

**Faculdade De Tecnologia
Pastor Enéas Tognini**

**Breno Juan da Fonseca Silva,
Dayana Ingrid Carata Choque,
Gabriel Leão Frigo,
Jackelyne Alicia Miranda Ramos,
Lucas Barbosa Defanti,
Lucas Matos Norbertino dos Santos,
Matheus Henrique Bianco Neres,
Michelle Mieko Coelho Koga**

Monitoramento de Causas de Morte por meios interativos

São Paulo - SP

2024

**Breno Juan da Fonseca Silva,
Dayana Ingrid Carata Choque,
Gabriel Leão Frigo,
Jackelyne Alicia Miranda Ramos,
Lucas Barbosa Defanti,
Lucas Matos Norbertino dos Santos,
Matheus Henrique Bianco Neres,
Michelle Mieke Coelho Koga**

Monitoramento de Causas de Morte por meios interativos

Relatório Técnico-Científico apresentado na disciplina de Projeto Integrador para o curso de Big Data para Negócios da Faculdade de Tecnologia Ipiranga “Pastor Enéas Tognini” (FATEC).

Orientador: Prof. Bruno Monserrat Perillo.

São Paulo - SP

2024

RESUMO

CHOQUE, D.I. Catarata *et al.* **Monitoramento de Causas de Morte por meios interativos.** 16 f. Relatório Técnico-Científico. Big Data para Negócios - **Pastor Enéas Tognini, Faculdade De Tecnologia**, São Paulo - SP, 2024.

O objetivo do presente trabalho é propor métodos simples de programação, *Big Data* e visualização de dados para monitorar causas de morte a partir dos dados extraídos do Sistema de Informação sobre Mortalidade Declaração de Óbitos (SIM-DO) disponibilizado pelo Sistema Único de Saúde (SUS). Como trabalho parcial realizamos a extração e o tratamento dos dados por meio da linguagem de programação *R*, utilizando o ambiente do *RStudio* e bibliotecas específicas para o Sistema de Informações sobre Mortalidade (SIM).

Palavras-chave: Relatório Técnico, Mortalidade, Causas de morte, Monitoramento, Visualização.

LISTA DE FIGURAS

| | |
|--|----|
| Figura 1 – Distribuição do Sexo na amostra | 11 |
| Figura 2 – Distribuição de Raça/Cor na amostra | 11 |

LISTA DE TABELAS

LISTA DE QUADROS

LISTA DE ABREVIATURAS E SIGLAS

| | |
|--------|--|
| BI | <i>Business Intelligence</i> |
| CEI | Causas Externas Inespecíficas de Mortalidade |
| ETL | Processo de Extração, Transformação e Carregamento |
| SIM | Sistema de Informações sobre Mortalidade |
| SIM-DO | Sistema de Informação sobre Mortalidade Declaração de Óbitos |
| SUS | Sistema Único de Saúde |

SUMÁRIO

| | | |
|------------|--|-----------|
| 1 | INTRODUÇÃO | 8 |
| 2 | DESENVOLVIMENTO | 9 |
| 2.1 | Objetivo | 9 |
| 2.2 | Fundamentação Teórica | 9 |
| 2.3 | Aplicação das disciplinas estudadas no Projeto Integrador | 10 |
| 2.4 | Metodologia | 10 |
| 3 | RESULTADOS | 11 |
| | REFERÊNCIAS | 13 |
| | APÊNDICES | 14 |
| | APÊNDICE A – CÓDIGO PARA EXTRAÇÃO DOS DADOS EM FOR- | |
| | MATO RMD | 15 |

1 INTRODUÇÃO

O monitoramento das causas comuns de morte é uma preocupação antiga, Malta *et al.* (2007). destaca que a construção de uma lista de mortes evitáveis é um primeiro passo para o monitoramento que, então, pode orientar as políticas públicas para a prevenção, como um exemplo pratico (Marson *et al.*, 2010) realizaram um estudo no Hospital Universitário de Londrina e achou uma possível relação entre mortes por trauma e a não aplicação de diretrizes de atendimento, dado este fato se conclui que uma parcela de mortes poderia ser evitada caso o problema fosse identificado e as diretrizes não fossem ignoradas.

Outro desafio é o registro dos dados nos sistemas, tanto pela não inclusão como pela qualidade dos dados, Muzy, Castanheira e Romero (2021) destaca que os dados disponibilizados no SIM muitas vezes são incompletos ou defasados e o preenchimento dos dados é feito de forma ambígua.

O objetivo deste relatório é propor uma solução para o problema do monitoramento dado os dados disponíveis pelo sistema SIM, o processo é feito de forma simples ao articular soluções de fácil acesso.

2 DESENVOLVIMENTO

2.1 Objetivo

O objetivo do presente relatório é apresentar uma solução simples que não demande um conhecimento técnico avançado, a partir dos dados disponíveis no SIM-DO será aplicado o processo de Extração, Transformação e Carregamento (ETL) da base de dados para a construção de um *Dashboard* que possibilite o monitoramento das causas de morte, dado as limitações computacionais o escopo será limitado ao Estado de São Paulo na janela temporal entre os anos de 2010 e 2014.

