РОССИЙСКИЙ УНИВЕРСИТЕТ ДРУЖБЫ НАРОДОВ

**Факультет физико-математических и естественных наук**

**Кафедра прикладной информатики и теории вероятностей**

ОТЧЕТ

по лабораторной работе № 7

дисциплина: Компьютерный практикум   
по математическому моделированию

Студент: Каримли Нуран

Группа: НК - 401

**МОСКВА**

2022 г.

**Лабораторная работа №7. Введение в работу с данными**

**Цель работы:**

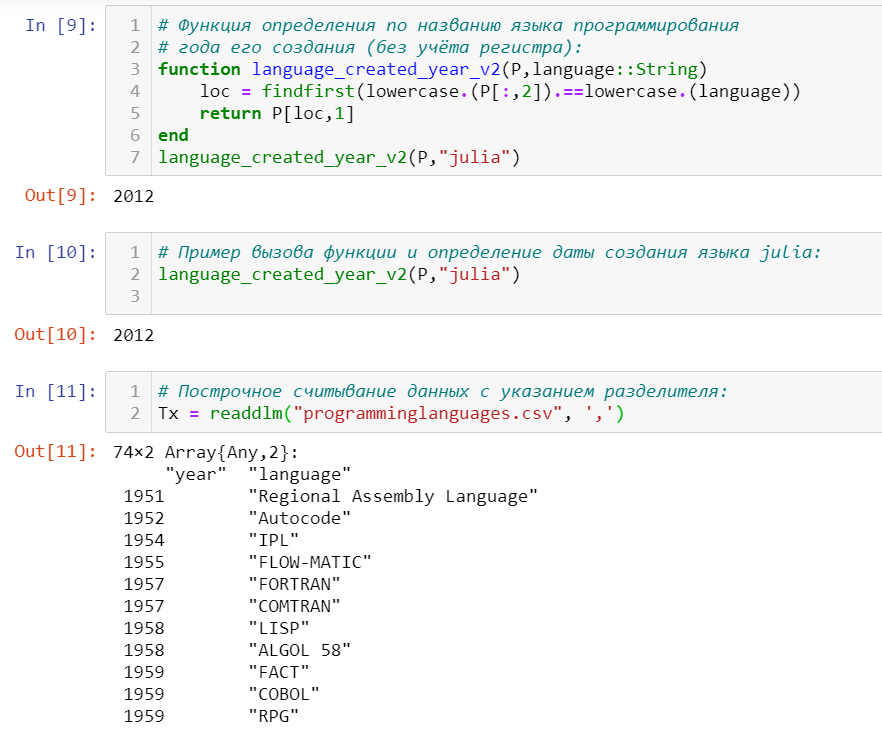
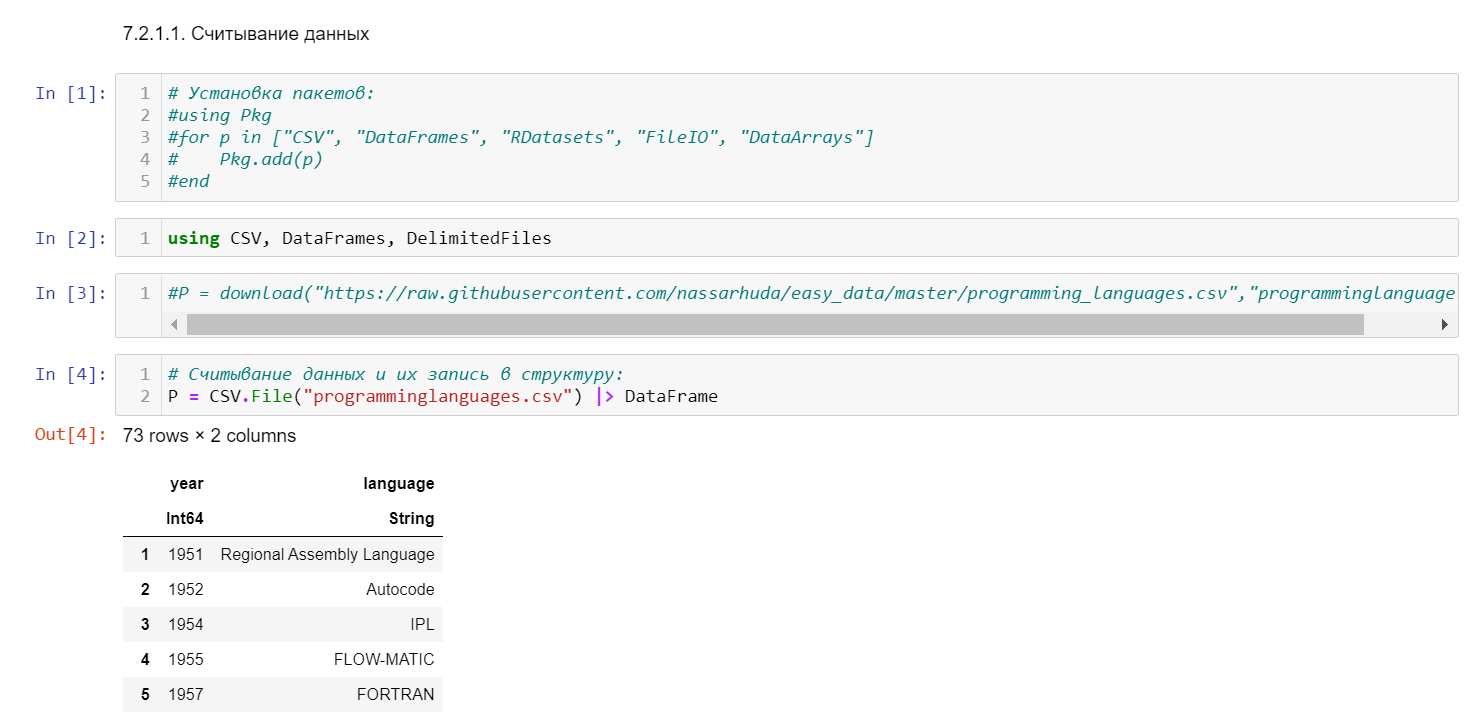
Основной целью работы является специализированных пакетов Julia для обработки данных.

