



---

**Importante:** R e SAS podem ser utilizados sempre que o estudante achar necessário. Em caso de outras linguagens, favor consultar o professor. Qualquer exercício pode ser solicitado para entrega após uma semana de divulgação, e por isso as soluções dos exercícios devem estar sempre organizadas para facilitar a entrega via Aprender. As respostas devem ser claras e incluir código de programação com explicações, caso seja utilizado.

---

1. **Exercício de Leitura:** Ler os Capítulos 1 e 2 de James et. al<sup>1</sup> (2021).
2. Para cada desafio do aprendizado estatístico descrito abaixo, encontre um exemplo de caso real e indique problemas com as técnicas multivariadas tradicionais.
  - (a) *Mining* (mineração,  $n$  muito elevado).
  - (b) *Scalability* (escalabilidade).
  - (c) *High Dimensional Data* (dados em alta dimensão,  $n > p$ ).
  - (d) *Pequenas amostras*.
  - (e) *Curse of Dimensionality* (Problema de dimensionalidade).
3. Fazer os seguintes exercícios do Capítulo 2 de James et al. (2021):
  - (a) Exercício 1.
  - (b) Exercício 2.
  - (c) Exercício 4.
  - (d) Exercício 7: para os itens (b) e (c) faça os cálculos manualmente e utilizando a função `knn` do pacote `class` do R.
    1. Faça um gráfico tridimensional dos pontos  $(x_{i1}, x_{i2}, x_{i3})$ . Apresente os pontos nas cores vermelho ou verde, conforme indicado por  $y_i$ .
    2. Indique se suas predições em (b) e (c) estão de acordo com o gráfico apresentado. Por quê?
  - (e) Exercício 8.
  - (f) Exercício 9.

---

<sup>1</sup>James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning (with Applications in R)*, Springer, 2021.