Отчёт к лабораторной работе

по дисциплине  
«Интеллектуальный анализ данных»

выполнил   
студент гр. ИС/б-18-1-з Демиденко А. А.  
зачётная книжка № 481483  
принял Шумейко И. П.

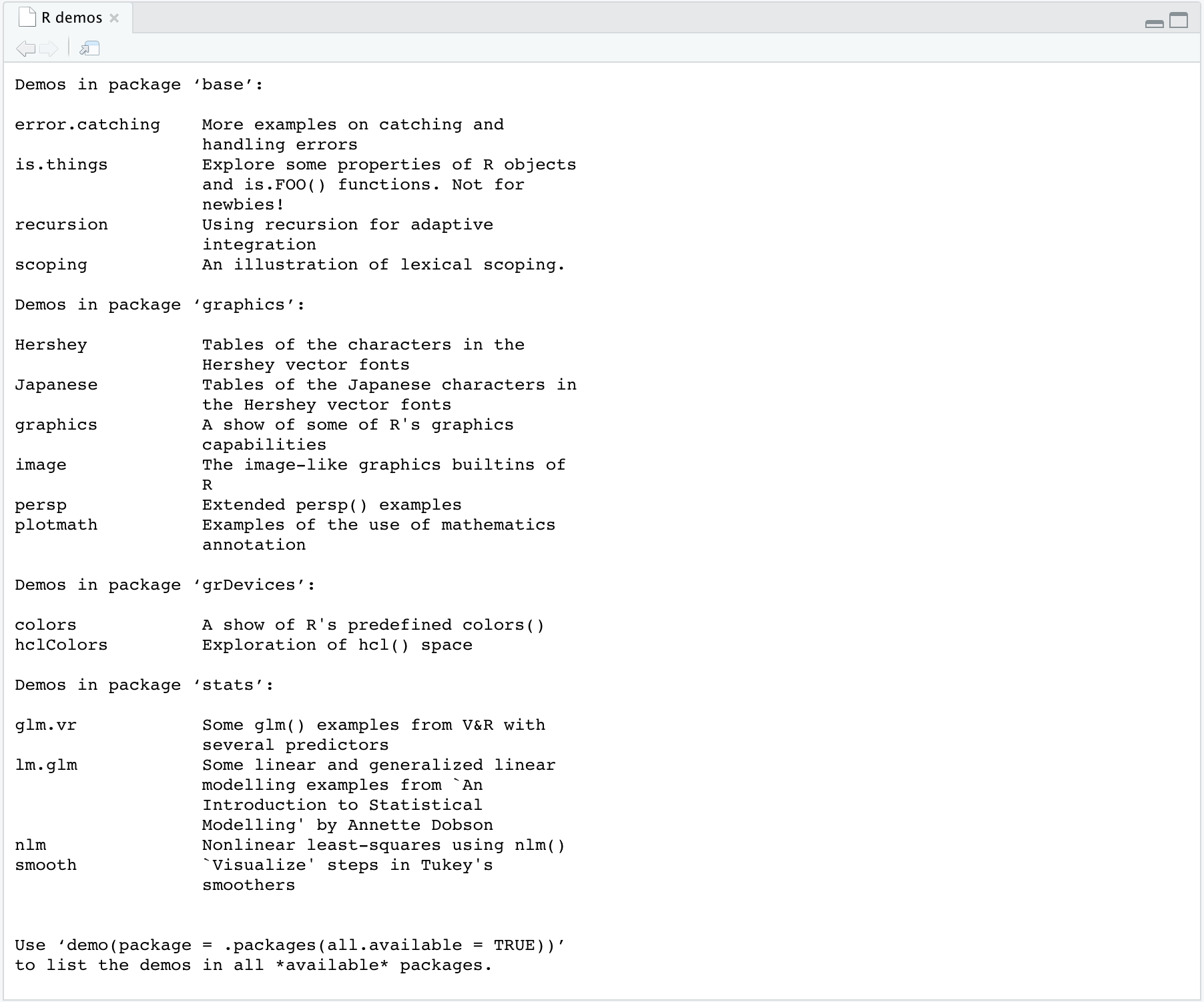
Лабораторной работа № 1  
«Исследование возможностей языка R для статистического анализа данных»

## Цель работы

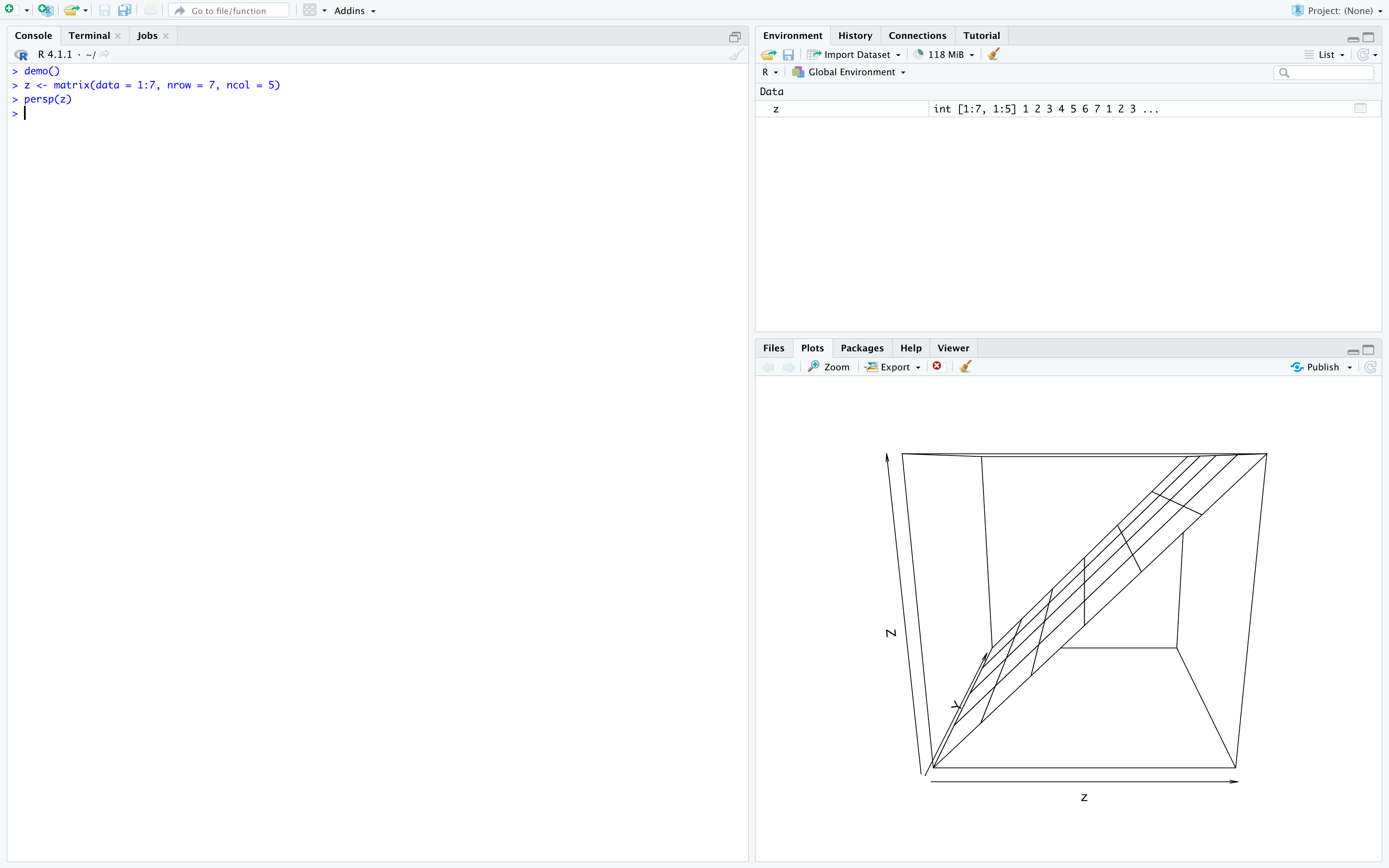
* Изучить основные особенности языка R
* Исследовать возможности языка R для работы с графикой