2.2 Fundamentação Teórica

O monitoramento de métricas para a saúde publica se mostra eficaz na produção de resultados, Filho *et al.* (2019) realizou um estudo em 60 cidades, participantes do projeto Dados para a Saúde, que busca melhorar o diagnóstico de causas de morte por meio da cooperação de equipes de vigilância de óbitos. As cidades participantes do projeto Dados para a Saúde apresentaram melhores resultados após a investigação das CEI, possibilitando uma análise detalhada da reclassificação para causas específicas, considerando sexo e faixas etárias.

Os resultados apresentados por Filho *et al.* (2019) mostram a importância do monitoramento em conjunto, adicionalmente destaca a melhoria contínua nos sistemas relevantes, como o SIM-DO.

O monitoramento das causas de morte são importantes insumos para a construção de indicadores de efetividade dos serviços de saúde, Malta e Duarte (2007) destaca a importância de uma lista de causas evitáveis para a construção de indicadores, porem a necessidade de levar em consideração fatores fora da cobertura dos sistemas de saúde, este problema necessita de uma abordagem mais complexa e fora do escopo deste trabalho.

A mortalidade por causas evitáveis também podem ser resultado de problemas sociais, que devem ser monitorados e enfrentados não só pelos sistemas de saúde, Malta *et al.* (2021) destaca que a mortalidade de adolescentes e adultos jovens apresentam tendência de alta por homicídios, e dados do Fórum de Segurança Pública apontam uma sub-notificação, em especial entre jovens negros que residem em periferias urbanas, deste modo problemas de completude do SIM-DO são importantes para políticas públicas em um escopo geral, e não apenas na área da saúde.

Dado o exposto o presente trabalho se propõe a apresentar uma solução base, ao se utilizar dos dados presentes no SIM, e que possa incorporar molharias a medida em que os problemas apresentados sejam melhor compreendidos e a completude dos dados seja melhorada.

2.3 Aplicação das disciplinas estudadas no Projeto Integrador

Durante a execução do projeto, aplicamos conhecimentos e técnicas adquiridas em disciplinas como ETL, Arquitetura de Dados, *Big Data Analytics*, Análise de Dados, Estatística I e II, Programação, *BI (Business Intelligence)* e *Storytelling*. Adicionalmente o fluxo de trabalho foi organizado por meio da ferramenta *Kanban*, a partir dos conceitos aprendidos em diversas disciplinas sobre metodologias de projeto.

Na fase de coleta e preparação de dados, utilizamos técnicas de limpeza e integração de dados aprendidas ao longo dessas disciplinas. Em seguida, na fase de análise, empregaremos técnicas estatísticas para extrair *insights* significativos dos dados. Finalmente, na fase de visualização e interpretação dos resultados, utilizaremos habilidades de programação para desenvolver *dashboards* interativos e relatórios.

2.4 Metodologia

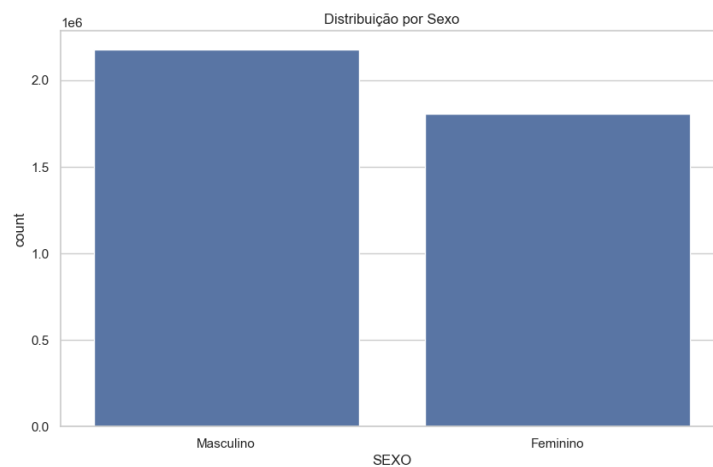
A extração será feita por meio dos *softwares R* e *RStudio*, onde foram utilizadas as funções da biblioteca desenvolvidas por Saldanha, Bastos e Barcellos (2019), esta biblioteca foi criada para facilitar a extração dos dados disponibilizados pelo SUS, incluindo o sistema SIM, adicionalmente são disponíveis funções de pré processamento que tratam os valores presentes nas tabelas e, a partir dos dados extraídos, inclui informações de municipalidade em um formato utilizável pelas ferramentas de mapas presentes na maioria dos *softwares* de visualização, toda a documentação e instrução de uso são disponíveis pela ferramenta de ajuda (documentação) da própria biblioteca.

De forma parcial a exploração dos dados foi feita por meio da linguagem *Python* em um arquivo no formato *Jupyter Notebook* de forma colaborativa por meio de versionamento pelo *GitHub*, que possibilita a colaboração independente das ferramentas de preferência de cada colaborador.

