Лабораторная работа №1

«Исследование возможностей языка R для статистического анализа данных»

1 ЦЕЛЬ РАБОТЫ

– изучить основные особенности языка R;

– исследовать возможности языка R для работы с графикой.

2 ХОД РАБОТЫ

1. Исследуем команду 'demo()' (см. рисунок 1). Demo - это удобный интерфейс для запуска некоторых демонстрационных R-скриптов. demo() дает список доступных тем.

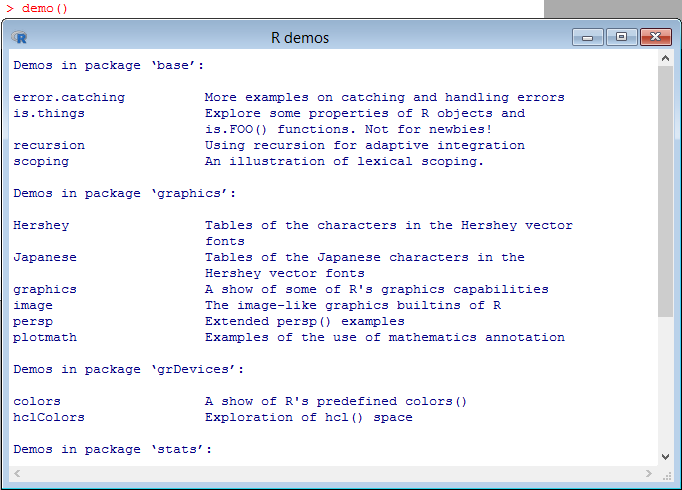


Рисунок 1 – Функция demo() для запуска демонстрационных программ

Создадим матрицу и используем команду persp(), которая рисует перспективные участки поверхности на плоскости x-y (см. рисунок 2).

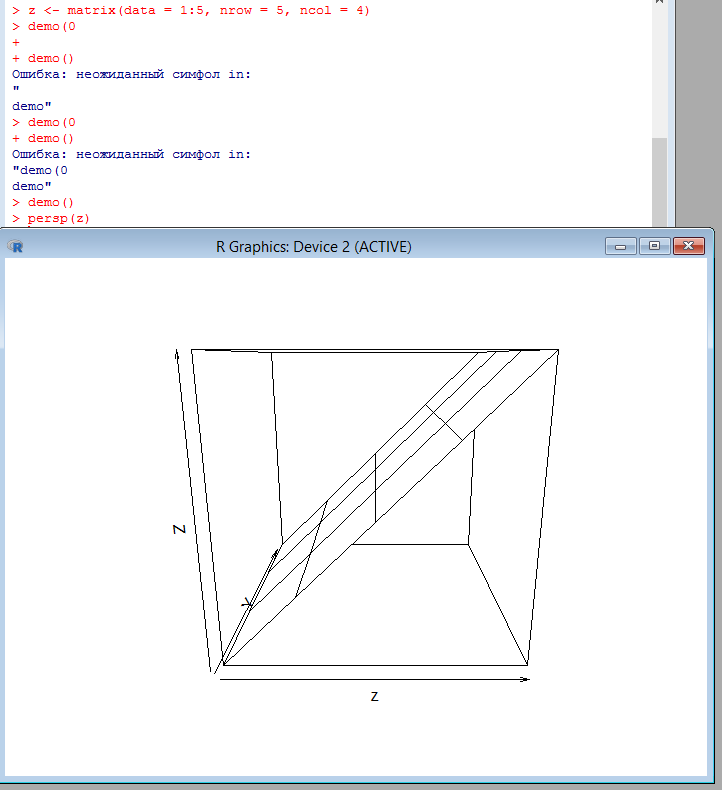
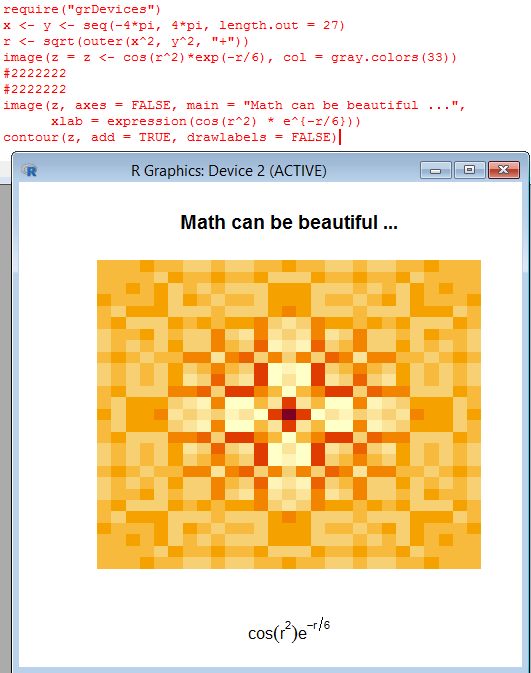


