



Bootcamp Data Analytics 2023

Challenge Frequência e Medidas



Desafio 1: Entendendo os estudantes.

Esse desafio tem como objetivo responder algumas perguntas através da análise de dados, aprofundando o conhecimento em frequência e medidas.

Os dados que deverão ser utilizado nesse script foram baixados do kaggle, e podem ser acessados através do link:

<https://www.kaggle.com/spscientist/students-performance-in-exams>

Este conjunto de dados consiste nas notas obtidas pelos alunos em várias disciplinas e suas informações pessoais.

Variáveis:

- **gender** - gênero do aluno
- **race/ethnicity** - raça/etnia dos alunos
- **parental level of education** - grau de escolaridade dos pais
- **lunch** - tipo de almoço
- **test preparation course** - indica se fez curso preparatório
- **math score** - nota na prova de matematica
- **reading score** - nota na prova de leitura
- **writing score** - nota na prova de escrita

Os dados são arquivos em .csv.

Dica / Passo a Passo:

1. Leia a base de dados e faça uma análise inicial, como por exemplo, verificar o tamanho da base de dados (quantidade de linhas e colunas) e o tipo das variáveis.
2. Alguma variável precisa ser alterada ou criada? Faça as alterações se julgar necessária.
3. Verifique a quantidade de alunos em cada grupo racial/étnico.
4. Qual a quantidade de alunos de cada gênero? Existem mais alunos do gênero masculino ou feminino?
5. Analise as notas de matemática dos alunos de forma mais completa possível.

Dica: Identifique o tipo de variável e faça todas análises que você sabe que pode ser aplicado para esse tipo de variável. Por exemplo, se for uma variável quantitativa calcule a média, mediana, quartis, etc. Já se for uma variável qualitativa faça a tabela de frequência, moda, etc.

Faça também análises gráficas das sugestões abaixo ou o que achar pertinente.

- Qual a nota média de matemática dos alunos segundo o grau de escolaridade dos pais?
- Faça a mesma análise para as notas de leitura e escrita.
- Verifique graficamente a relação da nota de leitura e de escrita.

Atividade extra:

Análise mais algum dado que seja interessante avaliar e represente graficamente, através de boxplot, histogramas, barplot dentre outros diversos de livre escolha.