

# Planejamento de Experimentos 2K

Mariana Freitas e Aline Pires

## O que são Experimentos $2^K$ ?

É um tipo de experimento que envolve  $k$  fatores, cada um com dois níveis, sendo amplamente utilizado para estudar o efeito de múltiplos fatores e suas interações em uma resposta de interesse. Como esse tipo de experimento pode ser aplicado a diversos números de fatores, o número total de combinações dos níveis (número de tratamentos) é dado por  $2^K$ .

## Como funciona?

Tendo os dados do experimento  $2K$  coletados e organizados, de acordo com as combinações e os números de repetições, é hora de analisar os efeitos dos fatores principais e de suas interações. Para isso, vamos construir uma tabela ANOVA, na qual vamos utilizar um teste  $F$  para definir se os efeitos são estatisticamente significativos.

# Aplicação

O dataset “Student Lifestyle” reúne informações sobre os hábitos diários e o desempenho acadêmico de 2.000 estudantes, coletadas por meio de uma pesquisa realizada com um Google Form. Abrangendo o período acadêmico de agosto de 2023 a maio de 2024, os dados refletem majoritariamente os estilos de vida de estudantes da Índia.

# Aplicação

O conjunto de dados contém as seguintes colunas:

- ▶ `student_id`: Identificador único de cada estudante.
- ▶ `study_hours_per_day`: Horas diárias dedicadas aos estudos.
- ▶ `extracurricular_hours_per_day`: Horas diárias envolvidas em atividades extracurriculares.
- ▶ `sleep_hours_per_day`: Horas diárias de sono.
- ▶ `social_hours_per_day`: Horas diárias dedicadas à socialização.
- ▶ `physical_activity_hours_per_day`: Horas diárias gastas em atividades físicas.
- ▶ `gpa`: Nota média acumulada (GPA) dos estudantes.
- ▶ `stress_level`: Nível de estresse relatado pelos participantes.

# Aplicação

Aqui, vamos usar apenas as variáveis `study_hours_per_day` e `physical_activity_hours_per_day` e `stress_level` como fatores na nossa análise dos efeitos na variável `gpa`. E, como são as duas primeiras são variáveis contínuas, vamos categorizá-las em dois grupos baseados nos seus quartis.

- ▶ Baixo tempo de estudo:  $\leq 6.5$  horas
- ▶ Alto tempo de estudo  $> 6.5$  horas
- ▶ Baixo tempo de atividade física:  $\leq 2$  horas
- ▶ Alto tempo de atividade física:  $> 2$  horas

## ANOVA

Agora, elaboramos o modelo levando em conta as interações entre os fatores e contruímos a tabela ANOVA abaixo. A tabela ANOVA irá nos auxiliar para comparar as médias dos grupos formado pelos fatores principais e pela interação desses fatores, verificando se existem diferenças significativas entre elas.

```
## Analysis of Variance Table
```

```
##
```

```
## Response: gpa
```

```
##
```

	Df	Sum Sq	Mean Sq	F value	
## stress	1	45.186	45.186	949.8332	< 2
## estudo	1	33.990	33.990	714.4792	< 2
## exercicio	1	1.132	1.132	23.7914	1.1
## stress:estudo	1	2.981	2.981	62.6622	4.0
## stress:exercicio	1	0.115	0.115	2.4173	
## estudo:exercicio	1	0.150	0.150	3.1453	
## stress:estudo:exercicio	1	0.003	0.003	0.0694	
## Residuals	1992	94.765	0.048		

```
## ---
```

# ANOVA

Observando a soma de quadrados (Sum sq), é possível perceber que os fatores stress, horas de estudo diária e a interação entre esses dois fatores são os fatores que explicam uma porção maior da variabilidade do modelo. As horas de exercício diárias explicam uma porção menor da variância e o restante das interações explica bem pouco.

Em seguida, é necessário realizar o teste F para verificar se os efeitos dos fatores principais e das interações são estatisticamente significativos. Para todos os efeitos e interação, teremos o F crítico de 3.846131.



## Teste F

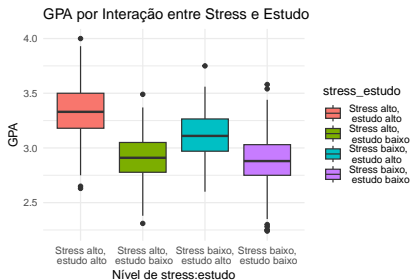
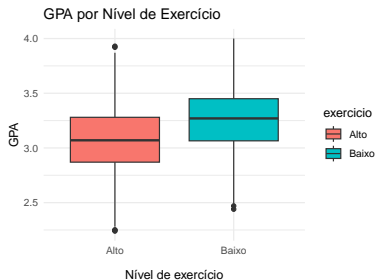
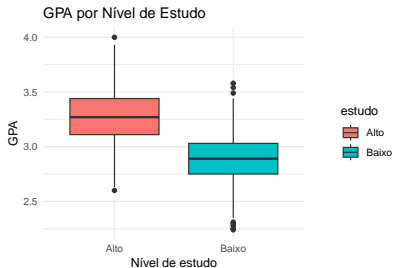
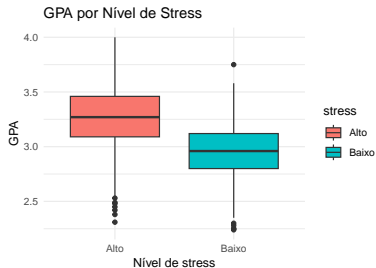
- ▶  $H_0$ : não há efeito significativo do fator/interação.
- ▶  $H_1$ : há efeito significativo do fator/interação.

##	F_Calculado	F_Critico
## stress	949.83318109	3.84
## estudo	714.47915877	3.84
## exercicio	23.79139694	3.84
## stress:estudo	62.66217777	3.84
## stress:exercicio	2.41730862	3.84
## estudo:exercicio	3.14529021	3.84
## stress:estudo:exercicio	0.06938032	3.84

Assim, concluímos que stress, horas diárias de estudo, horas diárias de exercício e a interação entre stress e estudo têm efeito significativo na GPA.

# Análise gráfica

Podemos observar esse efeito graficamente:



## O que falta

- ▶ Testar os pressupostos
- ▶ Teste de turkey ou outro de comparações múltiplas

