

PROJETO II

Atividade em R

Estatística para Computação 2019.1 - et586cc

1. Informações gerais: O segundo projeto de Estatística para Computação 2019.1 será realizado na plataforma **R Studio**, a qual foi abordada em sala de aula. O projeto vale 70% da 3ª nota da cadeira, sendo os outros 30% preenchidos pelo projeto I de seminários sobre a aplicação da estatística na computação.

2. O projeto: Deve ser feito em grupos de no **máximo 3 alunos**. Não aceitaremos projetos com mais integrantes. Será disponibilizada uma [planilha](#) onde deverá ser colocado o(s) nome(s) do(s) integrante(s) juntamente com seu(s) login(s).

O projeto consiste em elaborar um script que responda todas as questões pedidas a seguir referentes a [planilha](#) de “Episódios de Game of Thrones”. **A planilha deve ser baixada como arquivo .csv para ser carregada no R como foi ensinado em sala de aula.** A nota do projeto será proporcional a quantidade de questões solucionadas corretamente.

É importante:

- Colocar o nome dos integrantes no início do script na forma de comentário.

Ex: *# Rebeca Oliveira Silva (ros4) -----*

- Comentar antes que questão será resolvida no script, para facilitar a correção. Também é válido explicar seu raciocínio através de comentários.

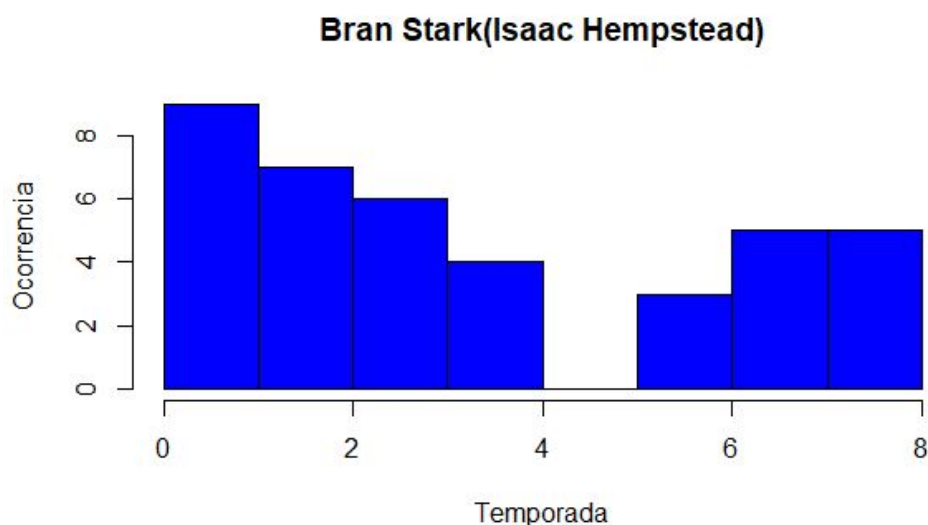
Ex: *# início da questão (1) -----*

nessa questão, o problema da média foi resolvido da seguinte forma (...)

- É necessário imprimir a resposta encontrada a cada questão resolvida. Se a questão pede a média de uma certa amostra, após realizar a função, você deve imprimir o resultado dessa função.
- **Não** é permitido copiar totalmente ou parcialmente a resolução de outras equipes. Caso ocorra, ambas serão penalizadas.

3. As questões: A seguir as questões que devem ser resolvidas:

- (1) Descarregue o arquivo .csv da planilha e imprima o dataframe obtido exatamente do jeito que ele se encontra.
- (2) Encontre a média, o desvio padrão e a moda das notas dos episódios.
- (3) Encontre a média, o desvio padrão e a mediana da audiência dos episódios.
- (4) Faça uma função que retorna apenas os nomes dos episódios que possuem notas maiores ou iguais a nove (9).
- (5) Faça uma função que retorna o nome dos episódios com menor e maior notas, nessa ordem para cada uma das temporadas. Por fim, faça um dataframe com cada episódio encontrado com as colunas TÍTULO, NOTA, TEMPORADA ordenados de forma crescente por temporada (de 1 até 8).
- (6) Faça uma função que retorne qual a temporada com o menor desvio padrão na audiência.
- (7) Faça uma função que retorne a média das notas dos episódios em que Brienne of Tarth(Gwendoline Christie) participa.
- (8) Faça uma função que retorne uma lista com os personagens que só apareceram em um único episódio na quarta (4ª) temporada.
- (9) Faça uma função que dado o nome de um personagem, cria um histograma onde mostra a frequência de aparição desse personagem a cada temporada. Não esqueça de dar um título e fazer ele de forma colorida, facilitando a visualização. Um exemplo para o personagem Bran Stark(Isaac Hempstead) seria:



4. A entrega: A entrega e apresentação do projeto será feita nos dias 7 e 12 de Junho. Cada grupo terá seu dia de apresentar, não precisando ir no outro dia (por exemplo, se você apresentar dia 12 não precisa aparecer no dia 7). As datas de apresentação de cada grupo ficarão disponíveis na planilha, **mas TODOS os arquivos(.R) dos projetos de todos os grupos devem ser enviados até às 13h do dia 7 de Junho por email com o assunto [PROJETO 2 ESTATÍSTICA]** para os emails rfri@cin.ufpe.br e ros4@cin.ufpe.br. A apresentação será feita para um monitor que esteja disponível. Entregas e apresentações depois dessas datas irão valer menos. Qualquer dúvida entre em contato com os monitores.