para obtenção da base de dados de consumo alimentar e nutrientes por indivíduo foi realizada conforme descrito no documento “**TRATAMENTO BASES DE DADOS**” dentro da pasta de algoritmos.

A base de nutrientes foi então filtrada para idade, relato plausível do consumo de energia e valores aberrantes (outliers) de altura. Todo o tratamento posterior na base de nutrientes, além da simulação de suplementação está descrito no código “**Cálculo de nutrientes e suplementação POF**”, também dentro da pasta de algoritmo.

A base de dados final da POF, de análise de nutrientes é a “**Base de Suplementação**” dentro da pasta das bases de dados. Após o filtro por idade, a base apenas de idosos também (“**Base POF unificada Idosos**”) foi gerada e utilizada para categorização por UF e, posteriormente, unificação com a base da CIHA para geração da “**DATASET\_FINAL**”.

- Análises descritivas: **Algoritmo → Análise\_Dataset\_Unificado**

Por fim, os procedimentos de análise por geolocalização e geração de mapas com os desfechos de interesse foram realizados no software Power BI, com o dashboard no link abaixo:

[https://app.powerbi.com/view?r=eyJrJmJlZG1iZDNDNDNDNmNWQ2LTQwOTgtNDYwNy05ODE2LWRIZDY1ZjIhMmU0YSJ9zFmNTI1liwidCI6IjhiNDNDNDNDNmNWQ2LTQwOTgtNDYwNy05ODE2LWRIZDY1ZjIhMmU0YSJ9](https://app.powerbi.com/view?r=eyJrJmJlZG1iZDNDNmNWQ2LTQwOTgtNDYwNy05ODE2LWRIZDY1ZjIhMmU0YSJ9zFmNTI1liwidCI6IjhiNDNDNmNWQ2LTQwOTgtNDYwNy05ODE2LWRIZDY1ZjIhMmU0YSJ9)

Todos os resultados foram agrupados e selecionados para inclusão no vídeo do *pitch* de modo a apresentar os insights da melhor maneira possível.