## Ход работы

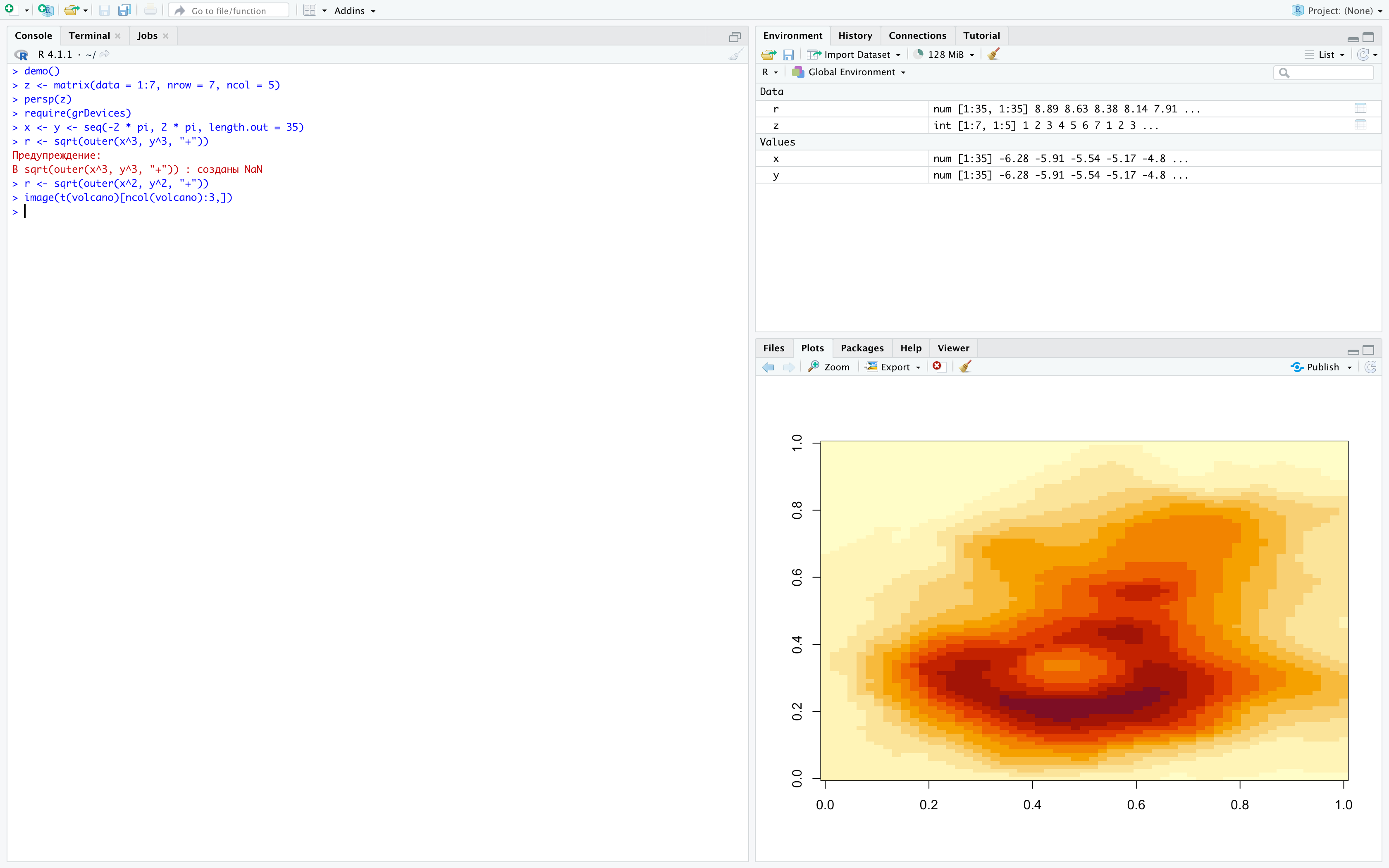
1. Исследуем команду `demo()` (см. рисунок 1). Demo - это удобный интерфейс для запуска демонстрационных R-скриптов. `demo()` выводит список доступных демонстраций.

