

Podstawy programowania R

Łukasz Wawrowski

Contents

Wprowadzenie	5
1 Wprowadzenie do R	7
1.1 R	7
1.2 RStudio	7
1.3 Pakiety	7
2 Struktury danych	11
2.1 Wektor	11
2.2 Macierz	11
2.3 Faktor/czynnik	11
2.4 Lista	11
2.5 Ramka danych	11
2.6 Braki danych	11
2.7 Import danych	11
3 Przetwarzanie danych	13
3.1 Pakiet tidyverse	13
3.2 Wybieranie kolumn	13
3.3 Filtrowanie	13
3.4 Dodawanie nowych zmiennych	13
3.5 Grupowanie	13
3.6 Podsumowanie	13
3.7 Łączenie zbiorów	13
3.8 Wąska i szeroka reprezentacja danych	13
4 Programowanie w R	15
4.1 Funkcje	15
4.2 Pętle	15
4.3 Instrukcje warunkowe	15
5 Wizualizacja danych	17
5.1 Wbudowane funkcje	17
5.2 Pakiet ggplot2	17

Wprowadzenie

Literatura podstawowa:

- Przemysław Biecek - *Przewodnik po pakiecie R*
- Marek Gagolewski - *Programowanie w języku R. Analiza danych, obliczenia, symulacje.*
- Garret Golemund, Hadley Wickham - *R for Data Science*

Literatura dodatkowa:

- inne pozycje po polsku
- inne pozycje po angielsku

Chapter 1

Wprowadzenie do R

GNU R to interpretowany język programowania oraz środowisko do obliczeń statystycznych i wizualizacji wyników [Wikipedia 2017].

Robert A. Muenchen - The Popularity of Data Science Software

1.1 R

Bazowa wersja R jest do pobrania ze strony r-project.org.

1.2 RStudio

RStudio to zintegrowane środowisko programistyczne (IDE) dla języka R dostępne za darmo na stronie RStudio.

Z R można także korzystać w Microsoft Visual Studio.

W codziennej pracy RStudio jest wygodniejsze, jednak długotrwałe obliczenia lepiej uruchamiać w zwykłym R.

1.3 Pakiety

Podstawowe możliwości R są dosyć ograniczone. Rozszerzają je pakiety, których obecnie jest ponad 12 tysięcy. Można je przeglądać według kategorii w CRAN Task Views lub w wygodnej wyszukiwarce METACRAN.

