

## Seazone Data Challenge

### Analista de Dados

O desafio consiste em analisar os dados de “details” e de “rating” de anúncios no Booking, com o intuito de responder uma série de perguntas. Leia todo o documento antes de começar a desenvolver o desafio.

#### Dados:

**desafio\_rating.csv** - Contém dados sobre as avaliações dos imóveis;

- hotel\_id: identificador do anúncio;
- hotel\_name: título do anúncio;
- url: url do anúncio no Booking;
- date: data a ser alugada;
- Comodidades, Limpeza, Conforto, Custo-benefício, Localização, Total, WiFi gratuito: colunas das categorias que receberam as notas;
- number\_of\_ratings: número de avaliações recebidas.

**desafio\_details.csv** - Contém características de cada imóvel/anúncio;

- hotel\_id: identificador do anúncio;
- room\_id: identificador do imóvel/quarto dentro de um anúncio;
- hotel\_name: título do anúncio;
- city\_name: nome da cidade onde está o anúncio;
- accommodation\_type: tipo de acomodação do anúncio;
- room\_name: nome do quarto/imóvel dentro do anúncio;
- room\_surface\_in\_m2: tamanho do quarto/imóvel em metros quadrados;
- room\_facilities: recursos disponíveis no quarto/imóvel;

Seu desafio consiste em analisar as bases de dados disponibilizadas. Responda as perguntas abaixo e, onde possível, embase sua análise com gráficos:

1. Ordene as cidades em ordem crescente de número de listings;
2. Ordene as cidades em ordem decrescente de metros quadrados;
3. Quais cidades têm mais avaliações?
4. Quais cidades têm a maior média de avaliações? E a menor média?
  - a. Existe alguma explicação para isso? Conjecture.
5. Existem correlações entre as características de um anúncio e a sua localização?
  - a. Quais? Explique.
6. Existem relações entre a nota do anúncio e os recursos disponíveis no imóvel?
  - a. Quais? Explique.
7. Existe alguma relação entre a nota recebida e a localização?
  - a. Quais? Explique.
8. O que você pode inferir sobre as notas dos imóveis?
9. Quais são os anúncios que te parecem críticos?
  - a. Explique.

10. Quais outras análises você faria desses dados? Use sua criatividade.
11. Como você projetaria um dashboard para mostrar essas informações?
12. Quais outras informações/dados você relacionaria com essas bases, caso tivesse acesso?
13. (Extra) Com base nesses dados e nos anúncios fornecidos, como você melhoraria as notas?

**Dicas:**

1. Recomendamos a utilização de Python e/ou R para a solução
  - a. Em python, recomendamos a biblioteca Pandas;
  - b. Caso opte pelo Jupyter Notebook, garanta que seu código pode ser executado por terceiros;
2. Dentre os pontos a serem avaliados no desafio, destacamos:
  - a. A qualidade do ReadMe do Github;
  - b. A organização, legibilidade e clareza do código;
  - c. O embasamento das respostas;
  - d. Utilização de gráficos e ferramentas visuais para explicação;
3. Utilize o site do Booking para entender melhor os dados;

**Entrega:**

Para entregar o projeto, você deverá criar um repositório contendo todos os códigos utilizados para a solução no seu perfil pessoal do Github (criá-lo caso ainda não tenha), e deixá-lo em modo público.

As respostas devem ser enviadas por email, com o link para o seu repositório do Github e um PDF no formato de um relatório contendo as análises solicitadas. Ao final do PDF, inclua um feedback sobre o desafio.