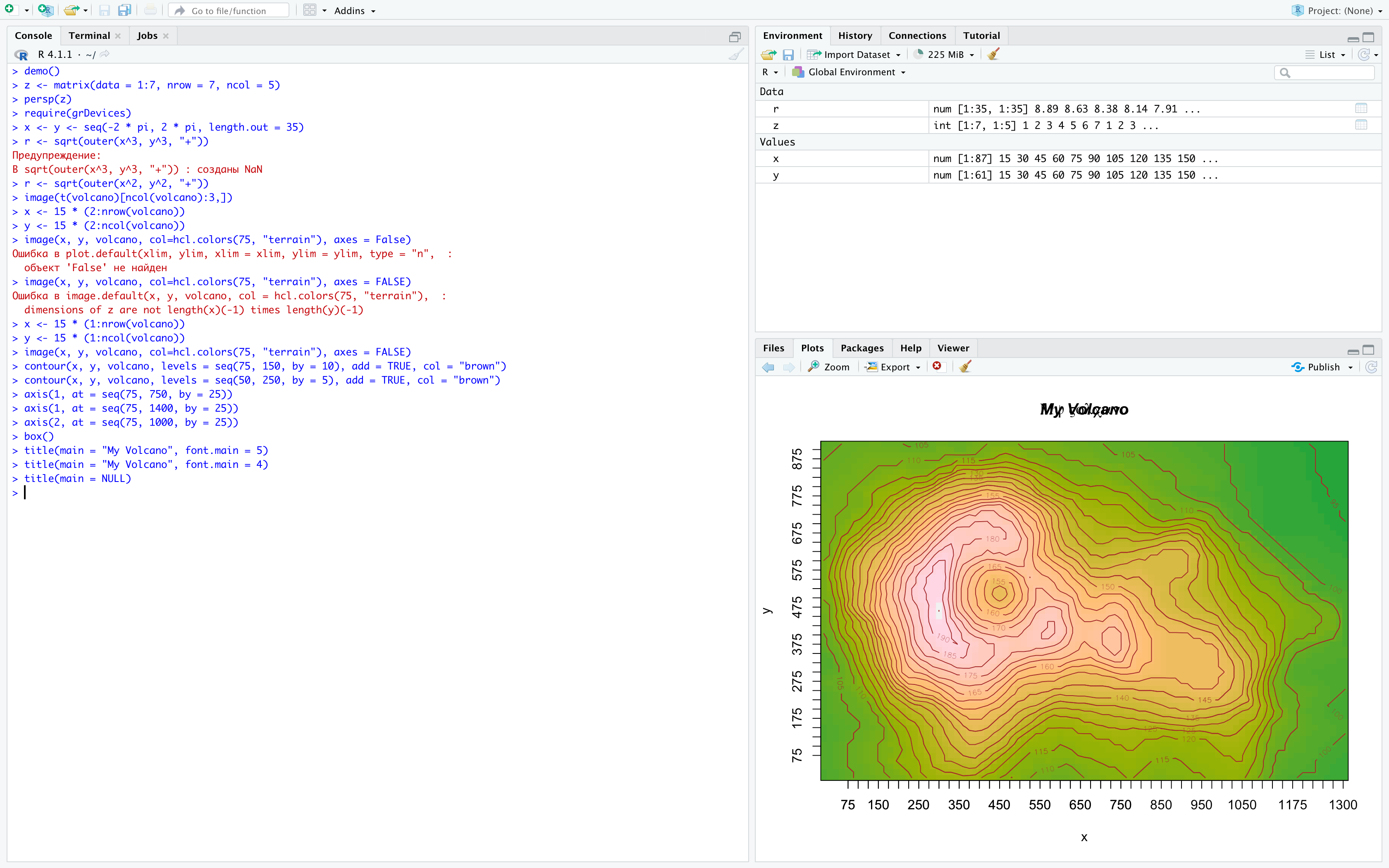
  
Рисунок 1 – Функция demo() для запуска демонстраций

Создадим матрицу и воспользуемся функцией `persp()` для отрисовки перспективных участков поверхности на плоскости x-y (см. рисунок 2).

  
Рисунок 2 – Результат выполнения функции persp()

Исследуем функцию image, создающую сетку из цветных или серых прямоугольников с цветами, соответствующих значениям в z. Это можно использовать для отображения трёхмерных или пространственных данных или изображений (см. рисунок 3 и 4).

  
Рисунок 3 – График, нарисованный функцией image()

  
Рисунок 4 – Демонстрация возможностей функции image()

1. Исследуем основные функции и команды языка R.

Рассмотрим присваивание и простые арифметические операции:

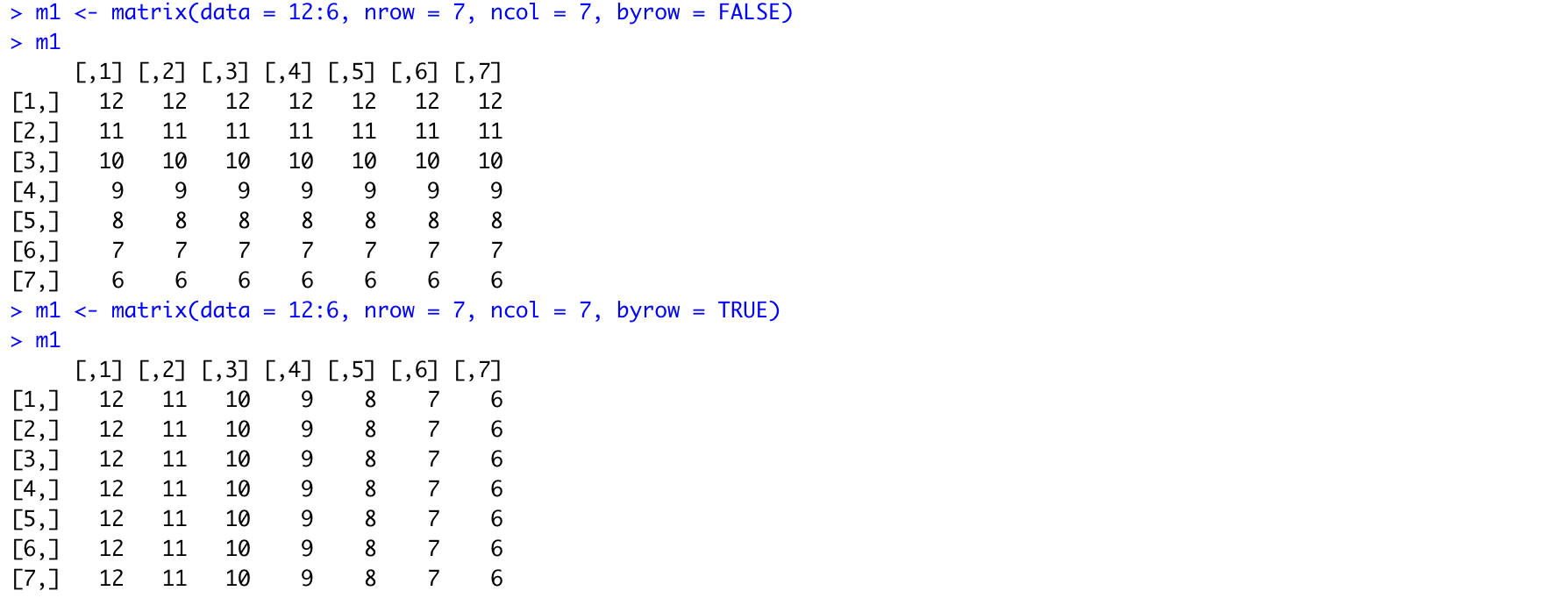


Вектор создаётся при помощи функции `c()`, которая объединяет несколько однотипных элементов:





