Министерство образования и науки Российской Федерации

Севастопольский государственный университет

Кафедра ИС

Отчет

По дисциплине: “Интеллектуальный анализ данных”

Лабораторная работа №1

“ Исследование возможностей языка R

для статистического анализа данных”

Выполнил:

ст.гр. ИС/б-32

Долженко И.А.

Проверила:

Сырых О.А.

Севастополь

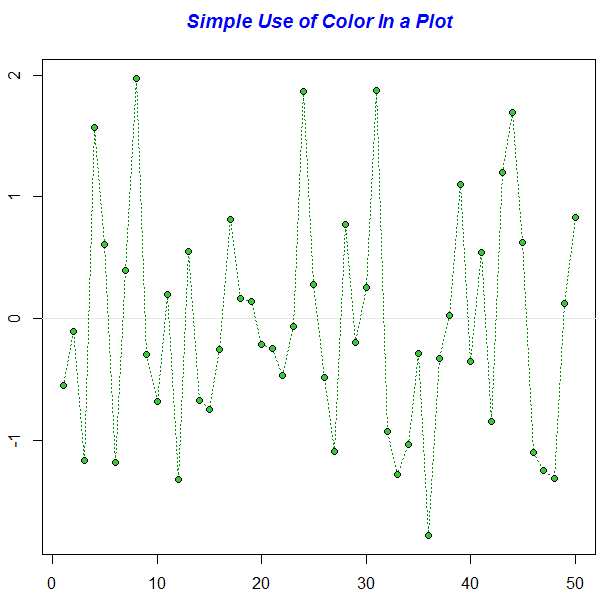
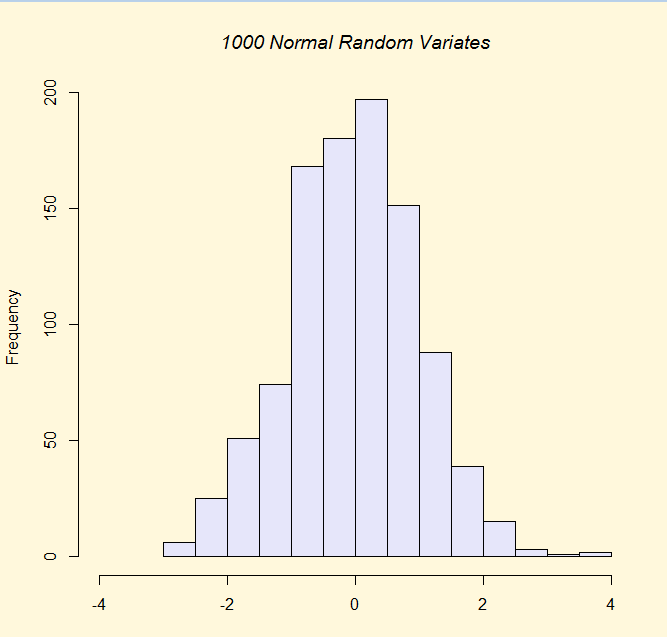
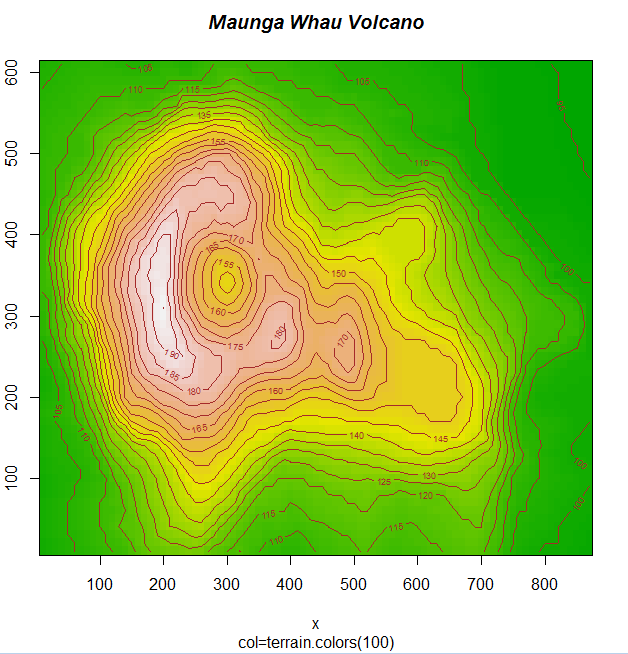
2020