**Ход работы:**

**Повторила примеры из раздела 7.2**

7.2.1. Julia для науки о данных

7.2.1.1. Считывание данных



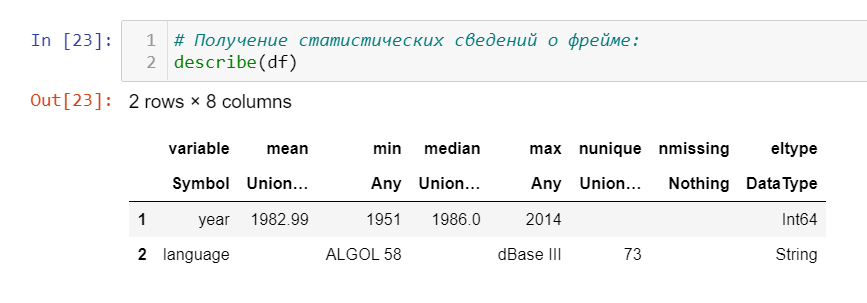
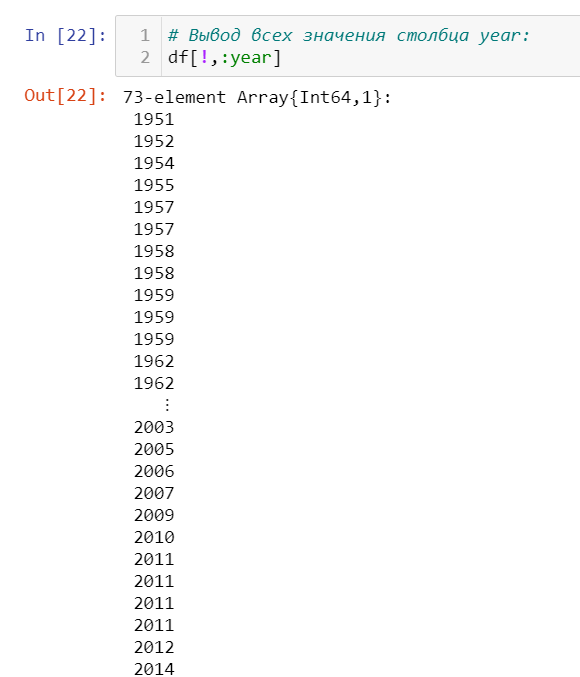
7.2.1.2. Запись данных в файл



