



# Bootcamp Data Analytics 2025

Challenge Teste de Hipótese



## Desafio 3: Suplementos Alimentares

**Dificuldade: Média.**

Imagine que você trabalha em uma empresa que fabrica suplementos alimentares. O time de produtos quer entender qual das três novas fórmulas de proteína em pó proporciona maior ganho de massa muscular em atletas.

Foi realizada uma pesquisa com vários atletas e coletados os seguintes dados no arquivo desafio\_03\_suplementos.csv:

- **id\_produto:** Código identificador do suplemento (Fórmula 1, 2 ou 3).
- **id\_atleta:** Código identificador do atleta que participou do estudo.
- **ganho\_massa:** Quantidade de massa muscular ganha (em kg) após 8 semanas de uso.
- **idade:** Idade do atleta.
- **frequencia\_treino:** Número médio de treinos semanais do atleta.

Questões a serem respondidas:

1. Faça uma análise exploratória dos dados:
  - Qual é a média e a mediana do ganho de massa para cada suplemento?
  - Como os dados estão distribuídos? Utilize histogramas e boxplots para visualizar as diferenças entre as fórmulas.
  - Há outliers no ganho de massa muscular? O que isso pode indicar?
2. Há diferença estatisticamente significativa no ganho de massa muscular entre as três fórmulas?
3. Existe correlação entre idade e ganho de massa muscular? Atletas mais jovens ou mais velhos têm resultados diferentes?

4. Atletas que treinam mais frequentemente ganham mais massa muscular independentemente do suplemento?
5. Existe alguma correlação entre idade, frequência de treino e a eficácia do suplemento?
6. Qual fórmula você recomendaria para atletas que treinam mais de 5 vezes por semana? Justifique sua escolha.