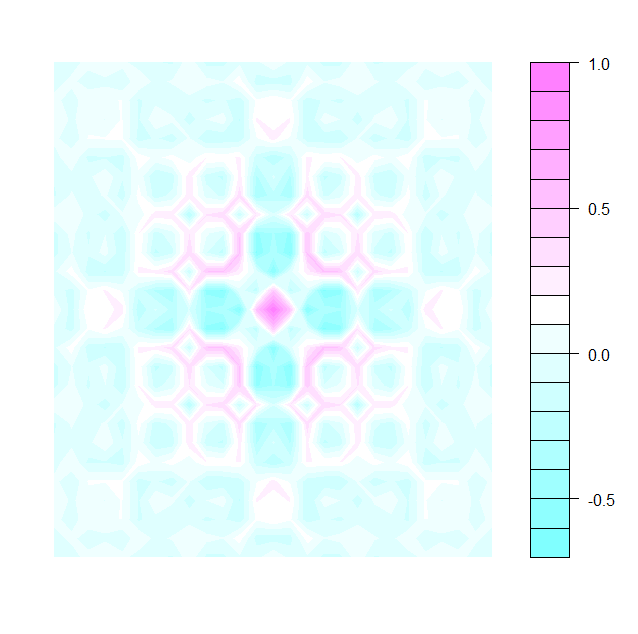
1 ЦЕЛЬ РАБОТЫ

Изучить основные особенности языка R; Исследовать возможности языка R для работы с графикой.

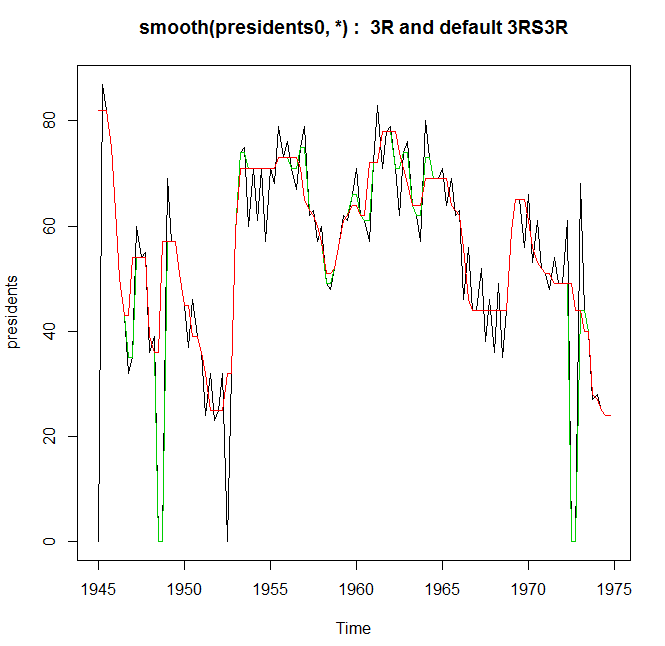
2 ХОД РАБОТЫ

1. Исследуем команду demo:

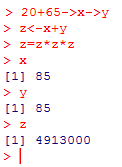
* demo(graphics)
* demo(image)



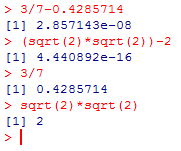
* demo(smooth)



1. Для присваивания используется символ «<‒» или «‒>» (можно также использовать традиционное «=»).



1. Порядок арифметических действий в R стандартный, знакомый со школьной математики. Скобки (раскрывающиеся изнутри наружу) позволяют этот порядок действий менять.



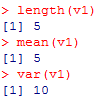
1. Команды разделяются точкой с запятой «;» или символом перевода на новую строку.



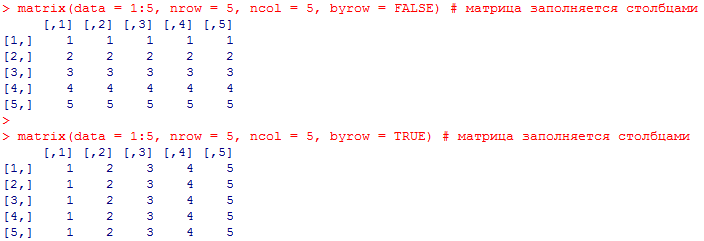
1. Кроме строк и чисел можно также создавать и сохранять векторы. Вектор создается с помощью функции c(), которая объединяет несколько однотипных элементов.



1. Можно получать различные свойства вектора.



1. Матрицы создаются с помощью команды matrix().



1. Основной функцией для рисования объектов в R является функция plot(x, y, …):

x – координаты точек графика, либо некоторая графическая структура, функция или объект, содержащий методы рисования.

y – y-координаты точек графика, если x – соответствующего типа.

