



PLATAFORMA DE MODELAGEM AUTOMÁTICA



Vaga Data Scientist



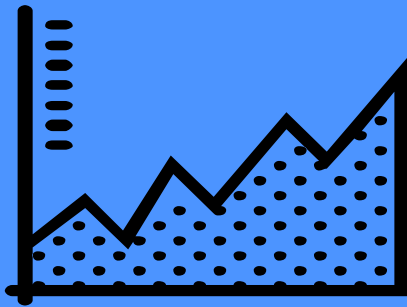
Regras gerais

Para enviar as respostas das questões:

- Gere um notebook/markdown em R ou Python para registrar as respostas e permitir validação dos modelos.
- Se possível, salve o arquivo gerado acima em um repositório (de acesso público) no github e nos envie o link. Caso queira proceder de outra forma, garanta que possamos abrir suas respostas e validar seus resultados.
- Usaremos esse exercício como base para a entrevista, assim poderá elaborar melhor suas respostas durante esse momento.



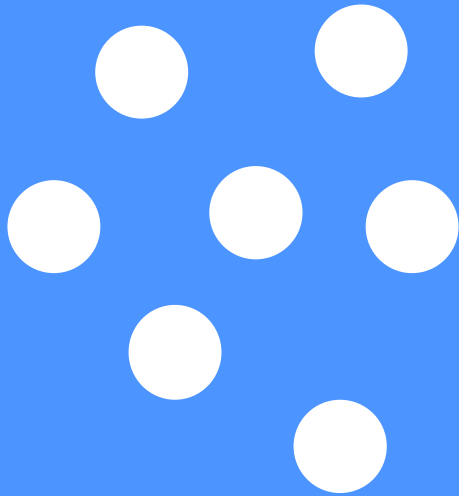
Time Series



A partir da série de inadimplência, disponível em:
<https://dadosabertos.bcb.gov.br/dataset/21082-inadimplencia-da-carteira-de-credito---total> :

1. Projete a taxa de inadimplência dos próximos 12 meses;
2. Como você defenderia essa projeção? Quais são, provavelmente, os principais drivers desses valores?

Classification



Nos mercados de crédito, um problema típico é saber se, dadas as informações de uma pessoa, ela pagará ou não sua dívida. A partir da base disponível em <https://www.kaggle.com/datasets/uciml/default-of-credit-card-clients-dataset?resource=download> gere um modelo preditivo de classificação (variável target: 'default.payment.next.month'), separe 30% da base para teste e traga métricas para validar o modelo em questão.