

Universidade Eduardo Mondlane  
Faculdade de Ciências  
Departamento de Matemática e Informática  
Programação Estatística  
Teste II

---

O presente teste tem duração de 4 horas. O teste deve ser feito no Rstudio (abra um novo R script). No seu R script, logo no inicio escreva o seu nome completo e no seu número de estudante )(Esta informação deve aparecer como um comentário). Não inclua no seu script nenhum output, apenas coloque o código para responder cada questão. Para cada questão, coloque primeiro a pergunta e de seguida a resposta (código). Salve o seu R script com o nome: PRIMEIRONOME\_ULTIOMONOME\_TESTEI. Após terminar o seu teste, submeta o seu teste **AQUI. PARA EVITAR CONSTRANGIMENTOS, SUMETA O SEU TESTE 5 MINUTOS ANTES DAS 21:00. O LINK PARA SUMISSÃO IRÁ FECHAR PONTUALMENTE AS 21:00**

---

### Como o R script deve estar organizado

```
#####  
## MODULO: PROGRAMACAO ESTATISTICA          ####  
## ACTIVIDADE: TESTE II                     #####  
## NOME DO ESTUDANTE: RACHID MULEIA         ####  
## NUMERO DO ESTUDANTE: 00002023            ####  
#####
```

1. Lembre-se das funções `head` e `tail`. Escreva uma função chamada `middle` que mostre algumas linhas ao redor (aproximadamente) do ‘meio’ do conjunto de dados. Dica: use `nrow`, `print` e possivelmente `floor`. **(2 valores)**
2. Considere os dados no ficheiro `CORDEX_DATA`. O ficheiro contém dados sobre humidade relativa (`rh`), precipitação (`pr`) e temperatura (`tas`). Os dados estão apresentados para três cenários de emissão, `RPC26`, `RPC45` e `RPC85`. **(3 valores)**
  - Escreva um código para ler os dados de humidade relativa. Junte todos os cenários de emissão numa única base de dados. Na base final crie uma variável que indica o cenário de emissão. use a função `list.files()` para imprimir os nomes dos ficheiros e a função `lapply` para ler todos os ficheiros, e a função `do.call` para fazer a junção das bases.
  - Escreva uma função generalizada que permitira ler todos os demais ficheiros, precipitação e temperatura. Deve testar a função.
3. Os dados no exercício 2 são diários. Agrupe os dados de forma mensal. Aplique as funções da livreria `dplyr`, `mutate`, `group_by` e `summarise`. use a função `month` da livreria `lubridate` para extrair o mês e a função `year` para extrair o ano. Aplique a função `lapply` para fazer o exercício para todas as bases (humidade relativa, precipitação e temperatura). Para a humidade relativa e temperatura vão calcular a média, para a precipitação irão calcular a soma ao longo do mês. A primeira coluna corresponde a data de extracção dos dados no formato `ano_mes_dia`. Use a função `ymd` da livreria `lubridate` para converter os dados para data. **(3 valores)**
4. Considere a pasta `AULAS`. Este ficheiro contém todos os slides das aulas teóricas. Escreva uma função que lê estes ficheiros todos e junta todos os ficheiros num só ficheiro. Se na pasta `AULAS` existir apenas um ficheiro pdf, a função não deve fazer nenhuma leitura. Se existir mais que dois deve ler todos os ficheiros e juntar num único ficheiro (Neste caso a função deve devolver um ficheiro pdf, que será gravado na pasta `AULAS`). Para este exercício irá usar a livreria `pdftools`. Instale esta livreria incluindo dependências. Use para este exercício as funções, `pdf_length`, `pdf_subset`,

- e `pdf_combine()` (estas funções são apenas para manusear a leitura de ficheiros pdf). Aplique, preferencialmente, a função `lapply`, mas pode aplicar ciclos se quiser. **(3 valores)**
5. Escreva uma função que cria uma `data frame` cujas as colunas são variáveis aleatórias de uma distribuição probabilidade. A função deve receber três argumentos, o número de linhas da `data frame`, número de colunas e a distribuição da qual pretendemos simular. Estes são argumentos especificados por você(criador da função). Importa observa que as funções que irá usar simular os dados tem outros argumentos. Então, deverá importar esses argumentos usando `...` (Vimos este operador.). Considere apenas três distribuições, Normal, Poisson e Bernoulli.**(3 valores)**
6. Considere o ficheiro `Zambezia_Projeccoes.xls`. Este ficheiro contem dados da população para a província da Zambézia, para vários anos. Escreva um ciclo `while` que faz a leitura das planilhas que estão neste ficheiro. O ciclo deve ler dados até 2035. Ao ler os dados, considere apenas os dados totais da província para as idades de 0 à 5 anos. Para este exercício vai usar a função `import` da livraria `rio`. Visite também a função `read_xls`, pois esta tem argumentos que foram passados para função `import`. **(3 valores)**
- Escreve um código que junta todos os dados lidos numa única base de dados
  - Calcule o numero total da população para cada ano.
7. Escreva uma função que recebe uma matriz a média, desvio padrão, máximo e mínimo para cada coluna ou linha. Coloque um argumento na função que permita o usuário escolher a dimensão para qual quer calcular estas estatísticas. Se conseguir, permita que o usuario tenha a possibilidade de devolver para as duas dimensões.**(3 valores)**
8. Escreva uma função que verifique se um valor está acima de um valor limite. A função deve ter dois parâmetros: `x` que é o valor numérico a ser verificado e `limiar` que é o limite numérico. Faça com que a função retorne um valor lógico, `TRUE` se o valor estiver acima do limite e `FALSE` se for igual ou abaixo do limite.**(2 valores)**