

# BIOESTATÍSTICA

M.I. Eng. Biomédica

2015-2016

Aula Prática 4



# Organização de bases de dados

- Organização em tabela de valores
  - Folha de cálculo;
- Variáveis em colunas
  - Variáveis nominais ou ordinais codificadas com valores numéricos;
  - Variáveis quantitativas nas mesmas unidades;
- Sujeitos em linhas

# Base de dados

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Identificaçã	sexo	Idade	Raça	Profissão	Grau escolaridade	Est. Civil	Perimetro Abd	Ant familiares	inicio doença	DM	Hipercolesterolemia	Hipertrigliceridemia	HTA	Apneia Sono	Hipotiroidis
2	1	1	36	1	2	4	1	127	1	1	2	2	1	1	2	2
3	2	2	47	1	11	4	2	108	2	2	1	1	1	2	2	2
4	3	1	50	1	2	4	2	115	1	2	2	1	2	1	2	2
5	4	1	32	1	1	2	2	100	2	2	2	2	2	2	2	2
6	5	1	68	1	3	4	2	118	1	2	2	1	1	1	2	2
7	6	1	56	1	12	3	2	120	2	2	2	1	2	1	1	2
8	7	1	75	1	3	1	2	105	2	2	2	1	1	2	2	2
9	8	1	65	1	1	1	2	123	1	2	1	1	1	1	2	2
10	9	1	52	1	1	2	2	104	1	2	2	2	2	1	2	2
11	10	1	40	1	13	3	2	105	1	1	2	1	2	2	2	2
12	11	1	52	1	14	3	2	114	2	2	2	2	2	1	2	2
13	12	2	56	1	15	4	1	160	1	2	1	1	1	1	1	2
14	13	1	54	1	1	1	4	114	1	2	2	2	2	1	2	2
15	14	1	35	1	1	2	2	112	2	2	2	2	2	2	2	2
16	15	1	38	1	6	2	2	103	1	2	2	2	2	2	2	2
17	16	1	60	1	3	1	2	117	2	2	2	1	2	1	2	2
18	17	1	46	1	16	4	2	103	1	2	2	2	2	2	2	2
19	18	1	44	1	6	3	2	119	1	1	2	1	2	2	2	2
20	19	2	51	1	17	2	2	116	1	2	2	2	1	1	1	2
21	20	1	34	1	4	3	4	99	1	1	2	2	2	2	2	2
22	21	1	38	1	18	2	2	103	1	2	2	2	2	2	2	2
23	22	1	31	1	6	3	2	127	1	1	2	2	2	2	2	2
24	23	1	64	1	3	1	2	101	1	2	1	1	2	1	2	2
25	24	1	49	1	1	2	2	129	2	1	2	2	2	2	2	2
26	25	2	59	1	3	2	2	123	2	2	1	1	2	1	1	2
27	26	1	18	1	8	3	1	109	1	1	2	2	2	2	2	2
28	27	1	33	1	4	3	1	132	1	1	2	1	1	2	2	2
29	28	1	41	1	10	2	2	112	1	2	2	1	2	1	2	2
30	29	1	35	1	6	2	1	119	2	2	2	1	1	1	2	2
31	30	1	30	1	8	4	1	105	1	1	2	1	2	2	2	2
32	31	2	52	1	3	2	2	113	2	2	1	1	1	1	2	2
33	32	1	55	1	5	3	2	98	1	2	2	1	2	2	2	2