Остальные графические параметры. Перечислим некоторые из них:

а) параметр type позволяет изменять внешний вид точек на графике и может принимать одно из следующих значений:

"p" – точки (points; используется по умолчанию);

"l" – линии (lines);

"b" – изображаются и точки, и линии (both points and lines);

"o" – точки изображаются поверх линий (points over lines);

"h" – гистограмма (histogram);

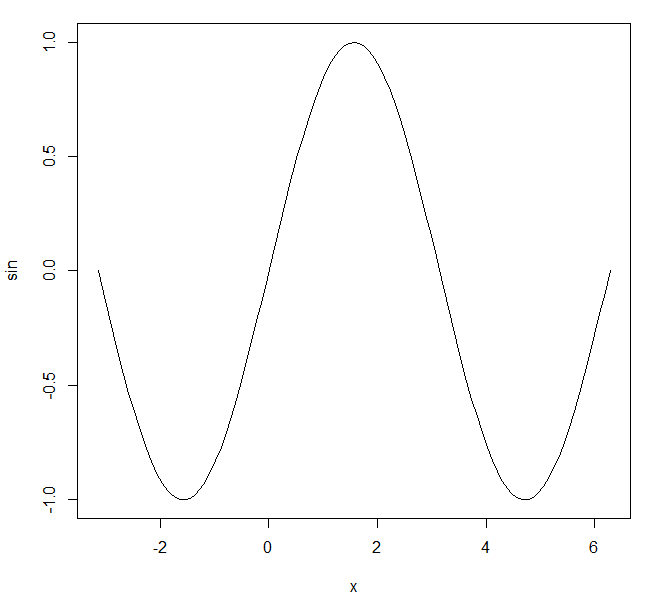
"s" – ступенчатая кривая (steps);

"n" – данные не отображаются (no points).

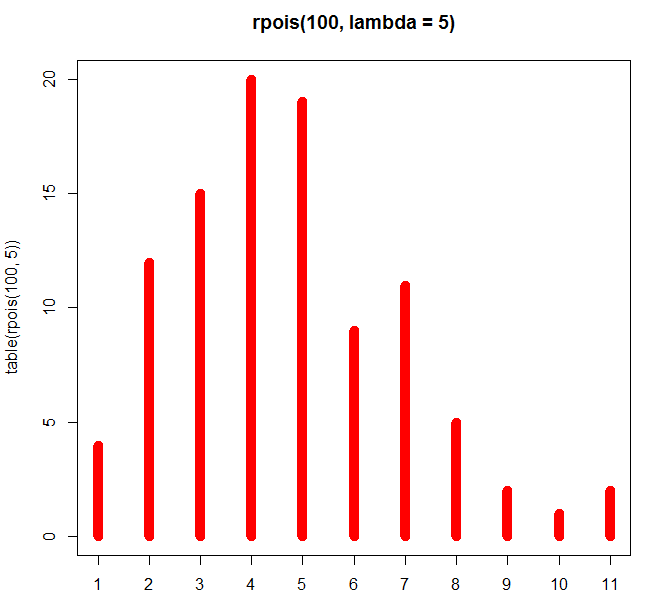
б) параметры xlabи ylabзадают название осей абсцисс и ординат, соответственно;

в) параметр mainзадаёт заголовок графика.

- plot(sin, -pi, 2\*pi)

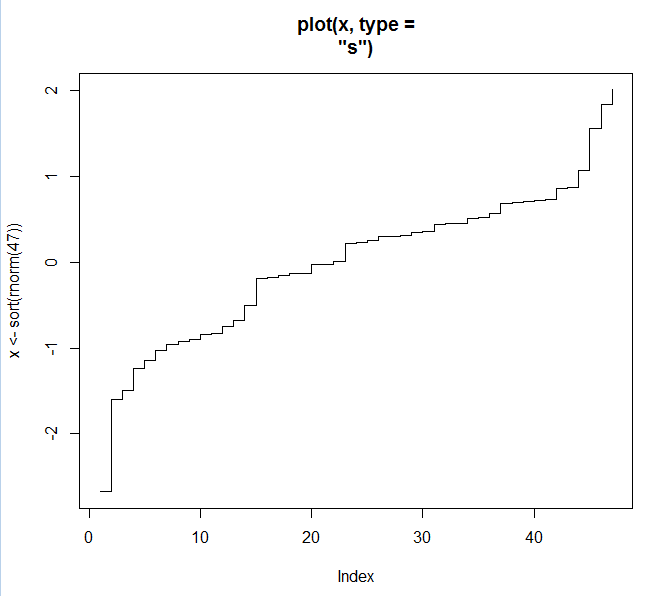


* plot(table(rpois(100, 5)), type = "h", col = "red", lwd = 10, main = "rpois(100, lambda = 5)")



