

R "— данные и графики

© А.Б. Шипунов*, Е.М. Балдин†

3 мая 2008 г.

1 R и работа с данными

Команды перемещения по директориям и поиск файлов:

```
> getwd()

[1] "/home/baldin/text/Articles/R/Rnw/R-2"

> setwd("./workdir")
> getwd()

[1] "/home/baldin/text/Articles/R/Rnw/R-2/workdir"

> dir()

[1] "1.txt"           "mydata-2.txt"     "mydata-3.txt"
[4] "mydata.txt"     "mydata-unicode.txt" "trees.csv"
[7] "x.rd"

> dir(pattern = glob2rx("*.txt"))

[1] "1.txt"           "mydata-2.txt"     "mydata-3.txt"
[4] "mydata.txt"     "mydata-unicode.txt"
```

Просмотр текстовых данных:

```
> file.show("mydata.txt")
```

Загрузка текстовых данных:

```
> read.table("mydata.txt", sep = ";", head = TRUE)
```

*e-mail: dactylorhiza@gmail.com

†e-mail: E.M.Baldin@inp.nsk.su

```
  a b c
1 1 2 3
2 4 5 6
3 7 8 9
```

Относительная адресация:

```
> read.table("../workdir/mydata.txt", sep = ";", head = TRUE)
```

```
  a b c
1 1 2 3
2 4 5 6
3 7 8 9
```

Указание кодовой страницы:

```
> read.table(file("mydata-unicode.txt", encoding = "UTF-8"), sep = ";",
+           head = TRUE)
```

```
  а б в
1 1 2 3
2 4 5 6
3 7 8 9
```

Задание имён строк:

```
> file.show("mydata-2.txt")
```

```
> read.table(file("mydata-2.txt", encoding = "KOI8-R"), head = TRUE)
```

```
  а б в
раз 1 2 3
два 4 5 6
три 7 8 9
```

Что делать, если целая часть от дробной отделяется запятой:

```
> read.table(file("mydata-3.txt", encoding = "KOI8-R"), dec = ",",
+           h = T)
```

```
  а   б   в
раз 1.1 2.2 3.3
два 4.4 5.0 6.0
три 7.0 8.0 9.0
```

Конвертация данных из сторонних форматов:

```

> library(foreign)
> help(package = foreign)
Чтение данных из буфера обмена:
> read.table("clipboard")
Сохранение данных в собственном формате R:
> x <- "яблоко"
> save(x, file = "x.rd")
> rm(x)
> dir(pattern = glob2rx("*.rd"))
[1] "x.rd"
> load("x.rd")
> x
[1] "яблоко"

```

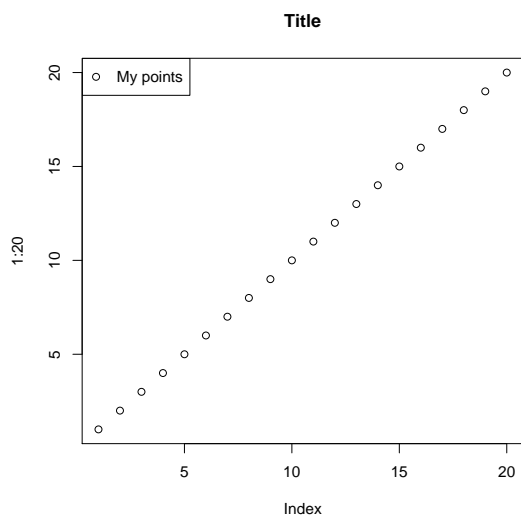
2 Графики

Устанавливаем шрифт для X Window (если он есть в системе, естественно):

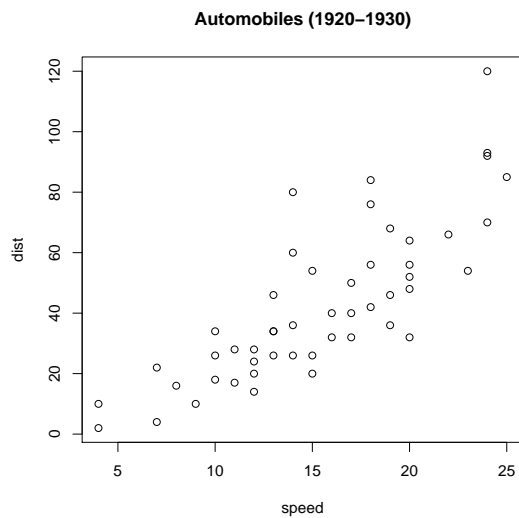
```

> X11(fonts = c("-rfx-helvetica-%s-%s-***%d-***-***-koi8-r",
+             "-adobe-symbol-medium-r-***%d-***-***-***-***"))
> setwd("..")
> plot(1:20, main = "Title")
> legend("topleft", pch = 1, legend = "My points")

```



```
> plot(cars)
> title(main = "Automobiles (1920-1930)")
```



2.1 Графические устройства

Создание png-файла:

```
> png(file = "1-20.png", bg = "transparent")
> plot(1:20)
> dev.off()
```

Создание pdf-файла:

```
> pdf("1-20.pdf", family = "NimbusSan", encoding = "KOI8-R.enc")
> plot(1:20, main = "Заголовок")
> dev.off()
> embedFonts("1-20.pdf")
```