Рисунок 2 – График, нарисованный с помощью функции persp()

Исследуем функцию image, создающую сетку из цветных или серых прямоугольников с цветами, соответствующими значениям в z. Это может быть использовано для отображения трехмерных или пространственных данных или изображений (см. рисунок 3 и рисунок 4).



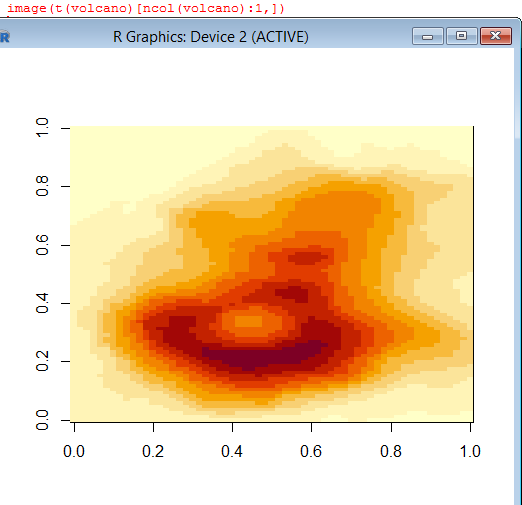


Рисунок 3 – График, нарисованный с помощью функции image()

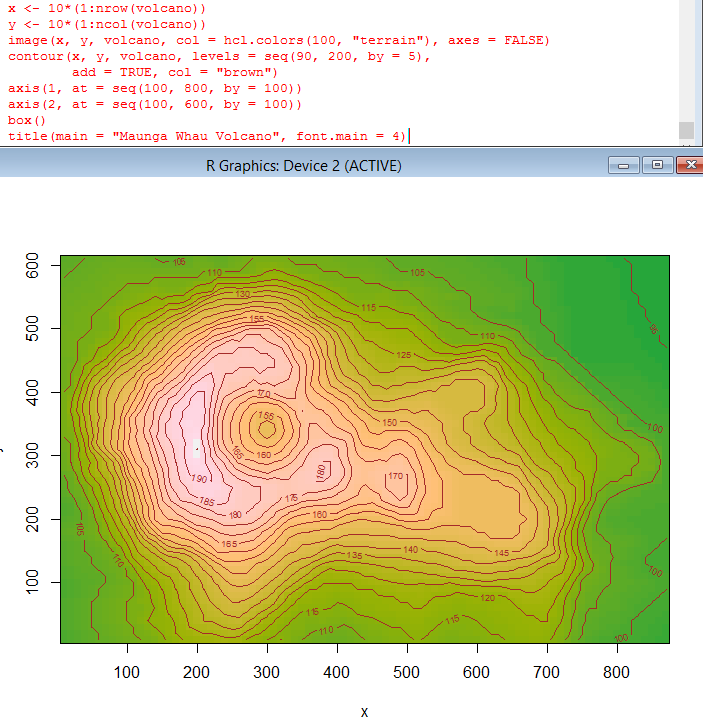
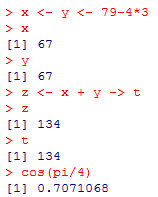


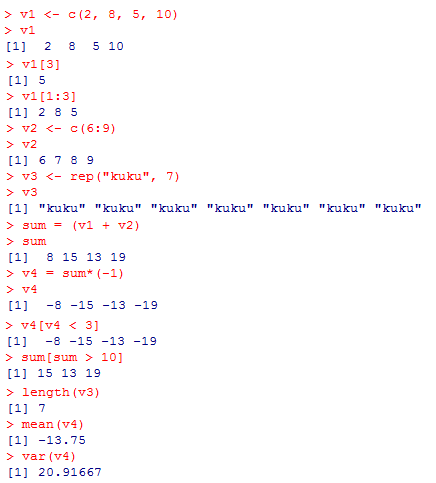
Рисунок 4 – Демонстрация возможностей функции image()

1. Исследуем основные функции и команды языка R.

