

## **1ª Lista de Exercícios**

### **Ciência dos Dados – 2º/2024**

Prof. Leonardo Costa Ribeiro

Alun@:\_\_\_\_\_

### **Introdução**

1. Explique o que é uma linguagem de programação. Não deixe de explicitar o papel do compilador e do interpretador.
2. De forma esquemática, apresente como é a estrutura dos comandos IF e FOR e como se define uma função em R e Python.
3. Usando R e Python, defina sua própria função (portanto, não use alguma outra função que já esteja definida na linguagem para esse fim) na qual se passe como entrada uma lista com 10 números e ela retorne o valor médio da lista.
4. Usando R e Python, defina sua própria função (portanto, não use alguma outra função que já esteja definida na linguagem para esse fim) na qual se passe como entrada uma lista com 10 números e ela retorne o maior número da lista.
5. Usando R ou Python, defina sua própria função (portanto, não use alguma outra função que já esteja definida na linguagem para esse fim) que gere os 50 primeiros valores da sequência de Fibonacci. Essa sequência tem várias aplicações em matemática e é encontrada codificada em diversos fenômenos da natureza. Ela começa com o par de números 0 e 1, e os seguintes são sempre a soma dos dois números anteriores.
6. A partir da sequência de Fibonacci que você construiu no exercício anterior, calcule, usando R ou Python, a razão entre dois números consecutivos à medida que a série vai crescendo. Essa razão parece convergir para um mesmo valor? Qual valor é esse?
7. Explique os conceitos de classe e instância associados com uma linguagem de programação orientada a objeto.
8. Usando Python, crie uma classe (que não tenhamos visto em aula) explicando, também, que tipo de objeto ela irá instanciar. Explique quais serão seus atributos e, pelo menos, um dos seus métodos. Apresente o código fonte que você usou para criá-la.

9. O DNA é composto por 4 bases: Guanina ('G'), Citosina ('C'), Adenina ('A'), e Timina ('T'). O RNA é composto por 4 bases, mas em vez de Timina ('T'), tem Uralina ('U'). Crie uma função para receber uma sequência de DNA de qualquer tamanho e converter para RNA. Por exemplo, pegar um texto na forma GCATATAC e retornar a conversão para RNA, que seria GCAUAUAC.

### **DataFrame**

1. Usando a planilha com os dados de PIB per capita do Maddison que está disponível no Moodle, crie um Dataframe contendo a razão em o PIB per capita de cada país e o dos EUA.
2. Encontre a lista de países que, pelo menos em um ano, teve a razão do exercício anterior acima de 0,5.
3. Calcule a razão média para cada país do exercício anterior para o período que tiveram a razão acima de 0,5.
4. Usando a planilha com a amostra de publicações no Google Scholar disponibilizada no Moodle, gere o quantitativo de publicações por tipo de documento e ano de publicação, por ordem decrescente do número de publicações.
5. Calcule também o número médio de citações das publicações por tipo de documento.
6. Usando a base REGIC.xlsx disponibilizada no Moodle, categorize a população municipal em cinco (5) categorias e conte o número de municípios em cada categoria para os municípios da região Sul do Brasil.