



ESCOLA POLITÉCNICA DA UNIVERSIDADE DE SÃO PAULO
DEPARTAMENTO DE ENGENHARIA MECÂNICA

PME3463 - Introdução à Qualidade

Prof. Dr. Walter Ponge-Ferreira

L1 - Introdução ao R

1 Exercício

Usando o R como uma calculadora, realize as seguintes operações:

- a) $1 + 2(3 + 4)$
- b) $4^3 + 3^{2+1}$
- c) $\sqrt{(4 + 3)(2 + 1)}$
- d) $\left(\frac{1+2}{3+4}\right)$

2 Exercício

Crie um vetor com os seguintes dados utilizando a função $c()$:

2 3 5 7 11 13 17 19

Use a função $length()$ para verificar o tamanho do vetor.

3 Exercício

Os preços do MINI Coopers variam entre vários revendedores. A relação a seguir apresenta a relação do preços em milhares de dólares desse veículo:

15.9 21.4 19.9 21.9 20.0 16.5 17.9 17.5

- a) Qual é o menor preço? Qual é o maior?
- b) Qual é o preço médio?
- c) Qual é a maior diferença de preços?

Realize as operações e repita usando as funções do R.

4 Exercício

Segundo o *New York Times* as vendas mensais do *Hummer - H2* nos Estados Unidos no ano de 2002 foram:

Janeiro	Fevereiro	Março	Abril	Maior	Junho
2700	2600	3050	2900	3000	2500
Julho	Agosto	Setembro	Outubro	Novembro	Dezembro
2600	3000	2800	3200	2800	3400

Pede-se:

- a) Carregue esses dados para variável *H2* do *Software R*.
- b) Use a função *cumsum()* para encontrar a soma cumulativa mensal das vendas de 2002. Qual é o volume total vendido no ano?
- c) Use a função *diff()* para estudar o incremento mensal nas vendas e determine o mês com a maior redução nas vendas.

5 Exercício

São definidos os seguintes vetores:

$$\begin{aligned}x &= c(1, 3, 5, 7, 9) \\ y &= c(2, 3, 5, 7, 11, 13)\end{aligned}$$

Adivinhe qual será o resultados das seguintes operações:

- a) $x + 1$
- b) $y * 2$
- c) $\text{length}(x)$ e $\text{length}(y)$
- d) $x + y$
- e) $\text{sum}(x > 5)$ e $\text{sum}(x[x > 5])$
- f) $\text{sum}(x > 5 | x < 3)$
- g) $y[3]$
- h) $y[-3]$
- i) $y[x]$ O que é *NA*?
- j) $y[y > -7]$

6 Exercício

Um processo de fabricação de chapas de aço produziu os seguintes resultados:

Onde os valores apresentados são as espessuras das chapas produzidas em dois dias consecutivos, *Dia 1* e *Dia 2*.

Pede-se:

- a) importar os dados para dois vetores de dados, *dia1* e *dia2*, correspondendo aos dados de espessura das chapas produzidas nos dois dias consecutivos. Depois agrupar os dados na forma de uma tabela, *data.frame*, a ser denominada de "placas";

Espessura	Dia	Espessura	Dia
0,821	Dia 1	0,678	Dia 2
0,846	Dia 1	0,742	Dia 2
0,892	Dia 1	0,684	Dia 2
0,750	Dia 1	0,766	Dia 2
0,773	Dia 1	0,785	Dia 2
0,956	Dia 1	0,759	Dia 2
0,840	Dia 1	0,708	Dia 2
0,913	Dia 1	0,789	Dia 2
0,737	Dia 1	0,732	Dia 2
0,793	Dia 1	0,804	Dia 2
0,872	Dia 1	0,758	Dia 2

Tabela 1: Dados das espessuras da chapas

- b) contar quantos dados estão disponíveis;
- c) determinar a maior e a menor espessura produzida em cada dia de produção;
- d) determinar a espessura média produzida em cada dia;
- e) determinar a espessura média produzida nos dois dias.

7 Exercício

Baixar o arquivo de dados do Excel "*Dados dos alunos.xlsx*", disponível no Moodle da disciplina.

Pede-se:

- a) converter o arquivo para o formato *csv*;
- b) importar os dados para uma sessão do R;
- c) visualizar os dados;
- d) verificar se existem dados não disponíveis;
- e) determinar os valores máximos e mínimos das idades, alturas e pesos;
- f) determinar média das idades, alturas e pesos dos alunos.

8 Exercício

Carregue o conjunto de dados "*primes*" da biblioteca *UsingR* para a área de trabalho do R. Esta é uma relação dos números primos entre 1 e 2003.

Pede-se:

- a) quantos número primos existem nessa faixa?
- b) quantos números primos existem na faixa de 1 a 100?
- c) quantos números primos existem na faixa de 100 a 1000?
- d) quanto vale a soma de todos os números primos disponíveis no conjunto?

9 Exercício

Carregue o conjunto de dados "*nym.2020*" da biblioteca *UsingR*. Essa tabela contém as marcas de tempo de vários participantes da maratona de Nova York de 2002.

Pede-se:

- a) Quantos valores estão disponíveis na tabela?
- b) Qual o tempo do último colocado em minutos? Converta o resultado para horas e minutos.
- c) Qual o tempo do primeiro colocado em minutos? Converta o resultado para horas e minutos.
- d) Organize os dados em ordem crescente de colocação.
- e) Quanto tempo demorou o 100º corredor para completar a prova?
- f) Quantos homens e quantas mulheres participaram da prova?
- g) Qual é a idade média dos homens? E das mulheres?
- h) Quanto tempo demorou o corredor mais velho para completar a prova? E o mais novo?
- i) Qual a cidade de origem do maior número de atletas?
- j) Exporte o conjunto de dados para um arquivo de texto separado por ponto e vírgula (*csv*) e importe o arquivo para o Excel.

10 Exercício

Carregue o conjunto de dados *mandms* da biblioteca *UsingR*. Essa tabela contém a distribuição de cores previstas para os pacotes de *M&M* em função dos diversos tipos de embalagens.

Pede-se:

- a) Qual das embalagens não contém uma das seis cores?
- b) Qual embalagem apresenta uma distribuição equilibrada de cores?
- c) Qual embalagem contém uma cor que supera a soma de todas as outras cores? Qual é essa cor?

Fonte:

Verzani, John. **Using R for Introductory Statistics**, Chapman & Hall, CRC Press, Florida, 2010.