Матрицы создаются при помощи функции `matrix()`:



Получим матрицу при помощи функции-комбинатора `cbind()`, которая объединяет столбцы:

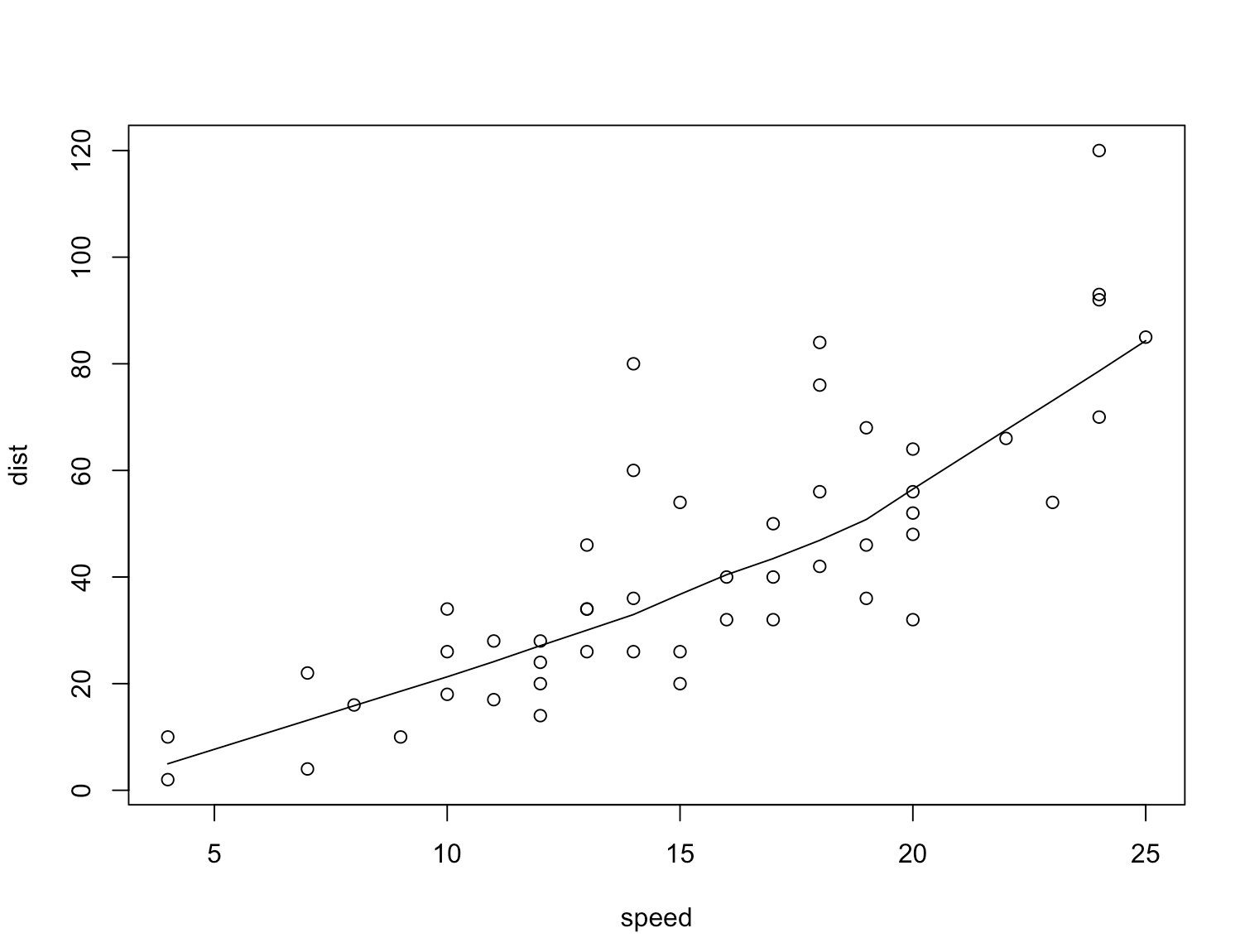


Получим тип элементов матрицы, класс объекта, а также проверим, является ли указанный объект матрицей и получим её размерность при помощи следующих функций:



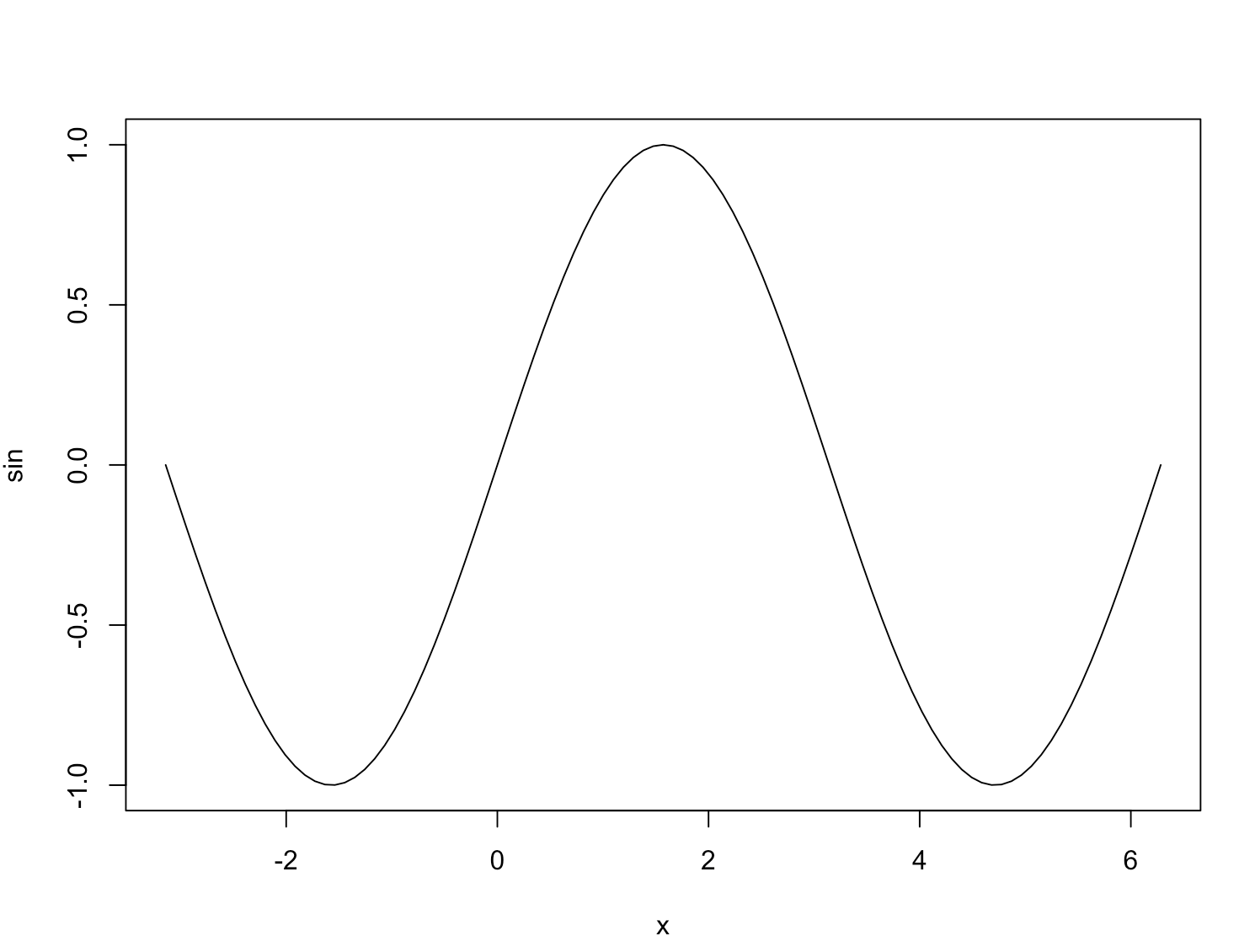
Построим график, используя функцию `plot()`; для этого необходимо подключить заранее установленный пакет в рабочее окружение (при помощи функции `require()`). Функция `lines(lowess())` проведёт функциональную линию по координатам сглаживания графика (см. рисунок 5).