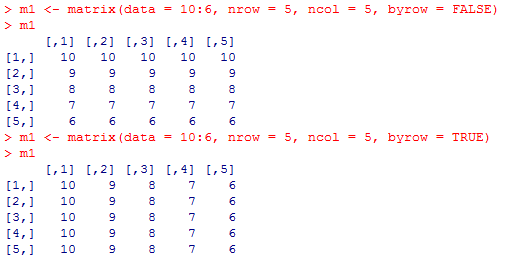
Рассмотрим присваивание и простые арифметические вычисления:



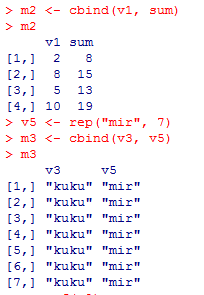
Вектор создаётся с помощью функции с(), объединяющей несколько однотипных элементов:



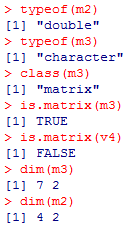
Матрицы создаются с помощью команды matrix():



Получим матрицу с помощью функции-комбинатора cbind(), объединяющую столбцы:



Получим тип элементов матрицы, класс объекта, а также проверим, является ли указанный объект матрицей и получим её размерность с помощью следующих функций:



Построим график, используя функцию plot(), подключив с помощью команды require() заранее установленный пакет в рабочее пространство, а также lines(lowess()) проведёт функциональную линию по координатам сглаживания графика (см. рисунок 5).

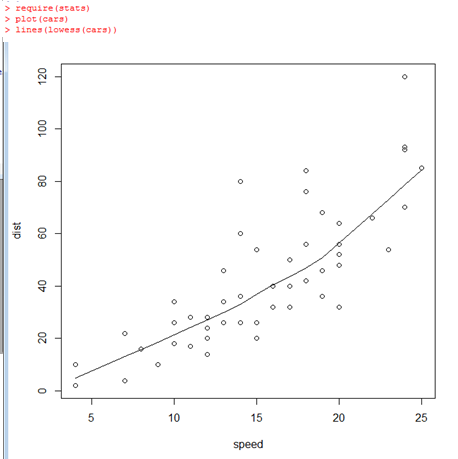


Рисунок 5 – Использование функции plot()

Построим график косинусоиды (см. рисунок 6):

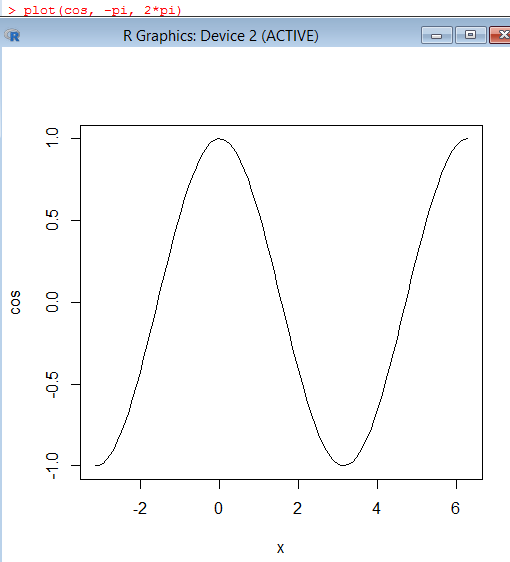


Рисунок 6 – График косинуса

Построим график дискретного пуассоновского распределения с помощью plot(table(rpois())) – обобщенной функции графика для (непредвиденных) объектов таблицы (см. рисунок 7).