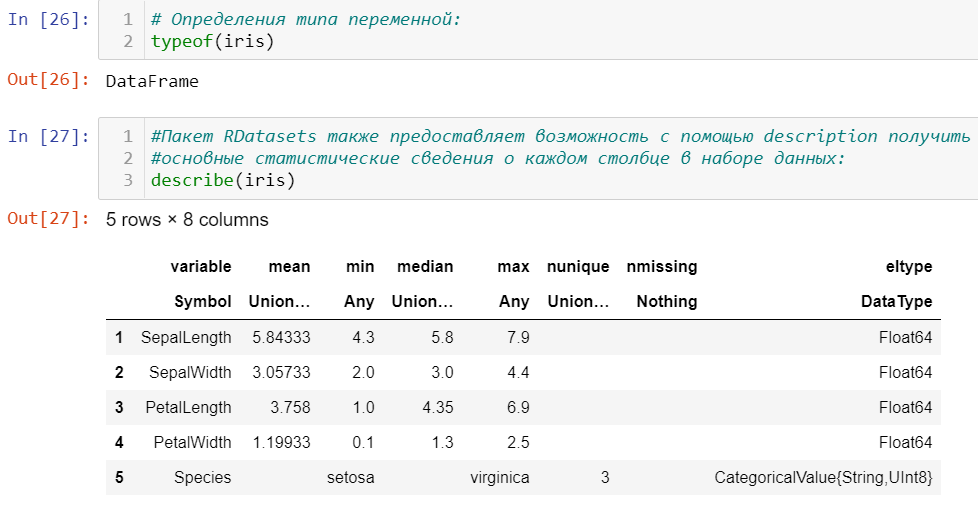
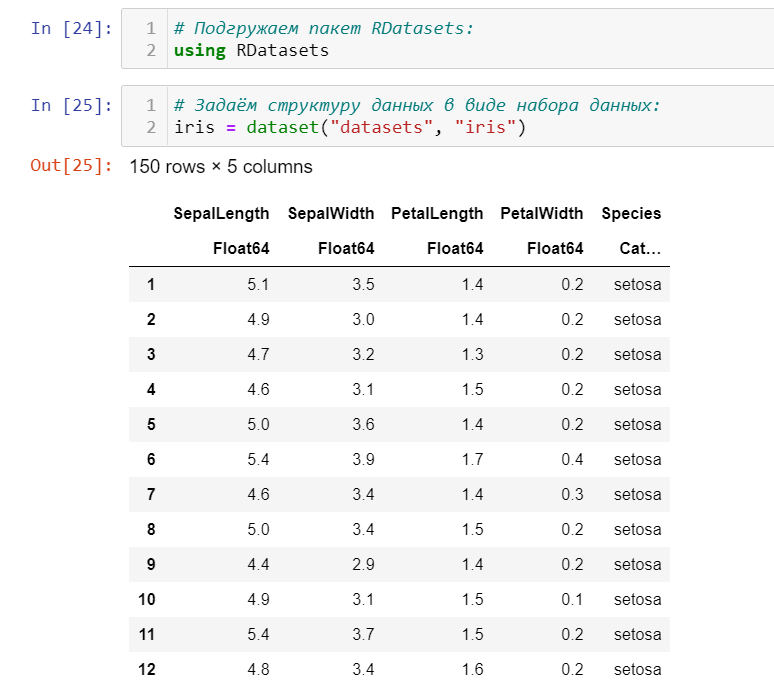
7.2.1.3. Словари