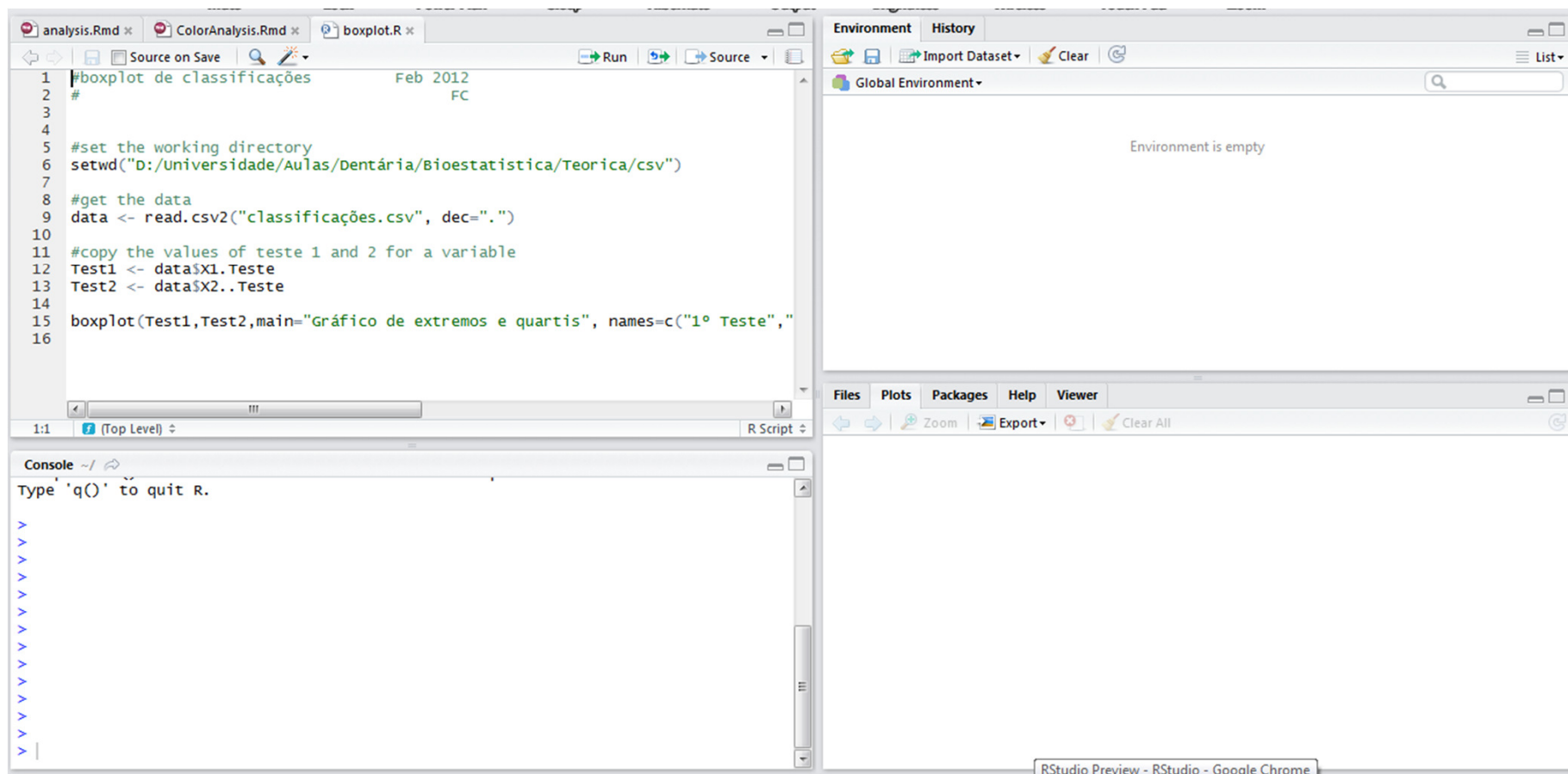
# Scripts em R

- Abrir o excel e criar o ficheiro teste.csv com a informação seguinte

Data

id	density	body_fat	age	weight	height	neck	chest	abdomen	hip	thigh	knee	ankle	biceps	forearm	wrist
1	1,0708	12,3	23	154,25	67,75	36,2	93,1	85,2	94,5	59	37,3	21,9	32	27,4	17,1
2	1,0853	6,1	22	173,25	72,25	38,5	93,6	83	98,7	58,7	37,3	23,4	30,5	28,9	18,2
3	1,0414	25,3	22	154	66,25	34	95,8	87,9	99,2	59,6	38,9	24	28,8	25,2	16,6
4	1,0751	10,4	26	184,75	72,25	37,4	101,8	86,4	101,2	60,1	37,3	22,8	32,4	29,4	18,2
5	1,0340	28,7	24	184,25	71,25	34,4	97,3	100	101,9	63,2	42,2	24	32,2	27,7	17,7
6	1,0502	20,9	24	210,25	74,75	39	104,5	94,4	107,8	66	42	25,6	35,7	30,6	18,8
7	1,0549	19,2	26	181	69,75	36,4	105,1	90,7	100,3	58,4	38,3	22,9	31,9	27,8	17,7
8	1,0704	12,4	25	176	72,5	37,8	99,6	88,5	97,1	60	39,4	23,2	30,5	29	18,8
9	1,0900	4,1	25	191	74	38,1	100,9	82,5	99,9	62,9	38,3	23,8	35,9	31,1	18,2
10	1,0722	11,7	23	198,25	73,5	42,1	99,6	88,6	104,1	63,1	41,7	25	35,6	30	19,2

# Scripts em R





## *Scripts* em R

- Abrir um ficheiro de dados

```
#set the working directory  
setwd("D:/Bioestatística/Teórica/csv")
```

```
#get the data  
data <- read.csv2("classificações.csv", dec=".")
```

# Scripts em R

- Abrir um ficheiro de dados

```
read.table(file, header = FALSE, sep = "", quote = "\"'",  
           dec = ".", numerals = c("allow.loss", "warn.loss", "no.loss"),  
           row.names, col.names, as.is = !stringsAsFactors,  
           na.strings = "NA", colClasses = NA, nrows = -1,  
           skip = 0, check.names = TRUE, fill = !blank.lines.skip,  
           strip.white = FALSE, blank.lines.skip = TRUE,  
           comment.char = "#",  
           allowEscapes = FALSE, flush = FALSE,  
           stringsAsFactors = default.stringsAsFactors(),  
           fileEncoding = "", encoding = "unknown", text, skipNul = FALSE)
```

```
read.csv(file, header = TRUE, sep = ",", quote = "\"",  
         dec = ".", fill = TRUE, comment.char = "", ...)
```

```
read.csv2(file, header = TRUE, sep = ";", quote = "\"",  
          dec = ",", fill = TRUE, comment.char = "", ...)
```

# *Scripts* em R

- Abrir o ficheiro teste no R
  - Determinar a média e o desvio padrão de
    - body\_fat
    - age
  - Criar uma nova variável
    - weight\_kg (1 lb = 0,45 kg)
    - height\_cm (1 inch = 2,54 cm)
    - IMC
    - IMC\_ordinal



# ggplot2

- Instalar o package
  - `install.packages("ggplot2");`
- Consultar <http://ggplot2.org/book/qplot.pdf>

# Gráficos

- Fazer um gráfico de dispersão entre
  - `body_fat` e `IMC` (qplot)
- Fazer um boxplot de
  - `age`