

Наука о данных для Цифровых гуманитарных исследований в R

Г. А. Мороз

Оглавление

1	Введение	5
1.1	Установка R и RStudio	5
1.2	Полезные ссылки	5
2	Introduction	7

Глава 1

Введение

1.1 Установка R и RStudio

В данной книге используется исключительно R (R Core Team, 2019), так что для занятий понадобятся:

- R (можно скачать [здесь](#))
- RStudio — IDE для R (можно скачать [здесь](#))
- и некоторые пакеты на R

Часто можно увидеть или услышать, что R — язык программирования для “статистической обработки данных”. Изначально это, конечно, было правдой, но уже давно R — это полноценный язык программирования, который при помощи своих пакетов позволяет решать огромный спектр задач. В данной книге используются следующая версия R:

```
sessionInfo()$R.version$version.string
```

```
## [1] "R version 3.6.1 (2019-07-05)"
```

Некоторые люди не любят устанавливать лишние программы себе на компьютер, несколько вариантов есть и для них:

- RStudio cloud — полная функциональность RStudio, пока бесплатная, но скоро это исправят;
- RStudio on rollApp — облачная среда, позволяющая разворачивать программы.

1.2 Полезные ссылки

В интернете легко найти документацию и tutorиалы по самым разным вопросам в R, так что главный залог успеха — грамматно пользоваться поисковиком,

и лучше на английском языке.

- книга (Wickham and Grolemund, 2016) является достаточно сильной альтернативой всему курсу
- `stackoverflow` — сервис, где достаточно быстро отвечают на любые вопросы (не обязательно по R)
- `RStudio community` — быстро отвечают на вопросы, связанные с R
- русский `stackoverflow`
- `R-bloggers` — сайт, где собираются новинки, связанные с R
- чат, где можно спрашивать про R на русском (но почитайте правила чата, перед тем как спрашивать)
- чат по визуализации данных, чат датажурналистов
- канал про визуализацию, дата-блог “Новой газеты”, ...

Глава 2

Introduction

You can label chapter and section titles using after them, e.g., we can reference Chapter 2. If you do not manually label them, there will be automatic labels anyway, e.g..

Figures and tables with captions will be placed in `figure` and `table` environments, respectively.

```
par(mar = c(4, 4, .1, .1))  
plot(pressure, type = 'b', pch = 19)
```

Reference a figure by its code chunk label with the `fig:` prefix, e.g., see Figure 2.1. Similarly, you can reference tables generated from `knitr::kable()`, e.g., see Table 2.1.

```
knitr::kable(  
  head(iris, 20), caption = 'Here is a nice table!',  
  booktabs = TRUE  
)
```

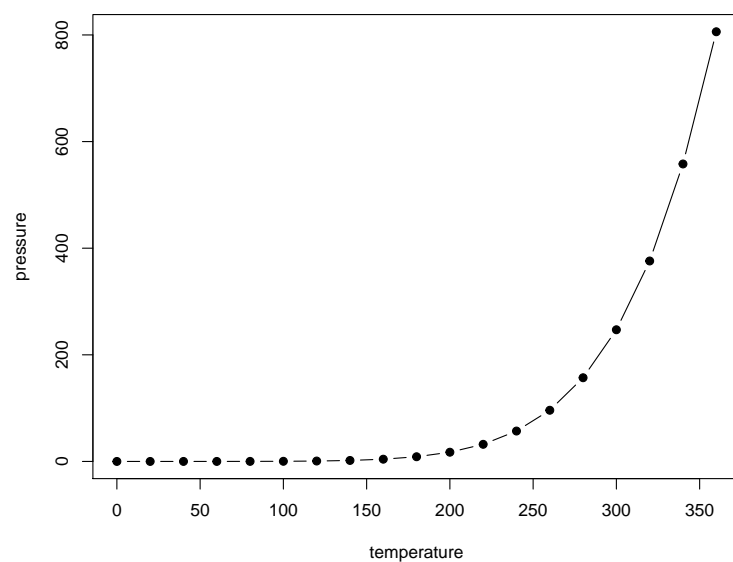


Рис. 2.1: Here is a nice figure!

Таблица 2.1: Here is a nice table!

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
4.6	3.1	1.5	0.2	setosa
5.0	3.6	1.4	0.2	setosa
5.4	3.9	1.7	0.4	setosa
4.6	3.4	1.4	0.3	setosa
5.0	3.4	1.5	0.2	setosa
4.4	2.9	1.4	0.2	setosa
4.9	3.1	1.5	0.1	setosa
5.4	3.7	1.5	0.2	setosa
4.8	3.4	1.6	0.2	setosa
4.8	3.0	1.4	0.1	setosa
4.3	3.0	1.1	0.1	setosa
5.8	4.0	1.2	0.2	setosa
5.7	4.4	1.5	0.4	setosa
5.4	3.9	1.3	0.4	setosa
5.1	3.5	1.4	0.3	setosa
5.7	3.8	1.7	0.3	setosa
5.1	3.8	1.5	0.3	setosa

Литература

R Core Team (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Wickham, H. and Grolemund, G. (2016). *R for data science: import, tidy, transform, visualize, and model data*. O'Reilly Media, Inc.