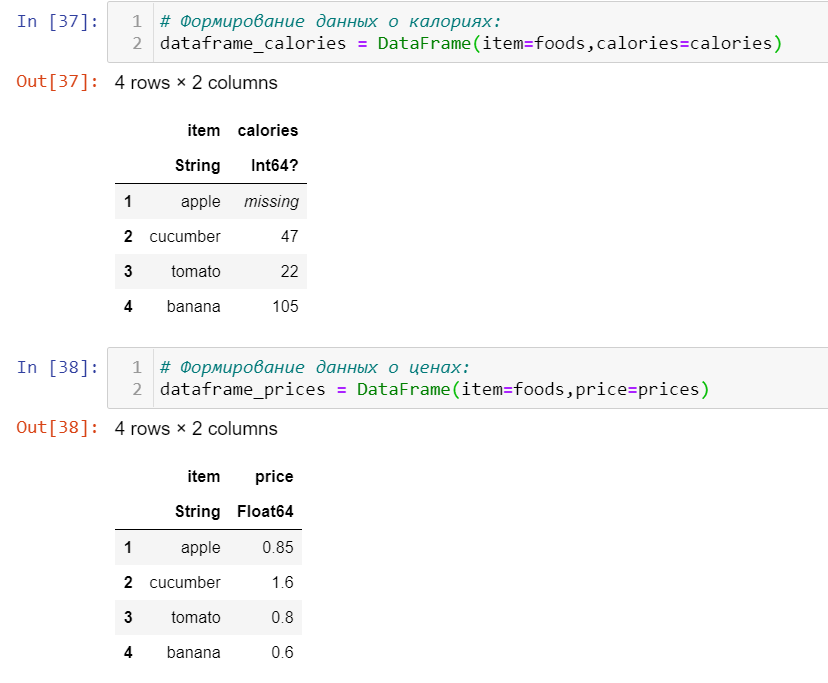
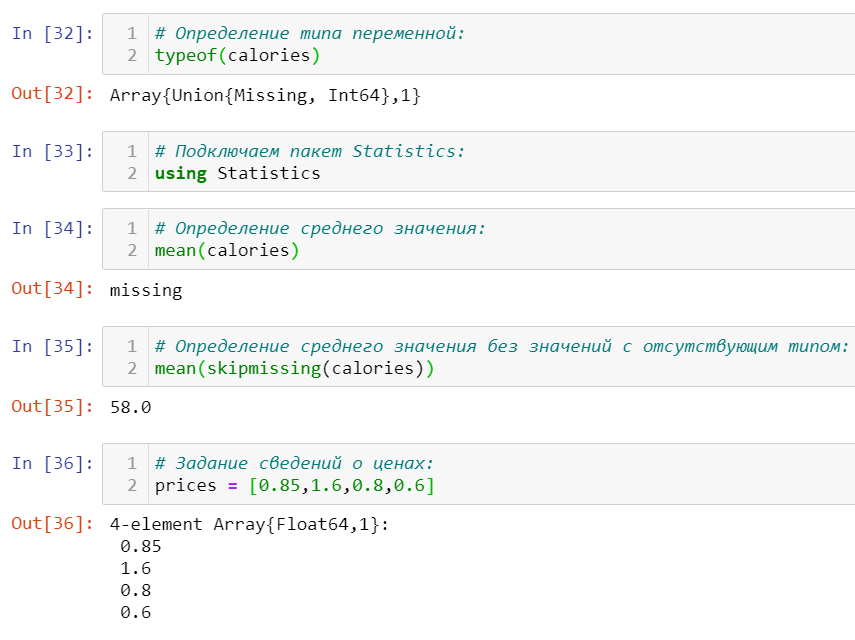
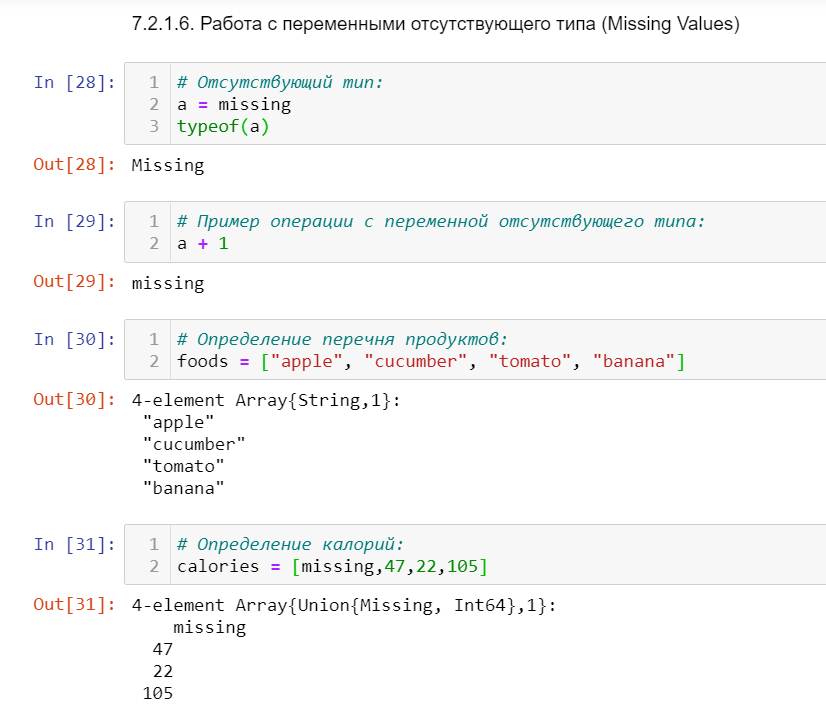
7.2.1.4. DataFrames



