

## Instruções para trabalho de grupo de métodos estatísticos para a bioinformática 2023\_24

1º. Pesquisar uma base de trabalho que contenha uma variável dependente contínua e várias variáveis independentes de diferentes tipos. Podem, por exemplo, utilizar as bases disponíveis em:

<https://dasl.datadescription.com/datafiles/>

<https://hbiostat.org/data/>

<https://www.kaggle.com/datasets?datasetsOnly=true>

2º. Efetuar uma análise exploratória das variáveis, descrevendo as variáveis envolvidas

3º Construir o modelo de regressão para a variável dependente considerada explorando as técnicas de seleção de variáveis

4º Avaliação das medidas de diagnóstico do modelo adotado

5º Verificação das condições de aplicabilidade

6º Compilar o trabalho para html ou pdf utilizando a função *knit document* do rstudio através do Markdown

7º Colocar o documento html ou pdf e os respetivos ficheiros que lhe deram origem (markdown e dados) num ficheiro .zip com a designação GrupoX\_Reg.zip, no fórum do grupo criado para o efeito no Blackboard