

Modelagem Estatística: Uma Jornada pelos *Tidy Models* em R

Tatiana Benaglia, Samara Kiihl
Departamento de Estatística - Unicamp

Resumo

Os *Tidy Models* no ambiente R oferecem uma abordagem sistemática e coerente para análise estatística e modelagem. Esta metodologia, parte integrante do ecossistema *tidyverse*, destaca-se pela sua ênfase em código legível e de fácil manutenção, facilitando a implementação de modelos complexos e a colaboração entre estatísticos e cientistas de dados.

A estrutura modular dos *Tidy Models* permite aos usuários grande flexibilidade na escolha e combinação de diferentes métodos e técnicas de modelagem, adaptando-se facilmente a diversas necessidades e tipos de dados. Além disso, a padronização dos processos de modelagem torna a mesma mais eficiente e reproduzível, aspectos fundamentais na pesquisa científica contemporânea. Em suma, os *Tidy Models* representam uma evolução significativa na prática da modelagem estatística, tornando-a mais acessível, robusta e adaptável às rápidas mudanças e exigências do campo da ciência de dados.

Neste sentido, este minicurso propõe uma introdução ao universo dos *Tidy Models*, utilizando o software R. O objetivo principal é apresentar tais ferramentas e destacar boas práticas no processo de modelagem. Nele, os participantes poderão aprimorar habilidades técnicas, ampliar sua perspectiva sobre as possibilidades modernas de modelagem estatística e fortalecer sua capacidade de conduzir análises de dados eficientes.

Conteúdo Programático

O minicurso será dividido em quatro segmentos principais:

1. **Introdução ao *Tidyverse*:** Visão geral do *Tidyverse* e introdução aos *Tidy Models*, destacando sua aplicabilidade e benefícios sobre outras abordagens em R.
2. **Fundamentos dos *Tidy Models*:** Exploração dos principais pacotes e funções, com orientações sobre a preparação eficaz de dados para a modelagem.
3. **Aplicação Prática:** Sessão interativa com exemplos práticos, incluindo modelos preditivos de regressão e classificação, utilizando conjuntos de dados reais.
4. **Discussão e Perguntas:** Espaço para esclarecimento de dúvidas e discussão sobre aplicações específicas.

Recursos

Este minicurso requer um laboratório com computadores com R e RStudio instalados.

Nível

Básico - Alunos de Graduação/Mestrado: Este tutorial requer apenas um conhecimento básico da linguagem R, tornando-o acessível a um amplo espectro de participantes interessados em aprimorar suas habilidades de modelagem de dados.