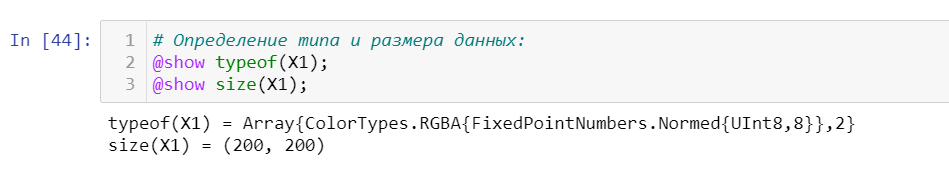
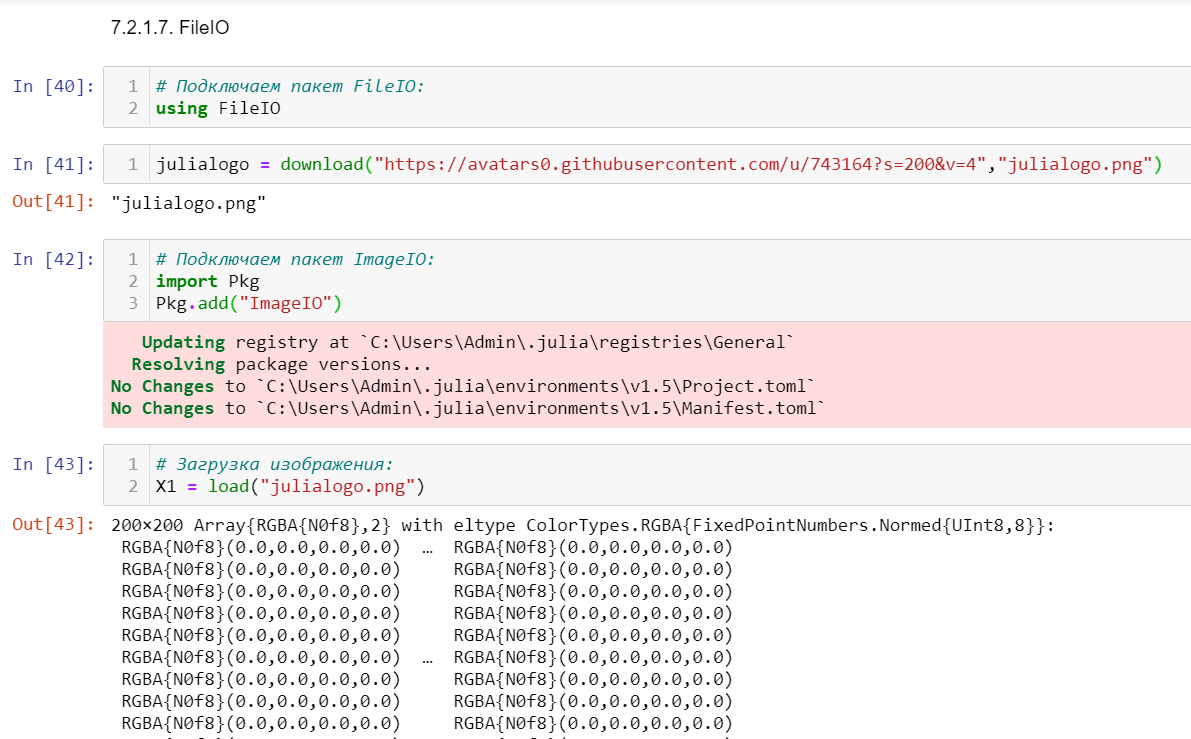
7.2.1.5. RDatasets



7.2.1.6. Работа с переменными отсутствующего типа (Missing Values)

