

# SME 0221 - Introdução à Inferência Estatística

## 1º Trabalho Prático - Estatística Descritiva

Grupo: \_\_\_\_\_

Região: \_\_\_\_\_

Aluno 1: \_\_\_\_\_ No.USP: \_\_\_\_\_

Aluno 2: \_\_\_\_\_ No.USP: \_\_\_\_\_

Aluno 3: \_\_\_\_\_ No.USP: \_\_\_\_\_

Aluno 4: \_\_\_\_\_ No.USP: \_\_\_\_\_

Aluno 5: \_\_\_\_\_ No.USP: \_\_\_\_\_

### OBSERVAÇÕES

- Este trabalho é para ser elaborado em grupo de 4 ou 5 pessoas e deverá ser apresentado em forma de relatório (máximo de 12 páginas).
- Todas as análises realizadas deverão ser feitas em R ou Python e o código deverá ser anexado no relatório (páginas adicionais às 12 páginas do relatório).
- Data de entrega do relatório: 25/05/2021.

- 1) Com o objetivo de avaliar a evolução da pandemia causado pelo COVID-19 nos diferentes Estados Brasileiros, considere os dados retirados de: <https://covid.saude.gov.br>.

Cada grupo ficará responsável por uma região:

Grupo 1 - Região Norte;

Grupo 2 - Região Nordeste;

Grupo 3 - Região Centro-Oeste;

Grupo 4 - Região Sudeste; e

Grupo 5 - Região Sul.

Faça uma descrição dos dados por Estado da região em estudo (máximo 3 Estados), considerando as principais métricas da análise descritiva e exploratória de dados. As variáveis essenciais para o estudo são:

- Casos novos por dia de notificação; e
- Óbitos novos por dia de notificação.

Para fins de comparação, utilize essas informações para cada 100.000 habitantes, considerando a seguinte fórmula:

$$\frac{\text{Dados} \times 100000}{\text{População do Estado}}$$

Para elaboração do relatório, com a apresentação e interpretação dos resultados, considere o seguinte guia:

- a) Classificar a variável segundo a natureza;
- b) Visualização gráfica da série de dados diários e acumulados, interpretando-os;
- c) Sumarizar as informações em tabela;
- d) Calcular a média, mediana e quantis;
- e) Calcular a amplitude total, amplitude interquartílica e o desvio-padrão. Ressaltar as informações de maior variabilidade (dentre as variáveis e os Estados);
- f) Visualização gráfica da série de dados diários (histograma e boxplot) para evidências de distribuição e de valores atípicos.