



# Bootcamp Data Analytics 2025

Challenge Probabilidade e Amostragem



## Desafio 1: Entendendo o consumidor.

Suponha que você trabalha em uma empresa de streaming e, com as mudanças no mercado, a equipe de marketing suspeita que o perfil dos clientes que cancelaram a assinatura mudou. Eles pediram sua ajuda para entender:

- 1 Se o perfil dos clientes que cancelam hoje é diferente do perfil dos clientes que cancelaram anteriormente.

Dica: Fazer um comparativo dos clientes que cancelaram nos últimos 6 meses com os que cancelaram a mais de 24 meses.

- 2 Quais são as principais diferenças entre os perfis, para que possam criar estratégias mais eficazes de retenção.

Para isso, eles fornecem a base de dados **clientes.csv**, que contém informações anonimizadas sobre os clientes ao longo do tempo.

### Dica / Passo a Passo:

1. Analise a distribuição de idade, tempo de assinatura, frequência de uso e região dos clientes que cancelaram e dos novos clientes no período de 24 meses.
2. Plote gráficos para identificar padrões nos clientes que mantêm e cancelam a assinatura.
3. Utilize testes estatísticos para verificar se há uma relação significativa entre tempo de assinatura e taxa de cancelamento.
4. Suponha agora que os dados recebidos são somente uma amostra da sua população total de clientes. Suponha que a sua "população" de clientes verdadeira é de 100000. Qual o tamanho da amostra necessária para chegarmos a conclusões com 95% de confiança e 2% de margem de erro? A base informada é suficiente? E com uma margem de erro de 10%?