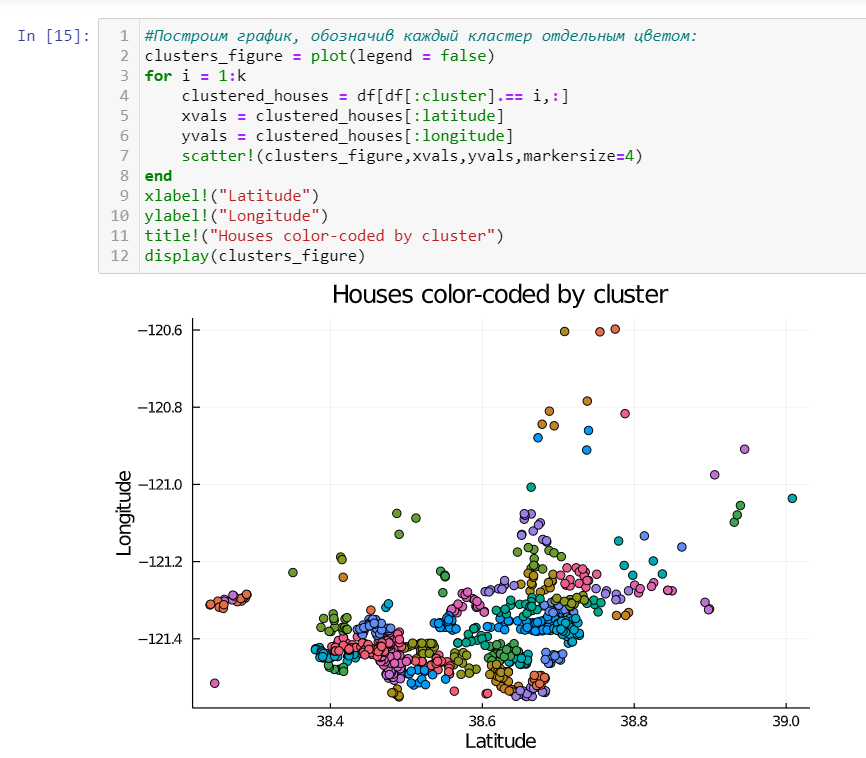
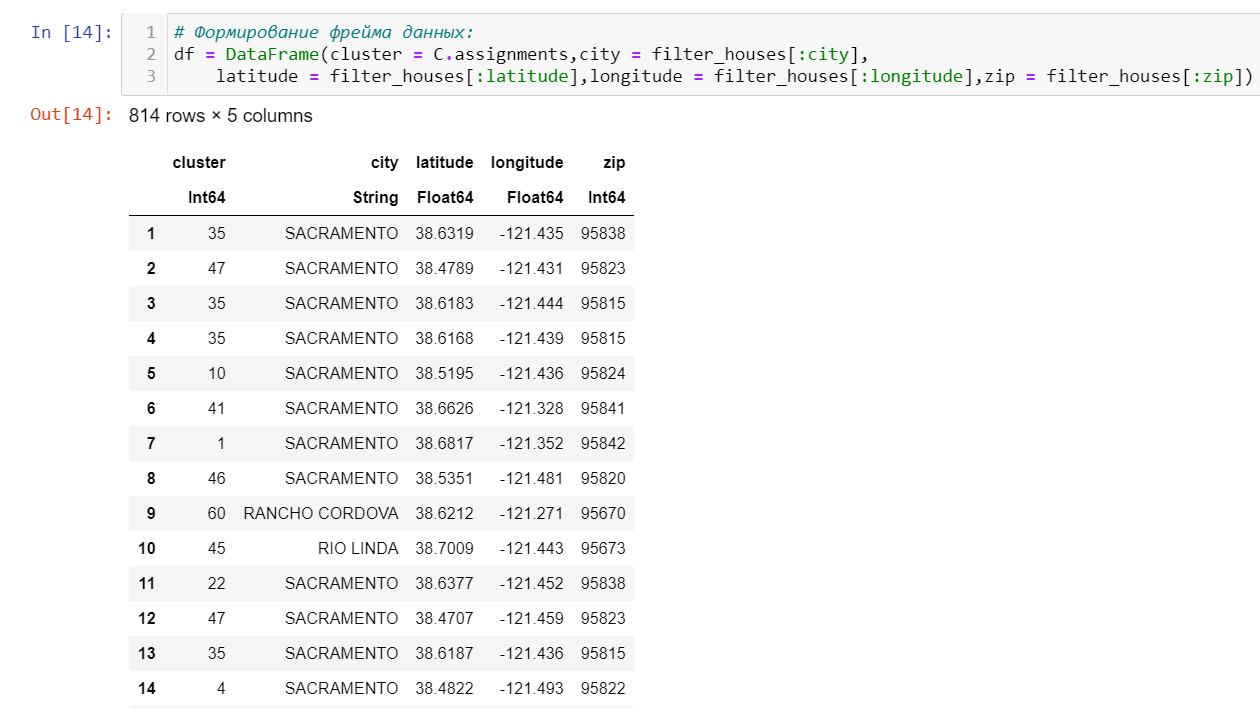
