

# Análise Exploratória dos Impactos da COVID-19

Inferência Estatística - PG Biometria

*Discentes: Maicon Centner Germano*

*Fernando Andrade*

*Douglas de Aquino Carrega*

PROF. THOMAS N. VILCHES

Universidade Estadual Paulista "Júlio de Mesquita Filho"  
Campus de Botucatu

## Sumário

<b>1</b>	<b>Data Set</b>	<b>3</b>
<b>2</b>	<b>Objetivo</b>	<b>5</b>
<b>3</b>	<b>Metodologia</b>	<b>6</b>

O dataset "dados-covid-sp" é uma fonte de dados abertos mantida pela SEADE-R (Fundação Sistema Estadual de Análise de Dados do Estado de São Paulo) e disponibilizada no GitHub. Ele contém informações detalhadas sobre a evolução da COVID-19 no estado de São Paulo.

Este conjunto de dados é atualizado regularmente e abrange um período específico, fornecendo uma visão abrangente dos impactos da pandemia na região. Os dados são fornecidos em formato tabular, com cada linha representando um registro individual.

As colunas do dataset incluem informações relevantes, como data de notificação, município, faixa etária, sexo, condição (caso confirmado, óbito, recuperado), além de outras variáveis relacionadas à pandemia. Essas informações permitem uma análise aprofundada dos aspectos demográficos, geográficos e temporais da propagação da doença em São Paulo.

O dataset é uma valiosa fonte de informação para pesquisadores, cientistas de dados e profissionais da saúde interessados em entender melhor a disseminação da COVID-19, identificar padrões e tendências, além de contribuir para a tomada de decisões estratégicas relacionadas ao controle e prevenção da doença no estado de São Paulo.

Ao explorar esse dataset, é possível realizar uma variedade de análises, incluindo análise exploratória de dados, visualizações gráficas, análises temporais e espaciais, além de investigar possíveis correlações com outras variáveis relevantes, como me-

---

didadas de controle da pandemia e taxas de vacinação.

No geral, o dataset "dados-covid-sp" é uma fonte confiável e abrangente que oferece insights valiosos para compreender o impacto da COVID-19 no estado de São Paulo e auxiliar no desenvolvimento de estratégias eficazes para combater a pandemia.

## Objetivo

O objetivo desta análise é explorar a distribuição geográfica dos casos de COVID-19 no estado de São Paulo. O foco será identificar as regiões ou municípios com maior incidência da doença e investigar possíveis correlações com fatores demográficos ou socioeconômicos.

Através da análise de dados, busca-se compreender como a COVID se espalhou geometricamente no estado de São Paulo e quais áreas foram mais afetadas. Além disso, procura-se investigar se existem fatores demográficos tais como densidade populacional, idade média da população, ou socioeconômicos como sendo índice de desenvolvimento humano ou nível de pobreza, que possam estar relacionados com a incidência da doença em diferentes regiões.

Ao identificarmos as áreas com maior incidência de casos, será possível direcionar os esforços e recursos de saúde de forma mais eficaz, implementando medidas de prevenção e controle adequados em tais regiões. Ademais, compreender as possíveis correlações com fatores demográficos e socioeconômicos contribuirá para uma melhor compreensão dos determinantes sociais da saúde relacionados com a COVID-19 em São Paulo.

Essa análise geográfica dos casos de COVID-19 no estado de São Paulo é fundamental para auxiliar no planejamento de ações de saúde pública e na adoção de medidas preventivas específicas de cada região, com o objetivo de mitigar a propagação do vírus e reduzir o impacto da doença na população.

## Metodologia

- Coleta e Preparação dos Dados: Baixamos o dataset "dados-covid-sp" e realizamos a limpeza dos dados.
- Análise Exploratória dos Dados: Exploramos a estrutura dos dados e identificamos variáveis demográficas e socioeconômicas relevantes.
- Visualização Geográfica: Utilizamos a biblioteca ggplot2 para criar mapas interativos que mostram a distribuição geográfica dos casos de COVID-19 em São Paulo.
- Análise de Correlação: Calculamos correlações entre os casos de COVID-19 e as variáveis demográficas ou socioeconômicas.
- Análise Espacial: Aplicamos técnicas de análise espacial para identificar padrões geográficos nos casos de COVID-19.
- Interpretação dos Resultados: Analisamos os resultados obtidos e interpretamos as visualizações e correlações encontradas.