3 RESULTADOS

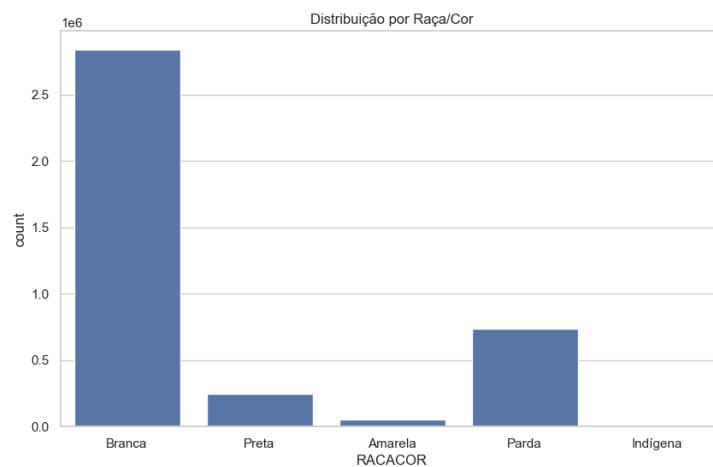
Como resultados parciais foi possível observar os dados de mortes em geral, foi observado que a mortalidade dos homens está em 54% (Figura 1), apontando um possível desbalanceamento a ser investigado, além disto há uma incidência de mortes de pessoas brancas de 73% (Figura 2), este é um número a ser considerado com cautela, dado a problemática envolvendo o preenchimento dos dados apontado por Muzy, Castanheira e Romero (2021).

Figura 1 – Distribuição do Sexo na amostra



Fonte: SIM-DO e Resultados da Pesquisa

Figura 2 – Distribuição de Raça/Cor na amostra



Fonte: SIM-DO e Resultados da Pesquisa

Para a conclusão do trabalho é necessária a construção do *Dashboard*, além de considerar as causas de morte de interesse inicial, no momento de conclusão deste relatório parcial está sendo utilizada a lista de causas evitáveis de morte proposta por Malta *et al.* (2007).

REFERÊNCIAS

- FILHO, A. M. S. *et al.* Melhoria da classificação das causas externas inespecíficas de mortalidade baseada na investigação do óbito no brasil em 2017. **Revista Brasileira de Epidemiologia**, SciELO Brasil, v. 22, p. e190011–supl, 2019.
- MALTA, D. C.; DUARTE, E. C. Causas de mortes evitáveis por ações efetivas dos serviços de saúde: uma revisão da literatura. **Ciência & Saúde Coletiva**, SciELO Public Health, v. 12, p. 765–776, 2007.
- MALTA, D. C. *et al.* Lista de causas de mortes evitáveis por intervenções do sistema único de saúde do brasil. Secretaria de Vigilância em Saúde, Ministério da Saúde, 2007.
- MALTA, D. C. *et al.* Mortalidade de adolescentes e adultos jovens brasileiros entre 1990 e 2019: uma análise do estudo carga global de doença. **Ciência & Saúde Coletiva**, SciELO Brasil, v. 26, p. 4069–4086, 2021.
- MARSON, A. C. *et al.* Mortes evitáveis em pacientes de trauma associadas a não adesão às diretrizes de atendimento. **Revista Brasileira de Terapia Intensiva**, SciELO Brasil, v. 22, p. 220–228, 2010.
- MUZY, J.; CASTANHEIRA, D.; ROMERO, D. Análise da qualidade da informação da mortalidade prematura por doenças crônicas não transmissíveis e sua utilização nos objetivos de desenvolvimento sustentável. **Cadernos Saúde Coletiva**, SciELO Brasil, v. 29, p. 152–164, 2021.
- SALDANHA, R. d. F.; BASTOS, R. R.; BARCELLOS, C. Microdatasus: pacote para download e pré-processamento de microdados do departamento de informática do sus (datasus). **Cadernos de Saúde Pública**, SciELO Public Health, v. 35, p. e00032419, 2019.

APÊNDICES

APÊNDICE A – CÓDIGO PARA EXTRAÇÃO DOS DADOS EM FORMATO RMD

```

---
title: "Extracao"
author: "Lucas Barbosa Defanti"
date: "2024-03-16"
output: html_document
---

```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
if (!"pacman" %in% rownames(installed.packages())){
 install.packages("pacman")
}
pacman::p_load("remotes",
"tidyverse",
"dplyr",
"devtools",
"Rcpp")

if (!"read.dbc" %in% rownames(installed.packages())){
 devtools::install_github('danicat/read.dbc')
}

if (!"microdatasus" %in% rownames(installed.packages())) {
 devtools::install_github('rfsaldanha/microdatasus')
}

library(dplyr)
library(tidyverse)
library(remotes)
library(microdatasus)
```

# Extraíndo
```{r extact}
dados <- microdatasus::fetch_datasus(
year_start = 2010,

```



```
year_end = 2024,
uf = "SP",
information_system = "SIM-DO")
```\n\n# Processamento  
```\n{r process}  
df <- microdatasus::process_sim(dados, municipality_data = TRUE)
rm(dados)
```\n\n# Head  
```\n{r peek}  
head(df)
```\n\n# Export  
```\n{r export}  
data.table::fwrite(df, "datasus.csv")
rm(df)
```\n
```