

# Introdução ao Machine Learning com R

Introdução ao curso



May de 2022

# Sobre a Curso-R

# Curso-R



**Athos Damiani**

*Curso-R  
Mestrando em  
Inteligência  
Artificial*



**William Amorim**

*Curso-R  
Doutor em  
Estatística*



**Fernando Corrêa**

*Curso-R e ABJ  
Mestrando em  
Estatística*



**Julio Trecenti**

*Curso-R, Terranova,  
ABJ, Confe  
Doutorando em  
Estatística*



**Daniel Falbel**

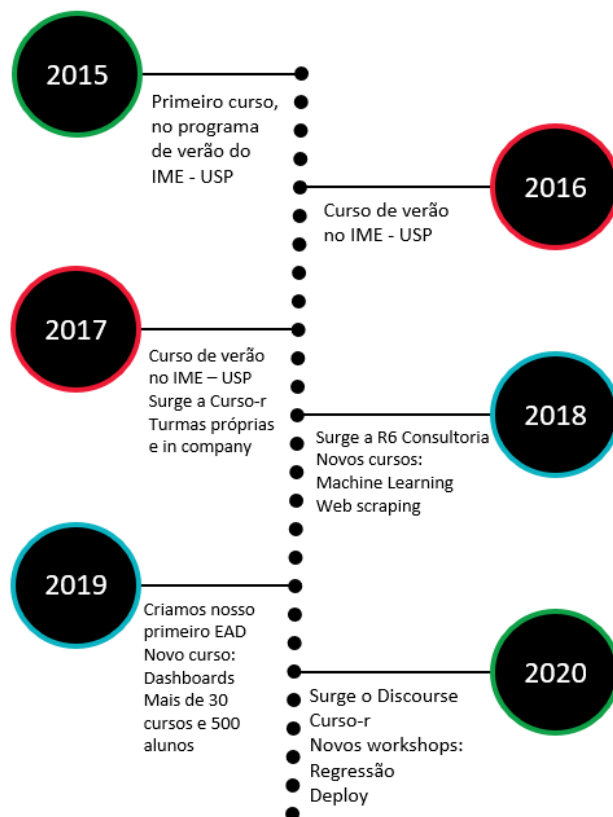
*Curso-R e RStudio  
Bacharel em  
Estatística*



**Caio Lente**

*Curso-R, Terranova,  
ABJ, Mestrando em  
Ciências da  
Computação*

# Linha do tempo



# Nossos cursos

## Programação em R

---

R para Ciência de dados I

R para Ciência de dados II

## Extração de dados

---

Web scraping I

Web scraping II

## Modelagem

---

Regressão Linear

Machine Learning

XGBoost

Deep Learning

## Comunicação e automação

---

Dashboards com R

Deploy

# Sobre o curso

# Dinâmica curso

- As aulas serão (quase totalmente) expositivas.
- Teremos diversos exercícios para serem feitos "em casa".
- Os exercícios não serão avaliações, portanto não terão prazo nem nota. O objetivo deles será gerar dúvidas em vocês, para que os professores possam melhor orientá-las(los) durante o curso.
- O certificado será emitido mediante uma **entrega final** (Kaggle InClass). O prazo da entrega final será de 30 dias após o fim do curso.

# Programa do curso

## Capítulo 1

- Introdução ao Machine Learning
- Tidymodels
- Regressão linear, Regularização
- Overfitting
- Validação cruzada e estratégias

## Capítulo 2

- Regressão logística, regressão vs classificação
- Data prep
- Métricas de performance para classificação: ROC, Acurácia

## Capítulo 3

- Árvores de Decisão
- Random Forest
- Boosting - XGBoost

## Capítulo 4

- Data prep parte II: Matriz X
- Estudo de Caso



# Não falaremos...

- Deploy em produção
- Redes Neurais/Deep Learning
- Séries Temporais
- Provas de Teoremas
- Shiny

# Tirando dúvidas

- **Sintam-se a vontade em fazer perguntas durante a exposição ao vivo..**
- Fora do horário de aula ou monitoria:
  - perguntas gerais sobre o curso deverão ser feitas no Classroom.
  - perguntas sobre R, principalmente as que envolverem código, deverão ser enviadas no [nosso discourse](#).

# Por que usar o discourse?

- Muito melhor para escrever textos que possuem códigos. Com ele, podemos usar o pacote `{reprex}`!
- Saber pesquisar sobre erros e fazer a pergunta certa é essencial para aprender e resolver problemas de programação.
- No discourse, teremos mais pessoas acompanhando e respondendo as dúvidas.
- Em um ambiente aberto, as suas dúvidas vão contribuir com a comunidade.

<https://discourse.curso-r.com/>

# Extras

- Teremos monitoria duas vezes por semana, sempre 30 minutos antes do início das aulas (seg e qui, das 18h30 às 19h)
- A gravação das aulas ficará disponível no Google Classroom por 1 ano após o final do curso.