* plot(x <- sort(rnorm(47)), type = "s", main = "plot(x, type =\"s\")")

points(x, cex = .5, col = "dark red")



3 ОТВЕТЫ НА КОНТРОЛЬНЫЕ ВОПРОСЫ

**1. Особенности языка R.**

R является и языком и программным обеспечением; его наиболее замечательный особенности:

- эффективная обработка данных и простые средства для сохранения результатов;

- набор операторов для обработки массивов, матриц, и других сложных конструкций;

- большая, последовательная, интегрированная коллекция инструментальных средств для проведения статистического анализа,

- многочисленные графические средства;

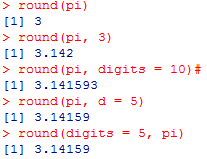
- простой и эффективный язык программирования, который включает много возможностей.

**2. Команда для получения подробной информации о функции в R.**

Для получения подробной информации (справки) о любой функции, необходимо выполнить команду help(<имя функции>) либо ?<имя функции>.

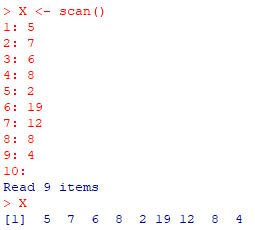
**3. Структура и особенности команды round() в R.**

Команда round() имеет два аргумента: число, которое нужно округлить, и значение digits, сообщающее, до какого знака округлять. Пример:



**4. Команды для работы с векторами в R.**

Вектор можно создать также при помощи функции scan(), которая "считывает" последовательно вводимые с клавиатуры значения.



Для создания векторов, содержащих совокупность последовательных чисел, удобна функция seq(). Так, вектор с именем S, содержащий совокупность целых чисел от 1 до 7, можно создать следующим образом:

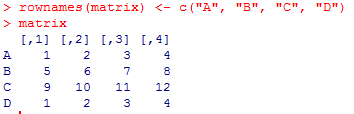


Для упорядочения значений вектора по возрастанию или убыванию используют функцию sort() в сочетании с аргументом decreasing = FALSE или decreasing = TRUE.

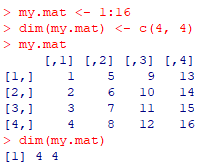


**5. Команды для работы с матрицами в R.**

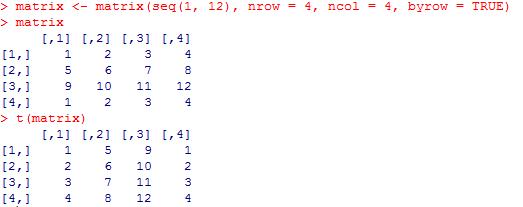
В качестве заголовков строк и столбцов создаваемой матрицы автоматически выводятся соответствующие индексные номера (строки: [1,], [2,], и т.д.; столбцы: [,1], [,2], и т.д.). Для придания пользовательских заголовков строкам и столбцам матриц используют функции rownames() и colnames() соответственно.



Функция dim() очень полезна. Она позволяет проверить размерность уже имеющейся матрицы (или таблицы данных), например:

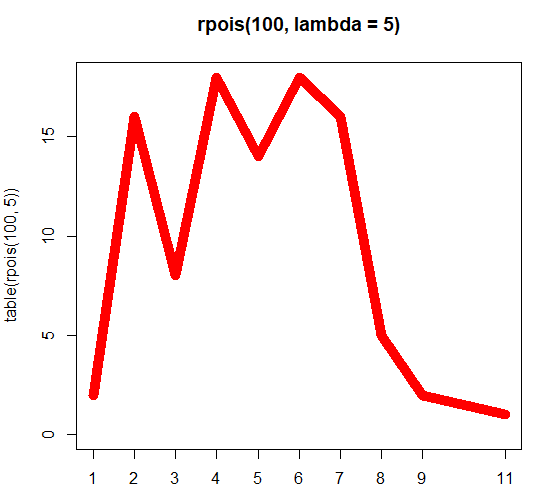


Матрицу можно транспонировать при помощи функции t().



**6. Работа с графикой в R.**

- plot(table(rpois(100, 5)), type = "l", col = "red", lwd = 10, main = "rpois(100, lambda = 5)")



ВЫВОДЫ

В ходе выполнения лабораторной работы были изучены основные особенности языка R. Исследована команда demo() для запуска демонстрационных команд и команда для получения подробной информации о функции - help(). Изучены команды для работы с векторами, матрицами и команды для работы с графикой.