  
Рисунок 5 – Использование функции plot()

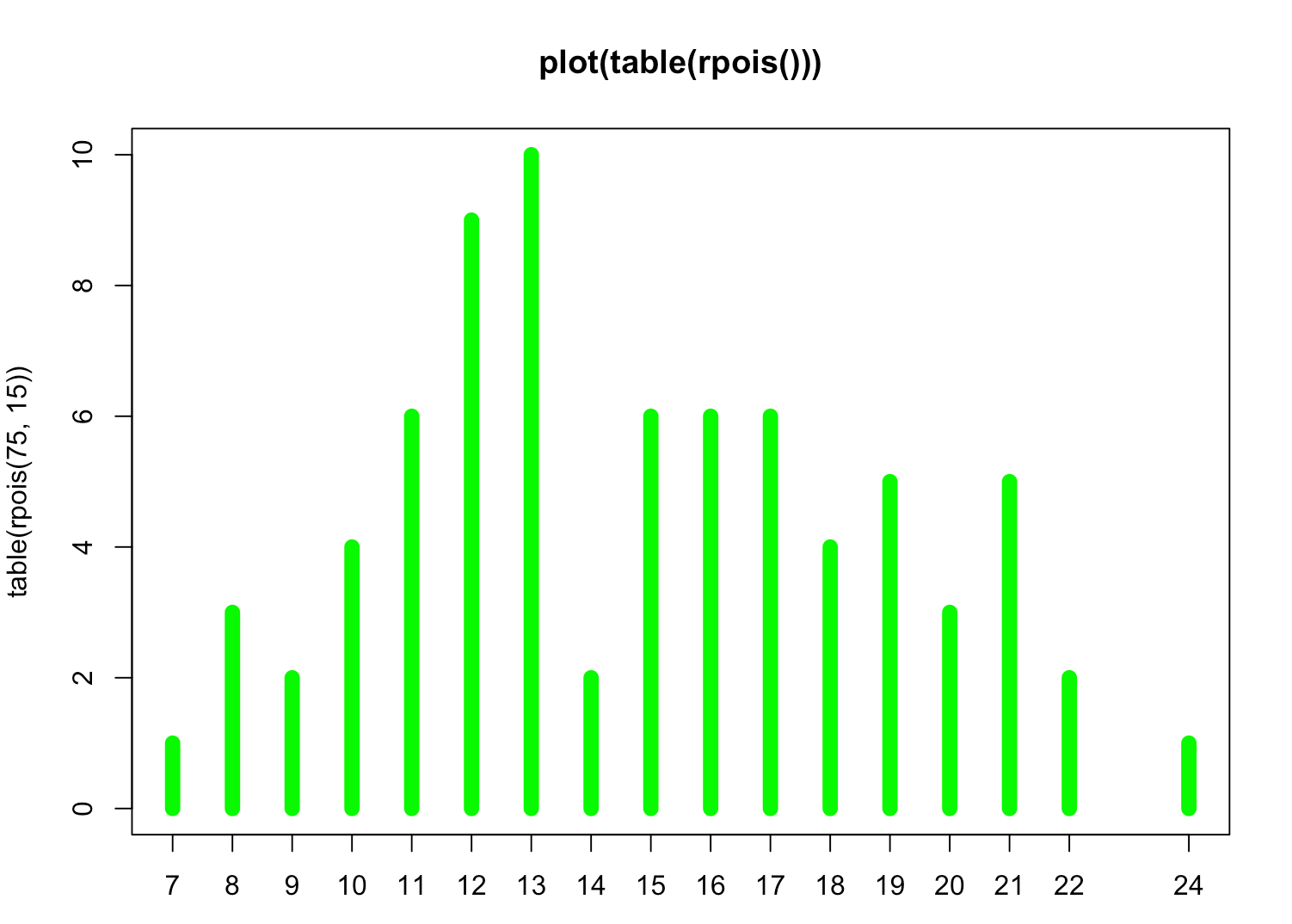
Построим график синусоиды (см. рисунок 6).