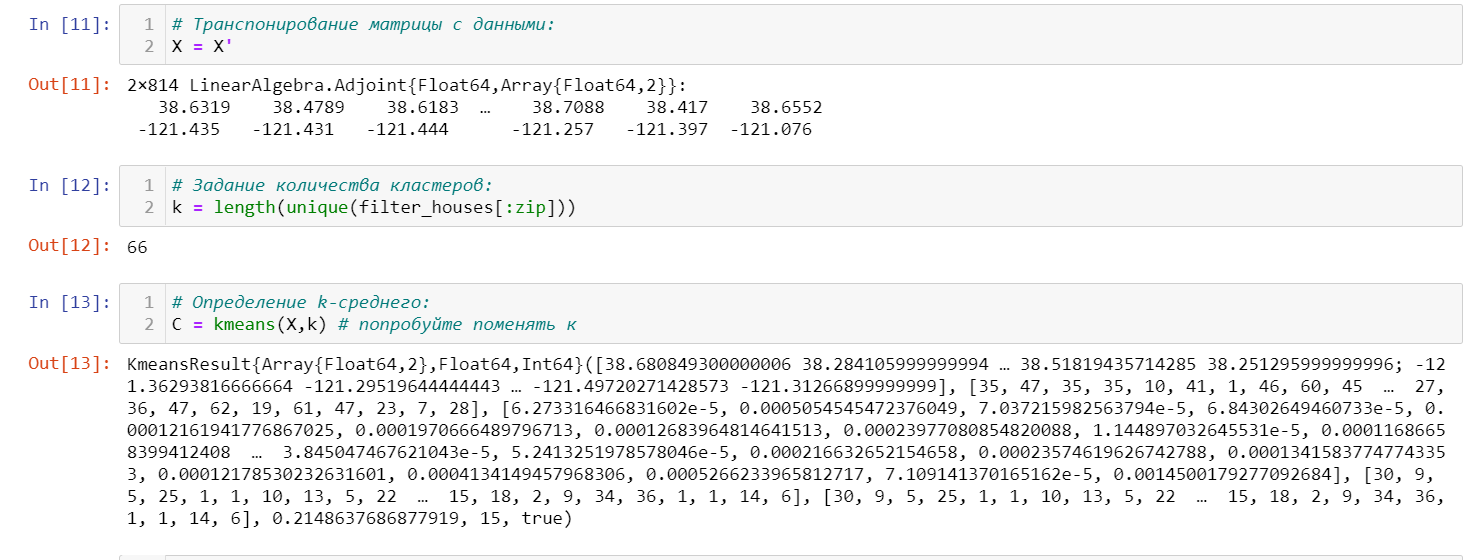
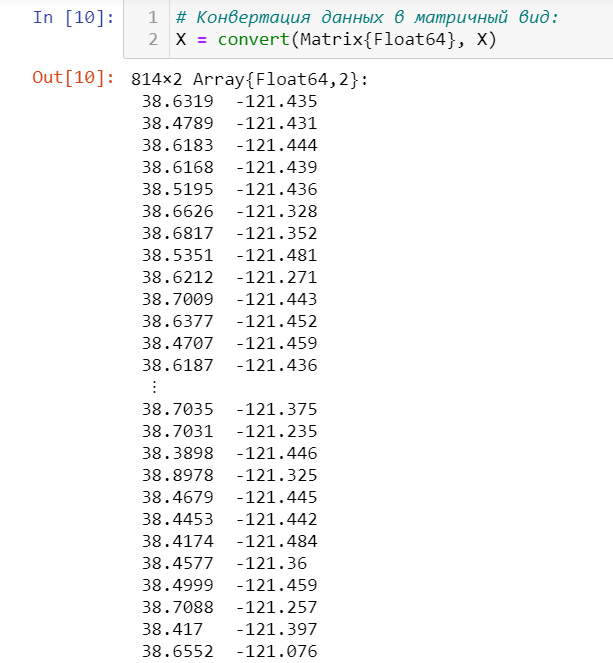
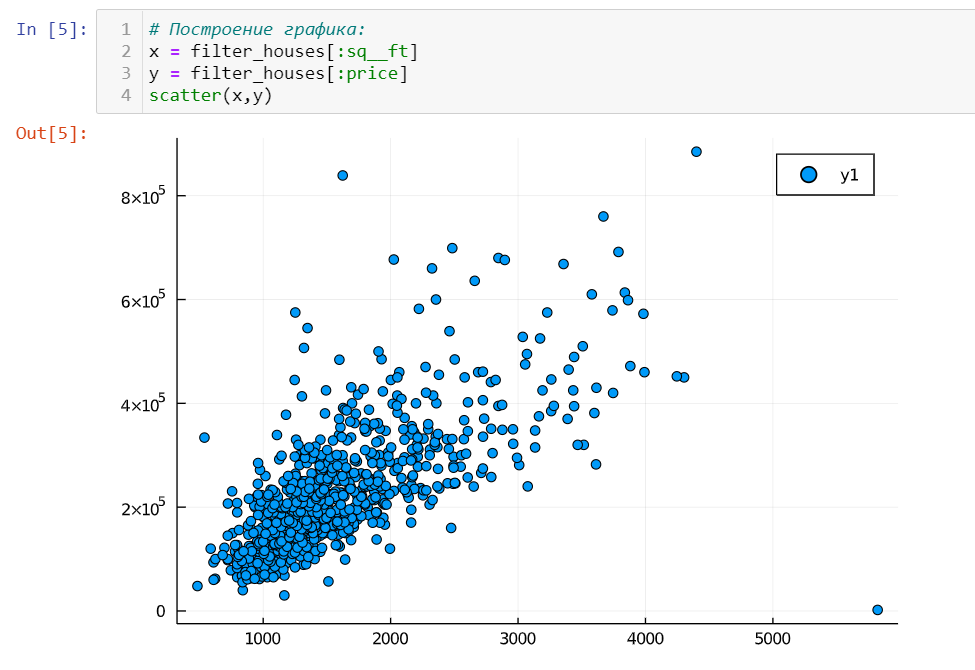
