

## TRABALHO DE GRUPO

### Objetivo do Trabalho

Análise estatística de um conjunto de dados, seguindo as seguintes etapas:

1. seleção de uma base de dados;
2. realização de uma análise descritiva e exploratória dos dados;
3. testar se uma dada amostra poderá ter sido extraída de uma população com uma distribuição especificada;
4. avaliação e quantificação da relação entre variáveis qualitativas, bem como modelação e análise da relação entre variáveis quantitativas.

### Fases do Trabalho

1. A primeira fase do trabalho é a recolha da informação. Devem construir/obter uma base de dados que tenha variáveis qualitativas e quantitativas. Podem recolher informação para construir a base de dados, utilizar dados que tenham ou ir buscar uma base de dados que já exista (um possível site para obtenção de base de dados pode ser: UC Irvine Machine Learning Repository).
2. A segunda fase do trabalho consiste em efetuar uma análise descritiva e exploratória dos dados (capítulo 1).
3. A terceira fase do trabalho consiste em sugerir distribuições e testar se os dados podem ser modelados pela distribuição sugerida (ponto 6.1 do capítulo 6).
4. A quarta fase do trabalho consiste em analisar os dados através de métodos estatísticos, como o teste de independência (ponto 6.2 do capítulo 6) e a regressão linear (capítulo 7), de modo a procurar identificar e quantificar relações entre as variáveis.

### Realização do Trabalho

O trabalho deve ser feito em grupo, preferencialmente grupos de 3 alunos da mesma turma de laboratório - inscrever o grupo no Moodle.

### Entrega do Trabalho

Deverá ser feito um relatório descrevendo as etapas do trabalho, os resultados obtidos e interpretação dos mesmos. O relatório deverá ter a estrutura de um artigo científico (ver o modelo disponibilizado no Moodle) e não pode ultrapassar as 20 páginas.

A entrega do trabalho é efetuada no Moodle, até ao dia 30 de junho de 2025 (segunda-feira), e deve entregar o relatório (em pdf), juntamente com o código utilizado para análise dos dados (ficheiro com extensão .R) e o ficheiro dos dados (ficheiro dados.txt).

### Apresentação do Trabalho

As quatro fases do trabalho terão de ser apresentadas na aula de laboratório:

- a primeira fase terá de ser apresentada na aula de laboratório da semana de 24 de março a 28 de março;
- a segunda fase terá de ser apresentada na aula de laboratório da semana de 7 de abril a 11 de abril.
- as datas das seguintes apresentações serão atempadamente divulgadas.

## **Avaliação do Trabalho**

As apresentações do trabalho são obrigatórias e a nota final individual (por elemento do grupo) do trabalho incluirá uma classificação individual (por elemento do grupo) das apresentações. A não realização das apresentações do trabalho significa uma classificação de zero valores no trabalho. A classificação do trabalho será uma nota entre 0 e 20 valores.