  
Рисунок 6 – График синуса

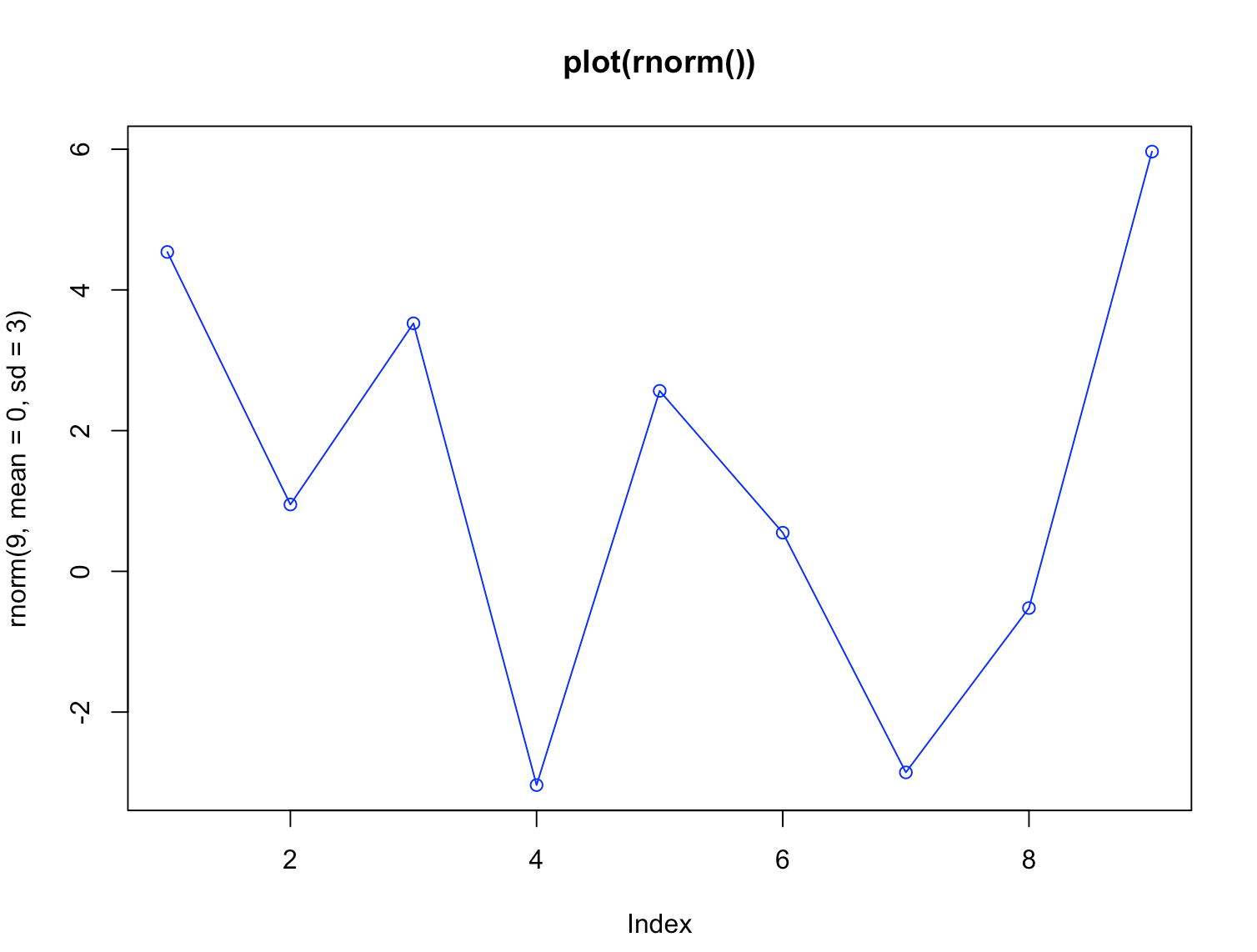
Построим график дискретного пуассоновского распределения при помощи `plot(table(rpois()))` – обобщённой функции графика для непредвиденных объектов таблицы (см. рисунок 7).



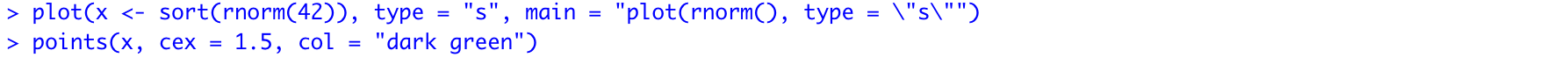
  
Рисунок 7 – График дискретного распределения Пуассона

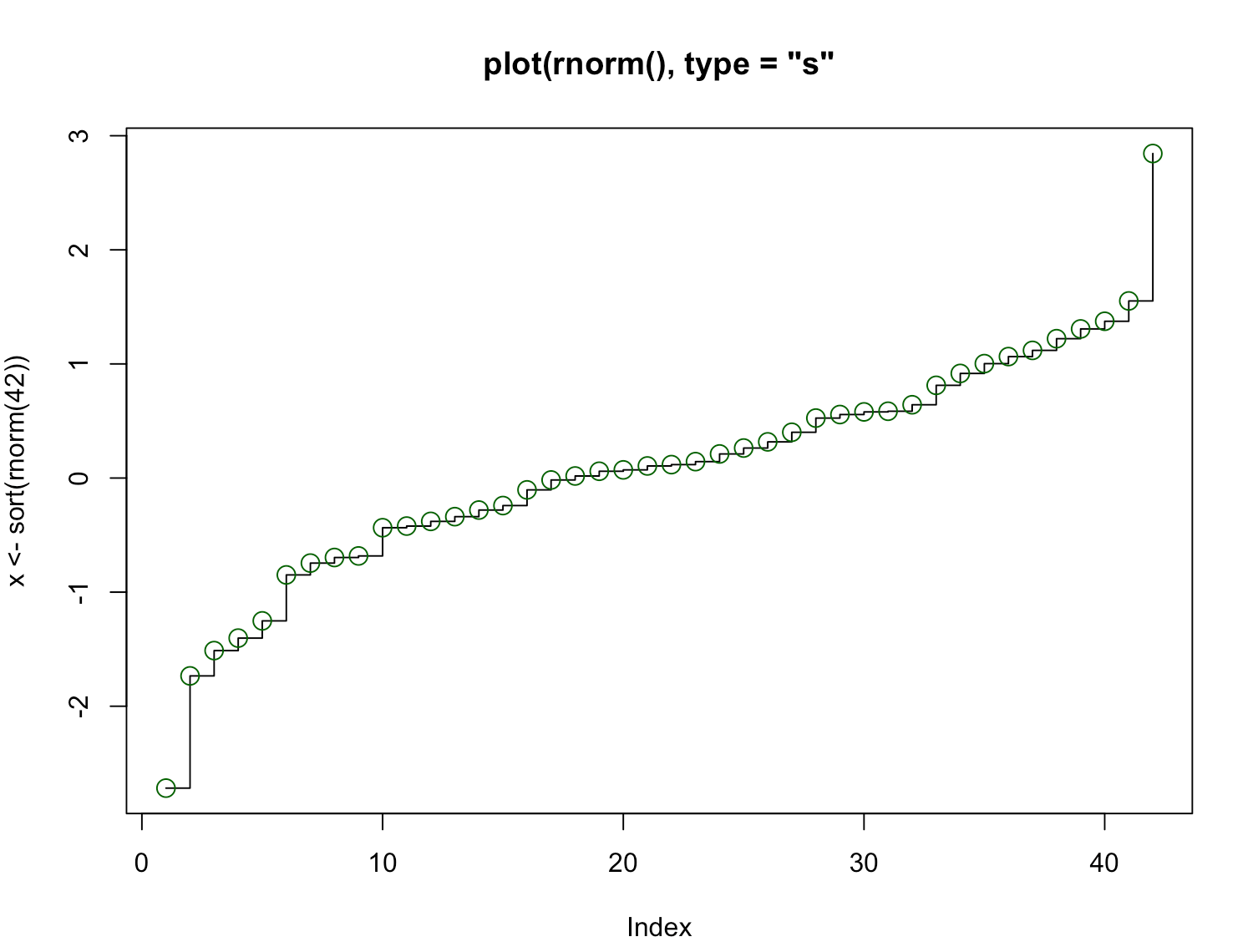
Функция `rnorm()` служит для случайной генерации совокупностей нормально распределённых чисел. Построим её график (см. рисунок 8).