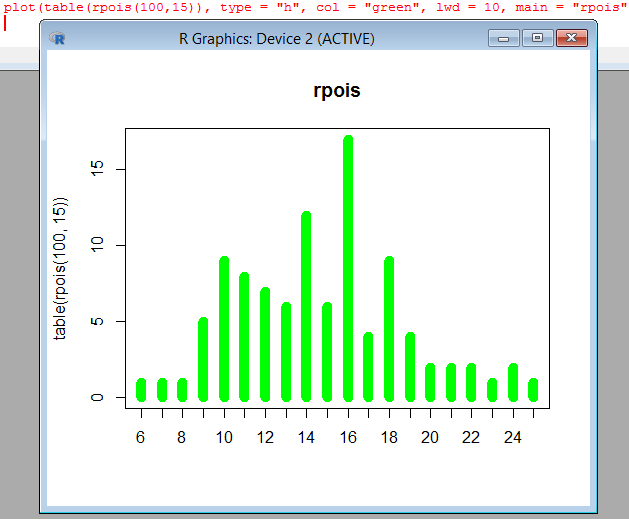


Рисунок 7 – График дискретного распределения Пуассона

Функция rnorm() служит для случайной генерации совокупностей нормально распределенных чисел. Построим её график (см. рисунок 8).

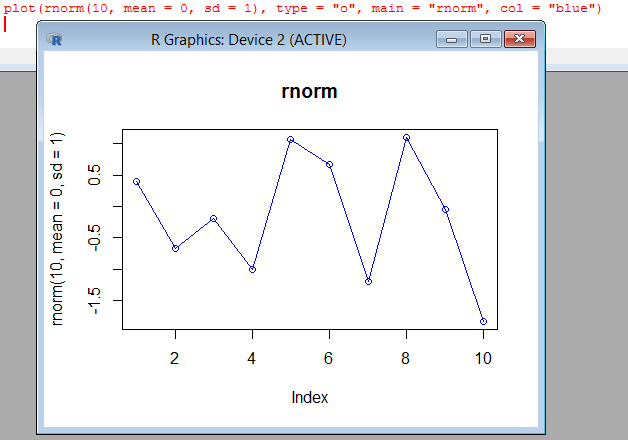


Рисунок 8 – График функции rnorm()

Построим график, тип которого – ступенчатая кривая (см. рисунок 9).

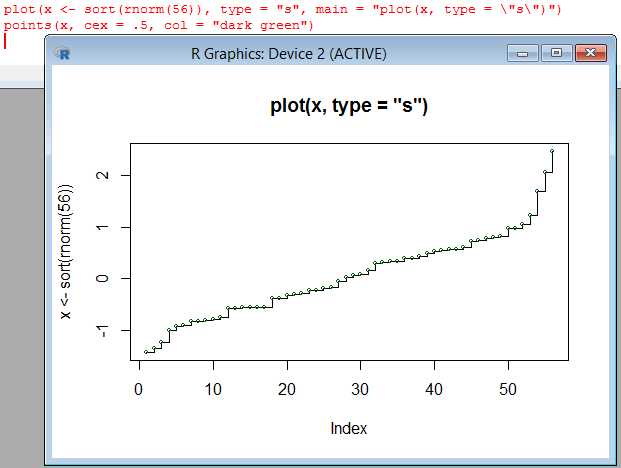


Рисунок 9 – График отсортированной генерации совокупностей нормально распределенных чисел

1. Ответим на контрольные вопросы.
   1. Особенности языка R:

– эффективная обработка данных и простые средства для сохранения результатов;

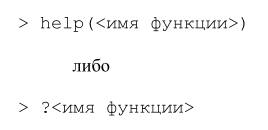
– набор операторов для обработки массивов, матриц, и других сложных конструкций;

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

– многочисленные графические средства;

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

* 1. Команда для получения подробной информации о функции в R:



* 1. Структура и особенности команды round() в языке R.

Данная команда служит для округления нецелочисленных значений. Она имеет два аргумента: число, которое нужно округлить, и значение digits, сообщающее, до какого знака округлять.

Примеры:



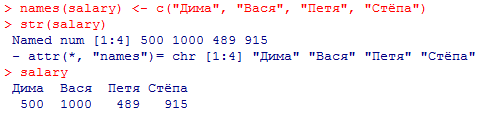


* 1. Рассмотрим команды для работы с векторами.