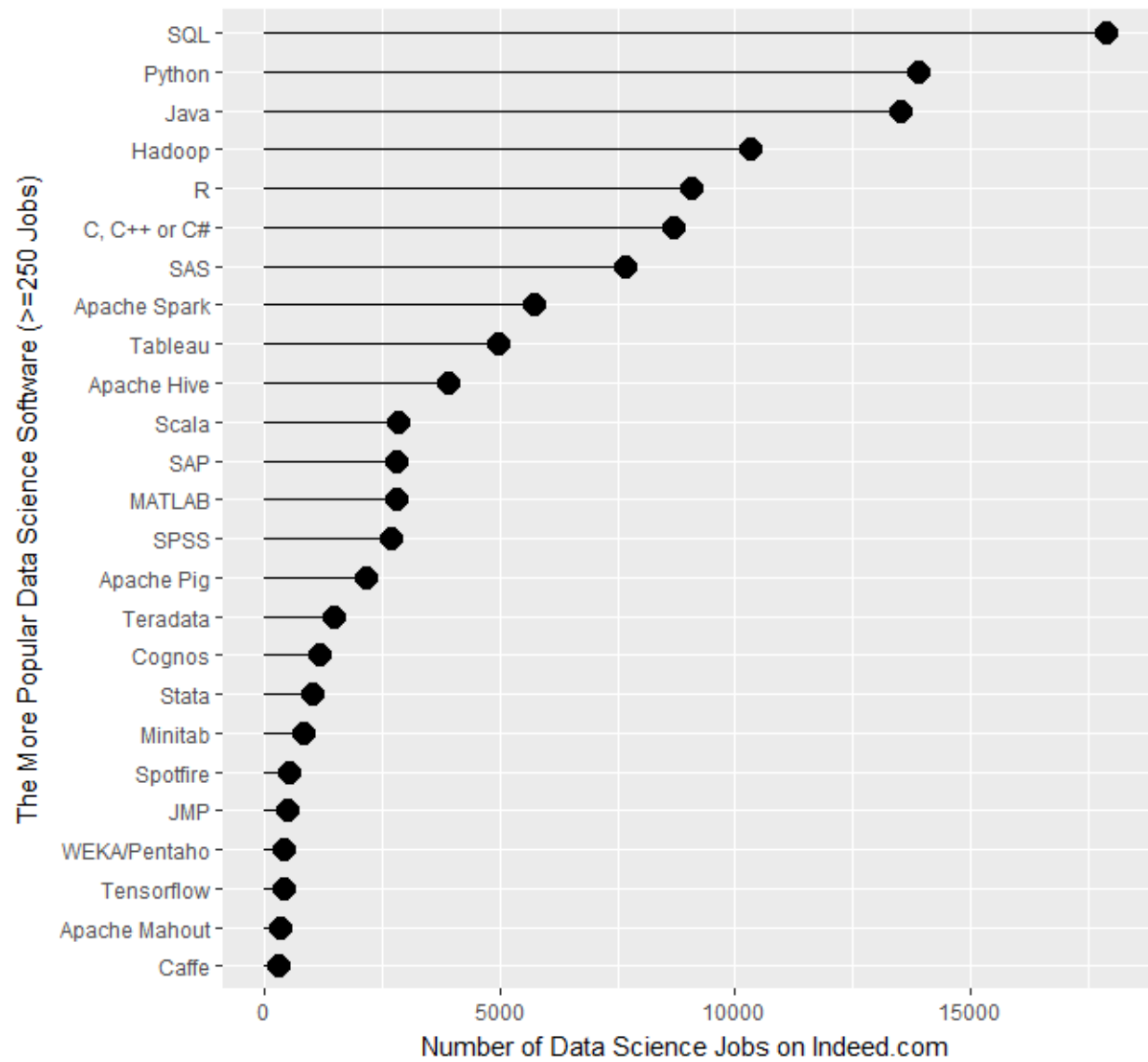


Figure 1.1:

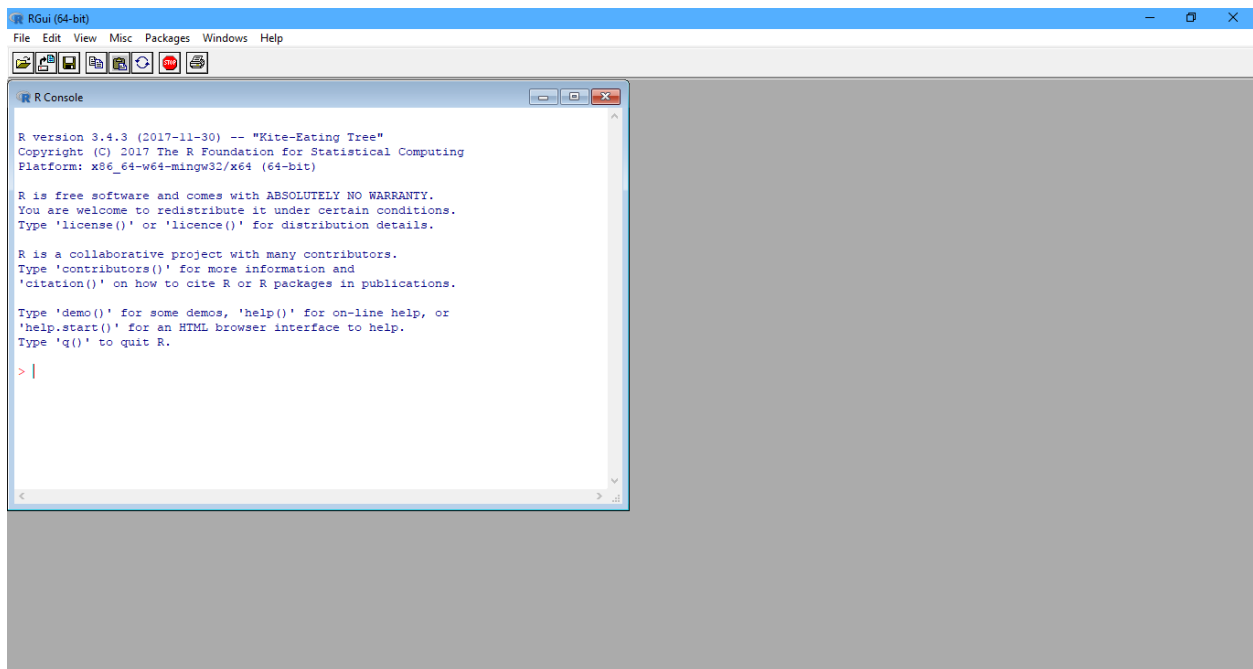


Figure 1.2:

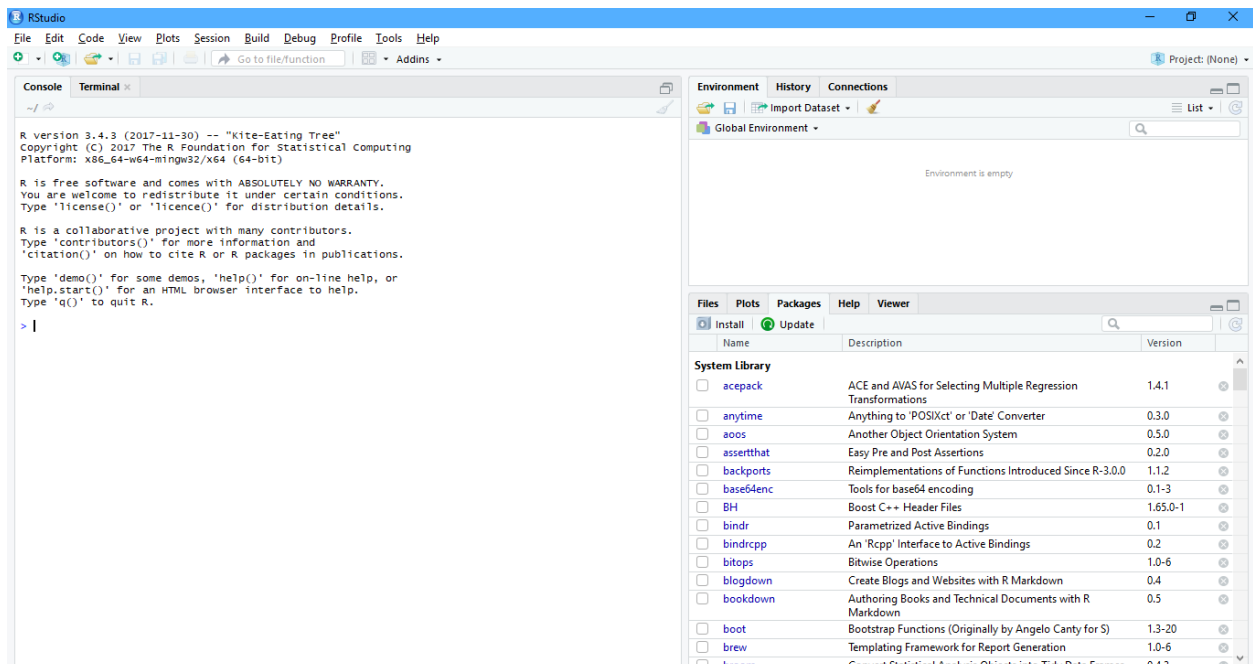


Figure 1.3:

Chapter 2

Struktury danych

2.1 Wektor

2.2 Macierz

2.3 Faktor/czynnik

2.4 Lista

2.5 Ramka danych

2.6 Braki danych

2.7 Import danych

Chapter 3

Przetwarzanie danych

3.1 Pakiet tidyverse

3.2 Wybieranie kolumn

3.3 Filtrowanie

3.4 Dodawanie nowych zmiennych

3.5 Grupowanie

3.6 Podsumowanie

3.7 Łączenie zbiorów

3.8 Wąska i szeroka reprezentacja danych

Chapter 4

Programowanie w R

4.1 Funkcje

4.2 Pętle

4.3 Instrukcje warunkowe

Chapter 5

Wizualizacja danych

5.1 Wbudowane funkcje

5.2 Pakiet ggplot2