  
Рисунок 8 – График функции rnorm()

Построим график типа «ступенчатая кривая» (см. рисунок 9).



  
Рисунок 9 – График отсортированной генерации совокупностей нормально распределённых чисел

1. Ответим на контрольные вопросы.

* Особенности языка R:

R — один из главных инструментов для анализа данных. Он создан статистиками для статистиков и необходим для качественной визуализации данных.

Это интерпретируемый объектно-ориентированный язык программирования. То есть с приложениями работает не ядро операционной системы, а программа-интерпретатор, поэтому разработчику перед запуском не нужно компилировать код в исполняемый файл.

* Команда для получения подробной информации о функции в R:

`help(<название функции>)` или `? <название функции>`

* Структура и особенности команды `round()` в R:

`round()` округляет значение, переданное в качестве её первого аргумента до указанного числа знаков (по умолчанию 0).

Имеет два аргумента: x – значение, которое необходимо округлить, digits – количество знаков до которых произвести округление.



* Команды для работы с векторами в R:

Создадим вектор при помощи функции `scan()`, которая "считывает" последовательно вводимые с клавиатуры значения:



Упорядочим значения вектора по убыванию при помощи функцию `sort()` в сочетании с аргументом decreasing = TRUE:



Получим краткую статистическую сводку по вектору при помощи функции `summary()`:

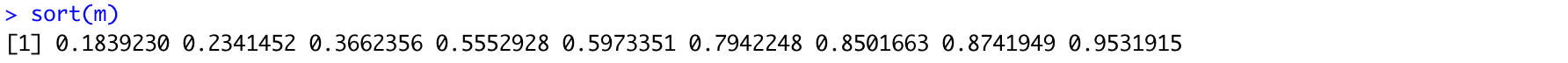


* Команды для работы с матрицами в R:

Создадим матрицу, заполнив её случайными величинами при помощи функции `runif()`:



Отсортируем матрицу (возвращается обычный вектор):



Транспонируем матрицу:



Получим определитель матрицы:

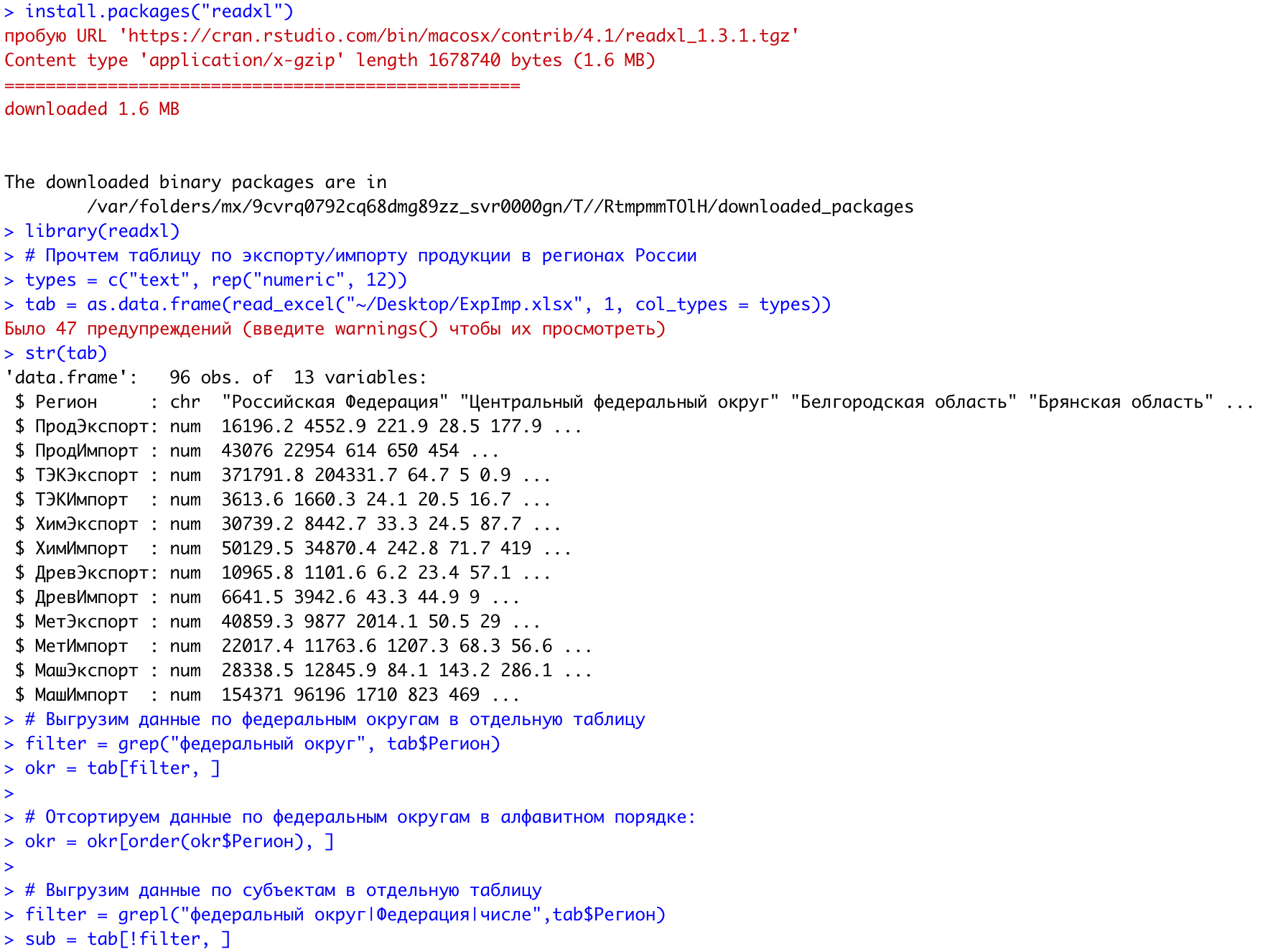


Получим краткую статистическую сводку по матрице при помощи функции `summary()`:



* Работа с графикой в R:

Подготовим исходные данные, мы будем работать с региональной статистикой Росстата: экспорт/импорт продукции по регионам России (млн дол. США) и объем сброса сточных вод по морям России (млрд м3):



Построим диаграмму рассеяния, она позволяет установить, есть ли зависимость между переменными, а также понять, как объекты дифференцируются по значениям переменных (см. рисунок 10):

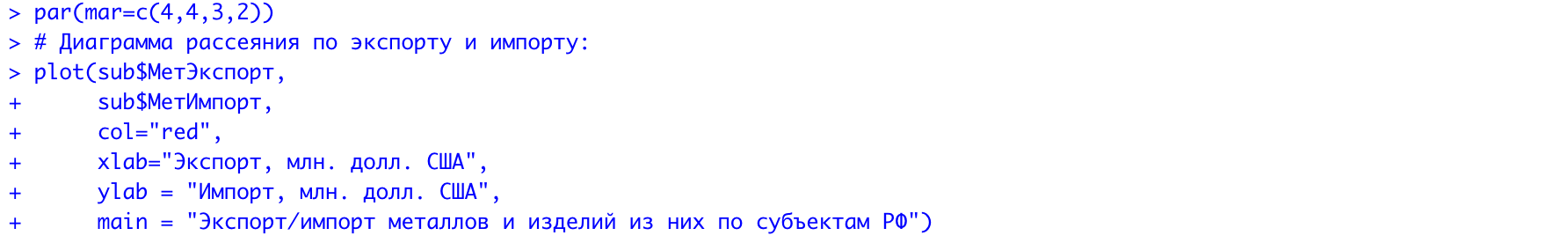


Рисунок 10 – Диаграмма рассеивания

Построим гистограмму распределения (см. рисунок 11):



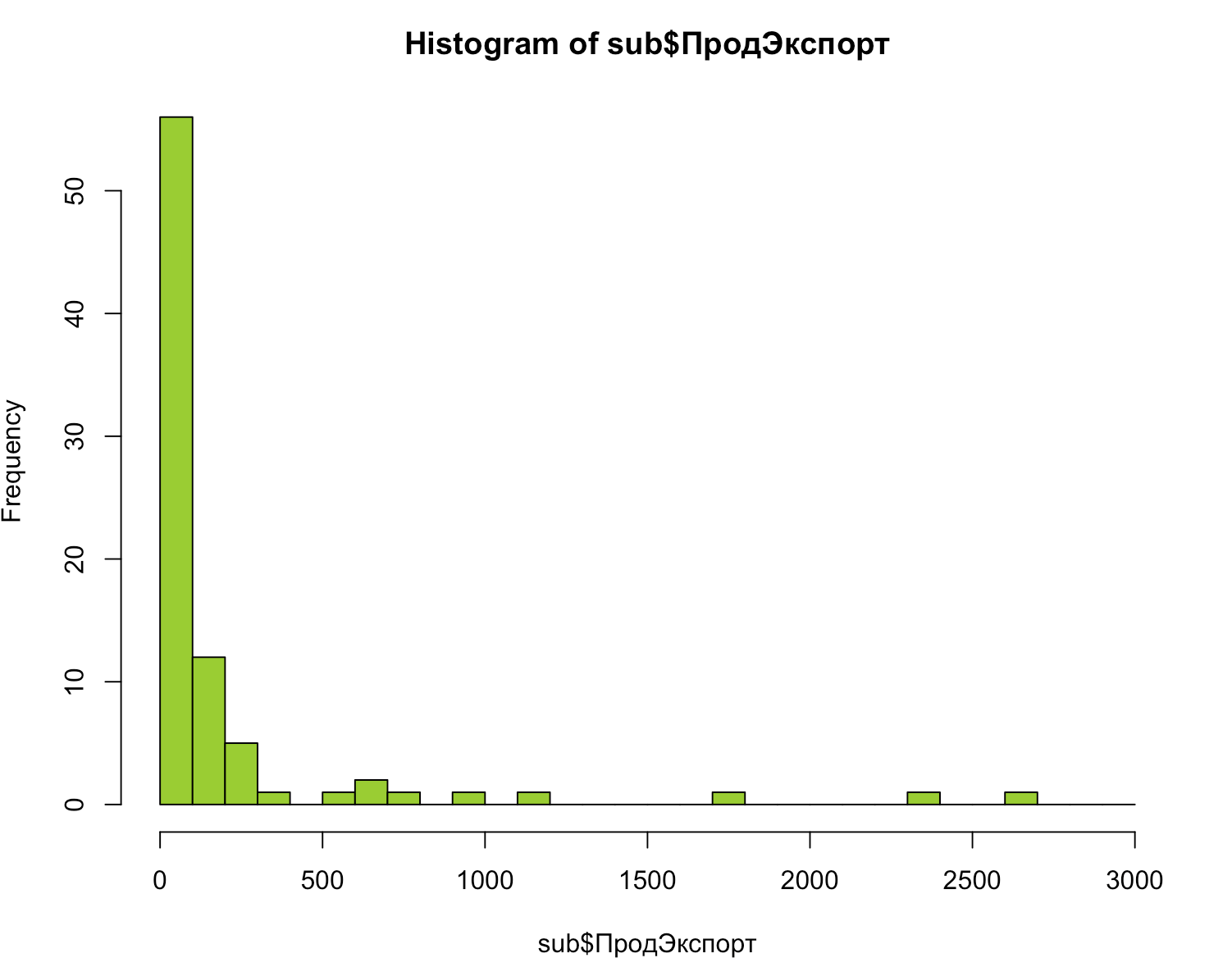


Рисунок 11 – Гистограмма распределения

## Выводы

В ходе лабораторной работы была изучена статистическая система анализа R, являющаяся и языком, и программным обеспечением. Был изучен синтаксис языка R, его особенности и основные функции, такие как создание матриц, построение графиков и настройка их отображения, работа с векторами и т.п.