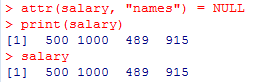
Создадим вектор salary, с индексами по умолчанию – порядковыми номерами, выводящимся с помощью команды str(salary):



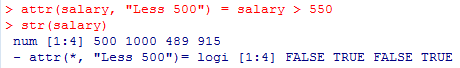
Создадим именованный вектор, как атрибут вектора salary и выведем вектор salary:



Удалим данный атрибут вектора salary, присвоив ему значение NULL:

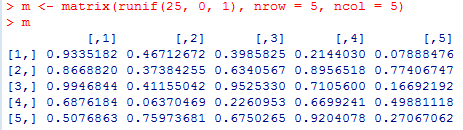


Придумаем свой атрибут с помощью команды attr() и зададим значение, по которому он будет определяться:

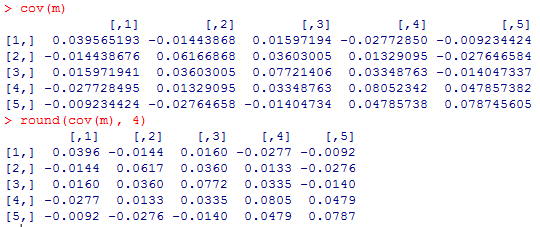


* 1. Рассмотрим команды для работы с матрицами.

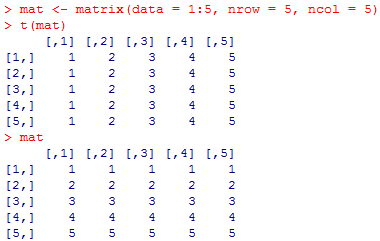
Создадим матрицу, заполнив её случайными величинами с помощью команды runif(n, min = 0, max = 0):



С помощью команды cov(x) рассчитаем ковариационную матрицу и округлим слишком громоздкие вещественные значения ранее рассмотренной командой round():



Произведём транспонирование матрицы командой t(x):

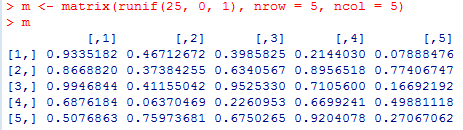


Рассчитаем сумму по каждому столбцу транспонированной матрицы командой colSums(x):



* 1. Рассмотрим работу с графикой в языке R.

На основе созданной в предыдущем разделе матрицы с рандомными значениями изобразим диаграмму рассеяния для всех возможных пар переменных из х, применив команду pairs(x) (см. рисунок 10):



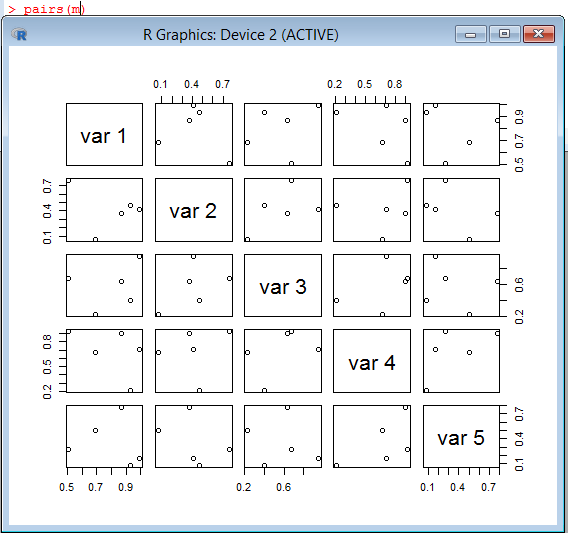


Рисунок 10 – Работа команды pairs(x)

Изобразим круговую диаграмму командой pie(x, labels = names(x)) (см. рисунок 11):

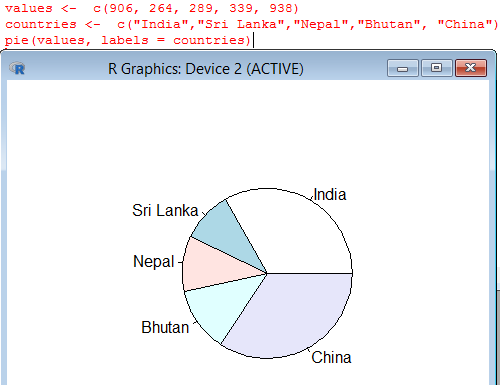


Рисунок 11 – Круговая диаграмма

ВЫВОДЫ

В ходе данной лабораторной работы была изучена статистическая система анализа – R, являющейся и языком, и программным обеспечением. Был изучен синтаксис языка R, с помощью различных использованных функций были исследованы способы работы с векторами, матрицами и графикой.