

## COMPETIÇÃO DE DATA SCIENCE - FASE 1

### INTRODUÇÃO

Neste primeiro case da competição de Data Science, você será Analista de Dados do time de Marketing da IME Jr. Seu escopo é analisar os dados reais vindos das nossas redes sociais e trazer insights pro time, de forma a agregar visão de negócio para as tomadas de decisão que virão a seguir.

Lembre-se que utilizaremos suas conclusões no nosso time de Marketing real.

**Cause impacto.**

### RECEBÍVEIS

Dados: [arquivo csv](#)

Recomendações: [planilha Google](#)

Exemplo de entregável: [Google Colab](#)

### ENTREGÁVEL: ATÉ 30 DIAS APÓS RECEBIMENTO

Jupyter Notebook - deixando evidente a resposta de cada questão e como você chegou nela.

### RESULTADO

Será divulgado um ranking com todos(as) os(as) competidores(as) com as pontuações desse primeiro desafio.

### DOCUMENTAÇÃO DA TABELA

Os dados dessa tabela são coletados manualmente pelo setor de Analytics do Time de Marketing da IME Jr nas plataformas de Insights do Instagram e do LinkedIn dos perfis da nossa EJ nessas redes sociais. Abaixo, o significado de cada coluna:

**Nome do Post:** Gerado pela concatenação de *Criativo sem espaço com as primeiras letras maiúsculas* + '\_' + *formato* + '\_' + *posicionamento* + '\_' + *plataforma* + '\_' + *data sem '/'*. Exemplo: DiaDoAmigo\_Imagem\_Feed\_Instagram\_12092021.

Cada nome de post é único, não se repetindo mesmo com stories de sequência.

**Data:** Data em que o post foi postado.

**Criativo:** Tema dos posts. Exemplos: MatemáticaESerHumanoALua, NovosMembros, etc.

**Posicionamento:** Feed, stories ou reels.

**Plataforma:** Instagram ou LinkedIn.

**Formato:** Imagem, vídeo ou gif.

**Alcance:** Quantidade de usuários que foram impactados por esse post.

Obs. Note que Alcance nunca pode ser somado, já que um mesmo usuário pode ser impactado por mais de um post. Se um mesmo usuário viu dois posts diferentes da IME Jr, vai ser contado duas vezes, perdendo, portanto, o significado dessa métrica.

**Impressões:** Quantidade de vezes que esse post foi impresso. Note que, diferentemente de Alcance, o número de impressões pode ser somado, já que, se um usuário viu o post mais de uma vez, será contabilizado mais de uma vez. A soma de Impressões significa a quantidade de vezes que nossos posts foram vistos.

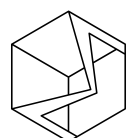
**Saves:** Quantidade de vezes que o post foi salvo.

**Cliques em Link:** Quantidade de vezes que clicaram no link do post. Note que essa métrica não faz sentido pra post no feed do Instagram, já que não é possível colocar links em posts.

**Comentários:** Quantidade de comentários no post.

**Reações:** Quantidade de reações do post. No Instagram, essa métrica significa Likes.

**Compartilhamentos:** Quantidade de compartilhamentos do post.



## QUESTÕES

Utilize os dados recebidos para responder as perguntas abaixo.

Cada pergunta possui uma pontuação específica - não divulgada. Busque responder a maior quantidade de questões que conseguir para alcançar mais pontos.

Procure escrever códigos limpos e utilize as células do notebook para comentários e explicações da resposta.

1. Trate os dados.
2. Qual o post mais reagido?
3. Qual o post mais comentado?
4. Qual o post mais compartilhado?
5. Qual o post mais salvo?
6. Qual o total de saves que nosso Instagram possui?
7. Em quais post alcançamos mais usuários?
8. Em quantos dias tivemos posts?
9. Quantos posts fizemos?
10. Qual o total de usuários alcançados no nosso Instagram?
11. Qual o formato de post com mais compartilhamentos?
12. O post sobre a federação da IME Jr teve melhor performance no Instagram ou no LinkedIn?
13. Qual a média de curtidas no Instagram? E no LinkedIn?
14. Há algum post com dados esquisitos? Quais as suas hipóteses?
15. Qual mês teve mais reações?
16. Quantos e quais posts tiveram mais do que 15 comentários?
17. Quantos posts do feed não possuem comentários? E quantos stories?
18. Quantos posts no feed temos no Instagram desde 29/10/2021?
19. Quantas vezes cada post foi visto por usuário? Qual o post mais visto por usuário?
20. Qual o post com a melhor performance de reações?
21. Qual(is) dia(s) da semana temos mais reações?
22. Qual o total de impressões que os posts com criativo DiaDoDesafio tiveram?
23. Quantos likes teve nosso post da Isabelle Tamarozzi? Ele foi um bom post? Por quê?
24. Qual o nome do nosso melhor story? Por quê?
25. Qual o nome do nosso melhor post? Por quê?
26. Posts que têm mais curtidas têm mais comentários também?
27. Crie uma coluna de nota para cada post. O cálculo da nota fica a seu critério.
28. Tivemos uma melhora ou piora de performance após a divulgação da nova identidade visual? Quais hipóteses você levanta pra isso?
29. Perdemos muito engajamento com posts sequenciais?
30. Qual post deveríamos repetir?
31. Qual formato a gente deve priorizar?
32. Compare nossa performance com outra EJ.
33. Como tomador(a) de decisão, quais seriam suas estratégias pro time de Marketing?

**Cause impacto.**