



**XXXIII SIC** SALÃO INICIAÇÃO CIENTÍFICA

<b>Evento</b>	Salão UFRGS 2021: SIC - XXXIII SALÃO DE INICIAÇÃO CIENTÍFICA DA UFRGS
<b>Ano</b>	2021
<b>Local</b>	Virtual
<b>Título</b>	AGRUPAMENTO E CLASSIFICAÇÃO APLICADOS À DADOS DO COVID-19
<b>Autor</b>	LORENZO MORAES COPETTI
<b>Orientador</b>	MARCIO VALK

**Nome do Autor:** Lorenzo Moraes Copetti

**Nome do Orientador:** Marcio Valk

**Instituição de Origem:** Universidade Federal do Rio Grande do Sul (UFRGS)

## **AGRUPAMENTO E CLASSIFICAÇÃO APLICADOS À DADOS DO COVID-19**

Classificação e clusterização de dados são dois desafios que possuem diversas aplicações em muitas áreas. Levando em conta que é um problema que está sendo cada vez mais abordado, principalmente com a popularização de machine learning e outras ferramentas, acabamos por ter vários métodos para utilizar na abordagem de um problema que necessite de classificação ou agrupamento, porém métodos que utilizem de inferência estatística e suas técnicas são poucos. O objetivo do trabalho consiste em explorar a metodologia do pacote do software 'R' chamado 'uclust', que possibilita encontrar clusters a partir de dados temporais, com a característica de utilizar inferência estatística para determinar se os agrupamentos encontrados são estatisticamente significativos. Uma das questões principais investigadas neste trabalho é se dados da evolução do número de casos e do número de mortes por COVID-19 por estado da federação ou por capitais têm dinâmicas similares, isto é, se podem ser agrupados em grupos que apresentam similaridade no comportamento da evolução temporal. Resultados indicam que clusters baseados na evolução do número de casos diferem dos clusters baseados na evolução do números de mortes, o que gera um questionamento importante. Porquê? Além disso, fatores como densidade e idh não apresentam correlação com a dinâmica encontrada nos clusters, o que leva a busca de outros fatores que influenciam diretamente nos desfechos.