# Regional and temporal patterns analysis of mortality in Brazil supported by a data warehouse.

*Débora Oliveira Santana[1], Juracy Araújo de Almeida Junior[1]*
[1] Universidade Federal da Bahia

**Abbreviated abstract:** Information on mortality for the years 2016 to 2020 adds up to more than 6 million registered deaths. Developing a data warehouse framework on this data helps consolidate it while maintaining quality and accuracy. In addition, clustering these data facilitates the recognition of mortality patterns to help support the development of public health policies that enable more effective management. The results show that the patterns have been maintained over the years, with the exception of the Covid-19 pandemic period, which had a strong impact on the age structure of the population.

**Related publications:**
– ALVES, D. d. S. B. *et al*, Mineração de dados na identificação de padrões de mortalidade no Brasil de 1979 a 2013 (2017)

**ECD**
UFBA

4th Conference on
**Statistics and
Data Science**
Salvador, Brazil (online)
December 1-3, 2022

# Problem

How to obtain a view of the panorama of mortality in Brazil and its influence on the demographic context to support the identification and decision of the necessary actions to reduce mortality in Brazilian regions?

### Social inequality

The Pan American Health Organization highlights the need to reduce inequalities in the integration of cross-cutting issues (equity, gender, ethnicity and human rights) as this directly influences global public health and, consequently, mortality rates.

+

### High mortality rates

The Brazilian epidemiological transition is characterized by high rates of morbidity and mortality from chronic non-communicable diseases that coexist with a high incidence and prevalence of infectious-parasitic diseases and external causes, especially homicides.

+

### Large mass of data

Information on mortality for the years 2016 to 2020 total more than 6 million deaths recorded in the Mortality Information System.

### Public health problem

The accelerated processes of urbanization and population aging require the Unified Health System to carry out analyzes and planning to identify the best strategies that guarantee the population's effective access to the system and directly affect social and economic determinants.
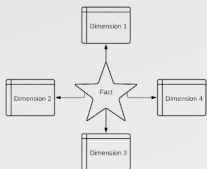
4th Conference on
**Statistics and Data Science**
Salvador, Brazil (online)
December 1-3, 2022

# Search architecture

## Mortality Information System

Identification of mortality data for the years 2016 to 2020.



**Excel**

## 05

## Data analysis

Identification of mortality patterns.

## 01

## 04

## Star schema

Construction of the data warehouse architecture.

**MySQL**

**pentaho**
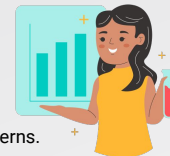
## Data grouping

Application of the clustering algorithm.

**python**

## 02

## 03

**Power BI**

## Data visualization

Dashboard creation.

**E C D**
U F B A
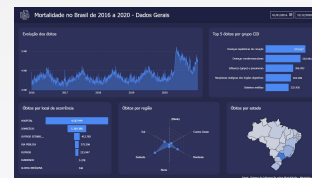
4th Conference on
**Statistics and
Data Science**
Salvador, Brazil (online)
December 1-3, 2022

# Results

From 2016 to 2020, mortality in Brazil **averaged 3.759 deaths** per year.

From March 2020 to December 2020 (**pandemic period**) the average number of deaths **increased by 17.5% compared to the same period in 2019**.



Top 5 óbitos por grupo da classificação internacional de doenças

| | |
|---|---|
| Doenças isquêmicas do coração | 573.617 |
| Doenças cerebrovasculares | 503.981 |
| Influenza [gripe] e pneumonia | 396.993 |
| Neoplasias malignas dos órgãos digestivos | 365.286 |
| Diabetes mellitus | 332.420 |

During the Covid-19 pandemic, the number of deaths of **elderly people over the age of 70 increased by 16%**.

Increase in male deaths between the **ages of 15 and 23** from **violent causes**.

In **women**, the death ranking refers to c**erebrovascular diseases**.

**Elderly**, identified as **white males** from the **Southeast**, are close to individuals who died from **ischemic heart disease**.

**Adults**, identified as **mixed-race males** from the **Southeast and Northeast**, are close to individuals who died as a result of **aggression**.

**Elderly**, identified as **mixed-race females** from the **Southeast and Northeast**, are close to individuals who died from **ischemic heart disease and cerebrovascular diseases**.

**Elderly**, identified as **white females** from the **Southeast**, are close to individuals who died from **influenza and pneumonia**.

| | | | | | | |
|---|---|---|---|---|---|---|
| Evolution of mortality in Brazil over 5 years | Mortality rate in Brazilian regions | 5 CID's with the highest incidence of mortality in Brazil | Behavior of historical series data in Brazilian regions | Data behavior in relation to the characteristics of individuals | Places of occurrence where there is a higher incidence of mortality | Groups of factors that directly influence the death of an individual |



Taxa de mortalidade por região (permilagem)

| Sudeste | Sul | Nordeste | Centro-Oeste | Norte |
|---|---|---|---|---|
| 7,07 | 6,90 | 6,32 | 5,62 | 4,81 |

The **highest daily average of deaths** over the five years was in the **Southeast**, with a value of 1,708 deaths per day.

**Aggressions** enter the **ranking** of the five reasons that occurred more deaths in the **North, Northeast and Midwest regions**.

**There are under-records in the data of the population that suffers from greater social inequality and, consequently, has less access to health services.**



Óbitos por local de ocorrência

| | |
|---|---|
| HOSPITAL | 4.567.449 |
| DOMICÍLIO | 1.360.388 |
| OUTROS ESTABEL... | 412.783 |
| VIA PÚBLICA | 275.296 |
| OUTROS | 223.947 |
| IGNORADO | 5.378 |
| ALDEIA INDÍGENA | 540 |

ECD
UFRA

4th Conference on
**Statistics and Data Science**
Salvador, Brazil (online)
December 1-3, 2022