**IMERSÃO DADOS**

Aula 3: Correlações, causalidade e relações entre genes.

## **PARTICIPE E VÁ MAIS FUNDO**

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

Nesta aula aprendemos como **cruzar informações dos dados**, analisando as relações de diversas colunas entre sí, iniciaremos relacionando os **dados de experimentos** e posteriormente os tipos **genéticos**. Em seguida utilizamos o **gráfico de dispersão** para investigar a relação entre os genes. Por fim, aprendemos o que é **correlação** discutindo também sobre **causalidade**, nosso objetivo é através da correlação levantar possíveis hipóteses relacionadas aos estímulos biológicos provocados pelos experimentos nas amostras celulares, por fim aprendemos como visualizar a correlação em um mapa de calor.

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

* Acesse o notebook da [Aula 03](https://drive.google.com/file/d/11QdqyySAjhpa_69VvYOfwAP2-fzZ9w4e/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 02** você encontra neste [notebook](https://drive.google.com/file/d/1NuOvnBnAFJQU48drGRiLfQUvkGvA3XT7/view?usp=sharing), a maioria das soluções utilizaram conceitos ensinados em aula, entretanto é importante salientar que existem diversas maneiras de resolver os exercícios.

Não esqueça que hoje às 18h30 teremos a nossa [Live no youtube](https://youtu.be/NMe7DOYDOpc), para falar sobre dados, portfólio e Github. Estaremos esperando você!

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, explore o [site Spurious](https://www.tylervigen.com/spurious-correlations) correlations para vc ver como a correlação muitas vezes não tem relação alguma com a causalidade.

Caso não tenha lido, dê uma olhada nos 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 adicionando 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!