

Bootcamp: Cientista de Dados

Desafio

Módulo 1: Análise Estatística de Dados

Objetivos de Ensino

Exercitar os seguintes conceitos trabalhados no Módulo:

1. Análise exploratória, teste de hipóteses, correlação e regressão.

Enunciado

O desafio é constituído de um problema de negócio contendo quinze questões, que serão resolvidas utilizando o R.

Será disponibilizado código R completo para execução de cada questão, a fim de que o aluno possa direcionar seus esforços no entendimento dos métodos estatísticos e interpretações. Mas àquele interessado em aprender o R, é uma excelente oportunidade e pode contar com apoio do professor.

Atividades

Uma empresa que trabalha com vendas em um e-commerce deseja estudar quais características (variáveis) dos seus clientes que impactam no consumo, ou seja, as variáveis que têm relação ao quanto o cliente consome (compra). Para este objetivo, foram selecionados aleatoriamente 29 clientes do banco de dados.

Foram selecionadas seis variáveis para este estudo:

- **Consumo** (variável contínua medida em reais) – É a variável resposta, nos diz qual foi o consumo de cada cliente.

- **Estado Civil** (variável categórica com três níveis) – Nos informa se o cliente é Casado, Divorciado ou Solteiro.
- **Gênero** (variável categórica com dois níveis) – Nos informa se o cliente faz parte do público masculino ou do público feminino.
- **Idade** (variável discreta medida em anos) – Nos informa a idade do cliente.
- **Renda Mensal** (variável contínua medida em reais) – Nos informa a renda mensal do cliente.
- **Possui Imóvel Próprio** (variável categórica com dois níveis) – Nos informa se cliente possui imóvel próprio ou não.