2.2 Графические опции

```
> postscript("2hist.eps", width = 6, height = 6, horizontal = FALSE,
+   onefile = FALSE, paper = "special")
> old.par <- par(mfrow = c(2, 1))
> hist(cars$speed)
> hist(cars$dist)
> par(old.par)
> dev.off()
> embedFonts("2hist.eps")
```

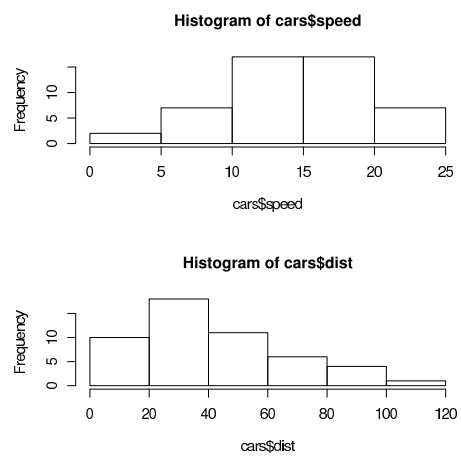


Рис. 1. Пример eps-файла (2hist.eps), созданного с помощью R

2.3 Разные типы диаграмм

Точечные диаграммы:

```
> dotchart(Titanic[, , "Adult", "No"], main = "People lost with the \"Titanic\"")
```

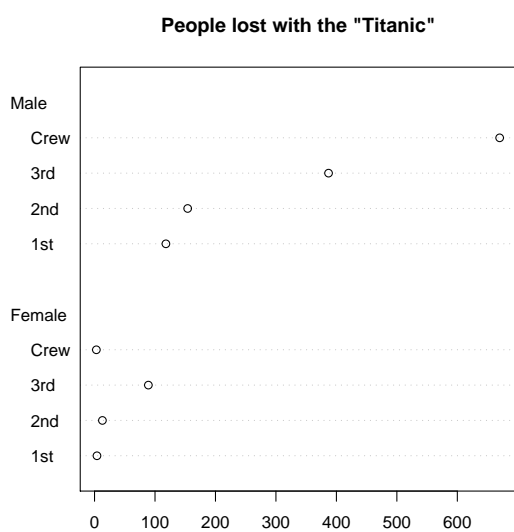
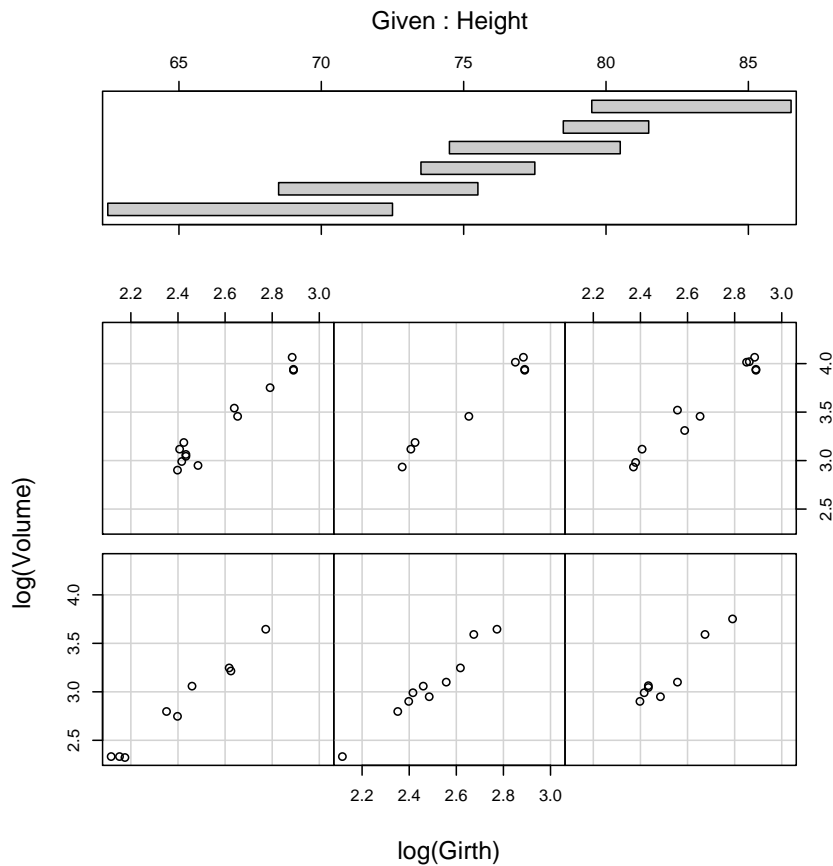


График-решётка:

```
> coplot(log(Volume) ~ log(Girth) | Height, data = trees)
```



2.4 Интерактивность

```
> plot(1:20)
> text(locator(), "Моя любимая точка", pos = 4)
```

3 Как сохранять результаты

Запись данных в файл:

```
> setwd("./workdir")
> write.table(file = "trees.csv", trees, row.names = F, sep = ";",
+   quote = F)
> setwd("../")
```

Запись результатов выполнения команды в файл:

```
> setwd("./workdir")
> sink("1.txt", split = T)
> print("2+2")
```

```
[1] "2+2"
```

```
> 2 + 2
```

```
[1] 4
```

```
> sink()
```

```
> setwd("../")
```