



UNIVERSIDADE ESTADUAL DE CAMPINAS
LABORATÓRIO CENTRAL DE TECNOLOGIAS DE ALTO
DESEMPENHO



Roteiro – Aula Prática:

Conceitos Básicos de R

Julho/2015

Sumário

Prática 1: Tipos de dados, operadores, atribuição de valores e variáveis.....	2
Prática 2: Comparação de valores.	3
Prática 3: Tabela Verdade e Condicional.	4
Prática 4: Vector e suas funções; Iterações.	5
Prática 5: Matrix.	6
Prática 7: Dataframe.	8
Prática 8: Manipulação de arquivos.	9
Em caso de dúvidas.	10

Prática 1: Tipos de dados, operadores, atribuição de valores e variáveis.

Encontrar a solução para a seguinte equação:

$$x = \frac{(100^2 + \sqrt{4000} - \log_2(345)).250}{230}$$

Dicas:

Você pode utilizar:

- ^: para exponencial;
- sqrt: para raiz quadrada;
- *: para multiplicação
- /: para divisão.

Prática 2: Comparação de valores.

Sabendo que:

$$x = 15;$$

$$y = \log_{10}(235);$$

$$z = (25)^3$$

Quais expressões abaixo são verdadeiras ou falsas?

a) $(x + z) \geq y$?

b) $(y + z) \leq x^2$?

c) $\log_2(z) == (x + y)$?

Prática 3: Tabela Verdade e Condicional.

Sabendo que:

$$x = 26^3 - (12 \cdot 10)^2;$$

$$y = \log_2(234.5) + (23 \cdot 16)/100;$$

$$z = 100 - 9^2 + 12^3$$

Quais expressões são verdadeiras?

a) Se $((x > y) \text{ AND } (z \leq x))$ {

$x = 1$

}

senão {

$x = 0$

}?

b) Se $((x + z) > 10 \text{ OR } (x + y) == (z - x))$ {

$x = 1$

}

senão{

$x = 0$

}?

c) Se $((z \geq x) \text{ AND } (y \geq x))$ {

$x = 1$

}

senão{

$x = 0$

}?

Prática 4: Vector e suas funções; Iterações.

Sabendo que:

$a = c(10, 2, 3, 14, 34, 200, 40)$

$b = c(0, 23, 65, 93, 72, 10, 11, 90)$

Calcule:

- a) Quais os tamanhos de cada *vector*?
- b) Quais os valores das somas dos primeiros 4 elementos de cada *vector*?
- c) E se somar os primeiros 4 elementos de *a* com os 4 últimos elementos de *b*, qual seria o valor?

Dicas:

- Comando "*length*" para saber o tamanho do vector.
- Comando "*sum*" pode ser utilizado para calcular a soma de vários elementos.
- Pode também utilizar iterações como "*for*" ou "*while*".

Prática 5: Matrix.

Considere a matriz de 3 colunas e 7 linhas:

```
m = cbind( a = c(100,122,33,54,64,20,4),  
          b = c(40,123,55,67,21,18,12),  
          c = c(12,21,44,23,87,50,123))
```

Calcule:

- A média e a mediana dos elementos da coluna “a”, “b” e “c” e concatene em uma nova matriz.

Dicas:

- Utilize os comandos *mean* e *median* para os cálculos de média e mediana.
- Para acessar uma coluna da matriz, lembre-se que:

matriz [linha, coluna]

Prática 6: List.

Faça listas contendo dados de 2 experimentos:

- a) Lista 1 = experimento da amostra 1 do grupo “controle” com 3 réplicas biológicas e tempos de 24 e 72 horas;
- b) Lista 2 = experimento da amostra 2 do grupo “teste” com somente uma réplica e tempos de 1, 2 e 3 horas.

Prática 7: Dataframe.

Considere:

```
genes = c("a","b","c","d");  
read_count_amostra1 = c(100,120,0,12);  
read_count_amostra2 = c(12,90,20,0);  
read_count_amostra3 = c(23,45,67,33)
```

Criar um *dataframe* que contenha os dados acima, também calculando os valores de soma dos *read counts* para cada gene.

Dica (opcional):

Caso queira colocar os genes como nomes das linhas em um *dataframe*, passe como parâmetro dentro do comando “*row.names = genes*”.

Prática 8: Manipulação de arquivos.

Antes de começar:

- Criar um *dataframe*:

```
genes=c("GAPDH", "COL12A1", "ACTB", "ANXA2", "SOD2", "MMP2", "FSTL1", "IGFBP4", "
AHNAK", "COL8A1")
rc_sample1 = c(1200,3425,23987,567,789,342,555,678,444,231)
rc_sample2 = c(5050,978,349,954,565,4787,344,66,737,100)
rc_sample3 = c(230,4646,5858,3432,567,8382,453,2020,100,102)

df = data.frame(rc_sample1,rc_sample2,rc_sample3,row.names=genes)
```

- Salvar o *dataframe* em um arquivo:

```
write.table(df, file = "arquivo_pratica8.txt", sep="\t")
```

Depois:

- Abrir o arquivo “*arquivo_pratica8.txt*”, calcular a média e a mediana por gene.

- Criar outro *dataframe* contendo:

- Todos os read counts por gene;
- Todas as médias por gene;
- Todas as medianas por gene.

- Salvar o novo *dataframe* como “*arquivo2_pratica8.txt*” separado por tabulação.

Dicas:

- Calcular a média usando iteração:

```
media[i] = mean(c(x$rc_sample1[i],x$rc_sample2[i],x$rc_sample3[i]))
```

- Calcular mediana usando iteração:

```
mediana[i] = median(c(x$rc_sample1[i],x$rc_sample2[i],x$rc_sample3[i]))
```

Em caso de dúvidas.

Use “?” antes do comando no R. Abrirá um tutorial contendo o conteúdo do comando, parâmetros que devem ser passados e até exemplos de como utilizar o comando.

Caso não encontre a solução para sua dúvida, há ótimos sites contendo tutoriais de R como:

- The R Project for Statistical Computing (<http://www.r-project.org/>)
- Quick-R (<http://www.statmethods.net/interface/index.html>)
- Tutorial TryR da CodeSchool (<http://tryr.codeschool.com/>)
- Tutorial de “Introduction to R” da DataCamp
(<https://www.datacamp.com/courses/introduction-to-r/>)