**IMERSÃO DADOS**

Aula 2: Estatísticas, visualização de dados e distribuições

## **PARTICIPE E VÁ MAIS FUNDO**

Seja bem-vindo e bem-vinda à segunda aula da Imersão Dados 3ª Edição!

Nesta aula aprenderemos a plotar e estilizar gráficos com as bibliotecas **seaborn** e **matplotlib**, utilizando recursos de seleção e tratamentos nos dados. Conhecemos o **histograma** para visualizar a média dos genes, utilizando a função **describe** para descrever estatisticamente os dados. Por fim, entendemos o **boxplot** e como tratar os **outliers**.

Para ver os desafios e todo o conteúdo que foi dado em aula siga os seguintes passos:

Acesse o notebook da [Aula 02](https://drive.google.com/file/d/1s5Ugj6D8OpgZztZI_By4dQ72mcsfucBl/view?usp=sharing).

* Na parte superior esquerda, clique em >File, logo depois em >Save a copy in Drive.
* Se você não estiver logado em uma conta **gmail**, um pop-up solicitará que você crie ou faça login em uma conta google.
* Feito o login, uma **cópia** da aula é criada em seu drive (pasta Colab Notebook, criada automaticamente).
* Abra o notebook e boa diversão.

Neste link, temos o Github do projeto onde está localizado a [base de dados e os notebooks das aulas](https://github.com/alura-cursos/imersaodados3).

As **respostas para os desafios da aula 01** você encontra neste [notebook](https://drive.google.com/file/d/1lsJd-K7WmfgnXKMHGkRLMsM3bcYuP-mD/view?usp=sharing), a maioria das soluções utilizaram conceitos ensinados em aula, entretanto é importante salientar que existem diversas maneiras de resolver os exercícios.

Se tiver dificuldades ao utilizar o Colab, acesse este [artigo da Alura](https://www.alura.com.br/artigos/google-colab-o-que-e-e-como-usar) para entendê-lo.

Para dar um mergulho ainda mais profundo, dê uma olhada nas documentações do [seaborn](https://seaborn.pydata.org/" \t "_blank) e [matplotlib](https://matplotlib.org/" \t "_blank), não esquecendo de também conferir os artigos sobre [histogramas](https://www.alura.com.br/artigos/o-que-e-um-histograma?utm_source=gnarus&utm_medium=timeline), a função [describe](https://www.alura.com.br/artigos/ampliando-a-analise-com-describe) e [boxplot](https://www.alura.com.br/artigos/melhorando-a-analise-com-o-boxplot).

Faça os **desafios deixados no final de cada aula** e não deixe de compartilhar seus resultados nas redes sociais marcando os instrutores e adicinando a **#ImersaoDados**

E não esqueça, agora no repositório do nosso [projeto](https://github.com/alura-cursos/imersaodados3) criamos a vitrine vai lá conferir os projetos incríveis que vocês estão desenvolvendo. E para ter seu projeto nesta vitrine, a única coisa que precisa é no repositório do seu projeto adicionar os tópicos alura e imersão dados. Para fazer isso é só clicar na engrenagem de configurações na parte superior direita do seu repositório ao lado de About.

Eu e toda equipe estamos presentes no nosso espaço de discussão dentro do Discord, não deixe de tirar suas dúvidas e participar da comunidade.

Esse projeto foi inspirado em um desafio do [Laboratory innovation science at Harvard](https://lish.harvard.edu/" \t "_blank) disponibilizando os dados em uma competição no [kaggle](https://www.kaggle.com/c/lish-moa" \t "_blank).

## Compartilhe seus resultados!

A gente fica muito feliz ao ver os seus resultados, onde vocês foram além. Para isso, você pode escrever artigos no Linkedin e no Medium com link para seu Colab, texto comentando o que fez de diferente e marcando nós, instrutores. Dá até para fazer um vídeo e encher a gente de orgulho!