

Fontes de Dados e Métodos de Coleta

1. Fonte de Dados Utilizada

A base de dados utilizada neste projeto é proveniente do relatório oficial do Vigitel Brasil 2023 — Vigilância de Fatores de Risco e Proteção para Doenças Crônicas por Inquérito Telefônico, publicado pelo Ministério da Saúde. Trata-se de uma pesquisa nacional realizada anualmente, que investiga comportamentos relacionados à saúde em todas as capitais brasileiras e no Distrito Federal.

Link direto para o relatório:

https://bvsms.saude.gov.br/bvs/publicacoes/vigitel_brasil_2023.pdf

2. Tipo de Dados

Os dados são públicos, secundários, não confidenciais e apresentados em formato de relatório PDF contendo tabelas percentuais por sexo, idade e outros fatores sociodemográficos. Todos os indicadores analisados se referem à população adulta das capitais brasileiras.

3. Indicadores Selecionados

Foram extraídos e estruturados os seguintes grupos de indicadores:

- Atividade física no tempo livre, inatividade física e deslocamento ativo (bicicleta)
- Alimentação: consumo de frutas, hortaliças, ultraprocessados e refrigerantes
- Doenças crônicas e fatores de risco: obesidade, sobrepeso, hipertensão, diabetes, depressão, tabagismo e álcool

4. Método de Coleta e Estruturação

Os dados foram copiados diretamente das tabelas do relatório oficial em PDF. As informações foram organizadas em uma planilha estruturada em três abas: Atividade_Fisica, Alimentacao e Doencas_Cronicas. Cada aba foi formatada com colunas padronizadas contendo os campos: Indicador, Total (%), Homens (%) e Mulheres (%). Todos os valores foram convertidos para formato numérico compatível com ferramentas de BI.

5. Ferramentas Utilizadas

- PDF Reader (Adobe Acrobat)
- Google Sheets (formatação dos dados)
- Looker Studio (visualizações interativas)
- Python + Pandas (estruturação inicial automatizada)