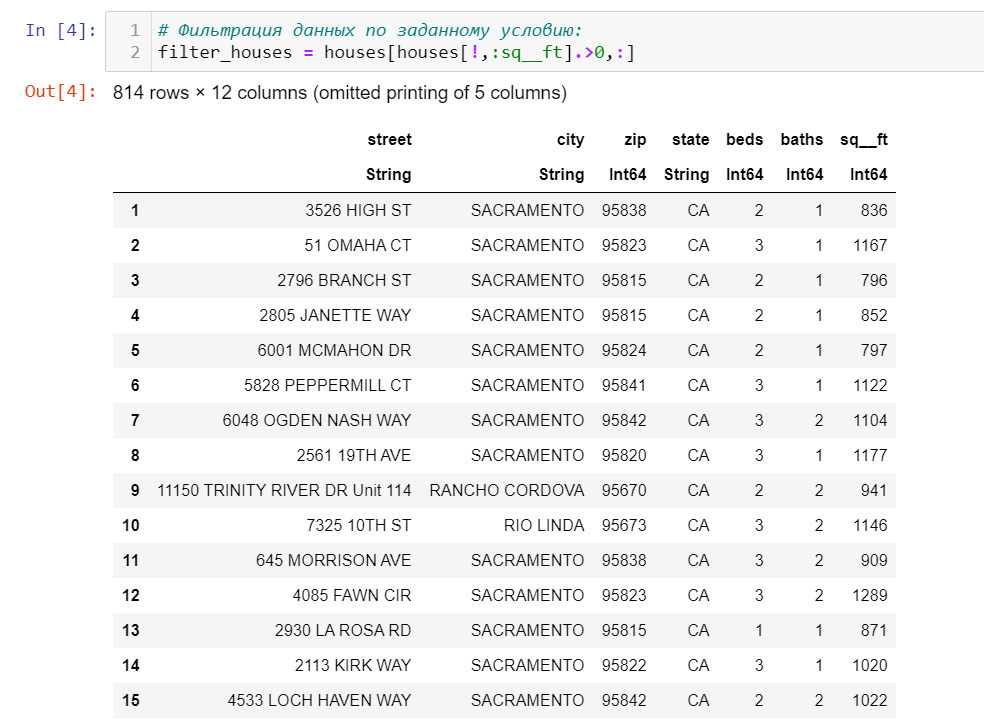
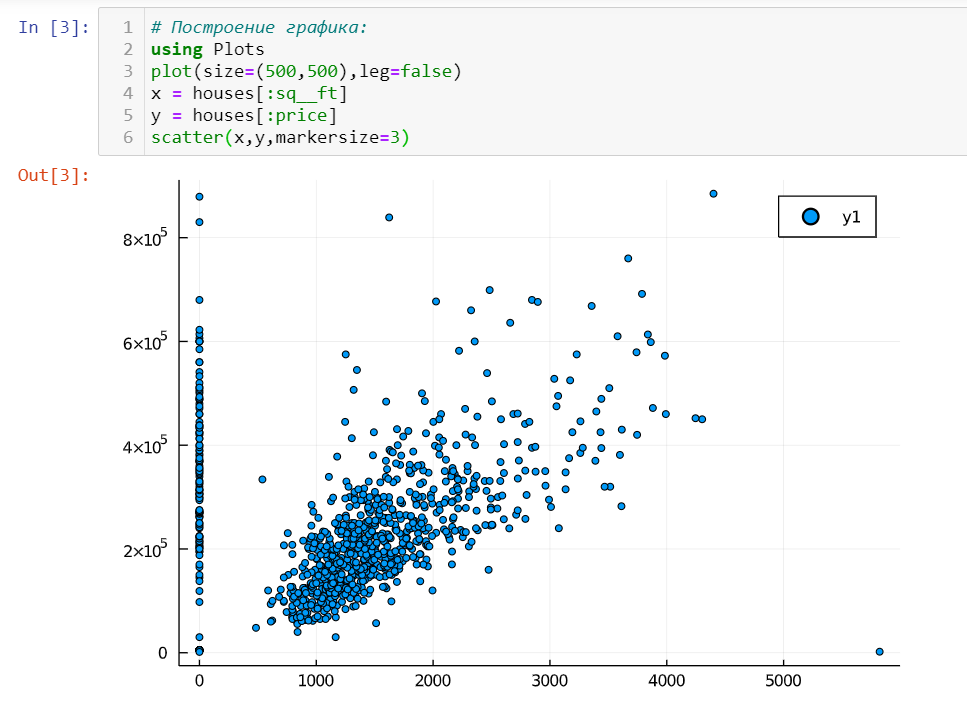
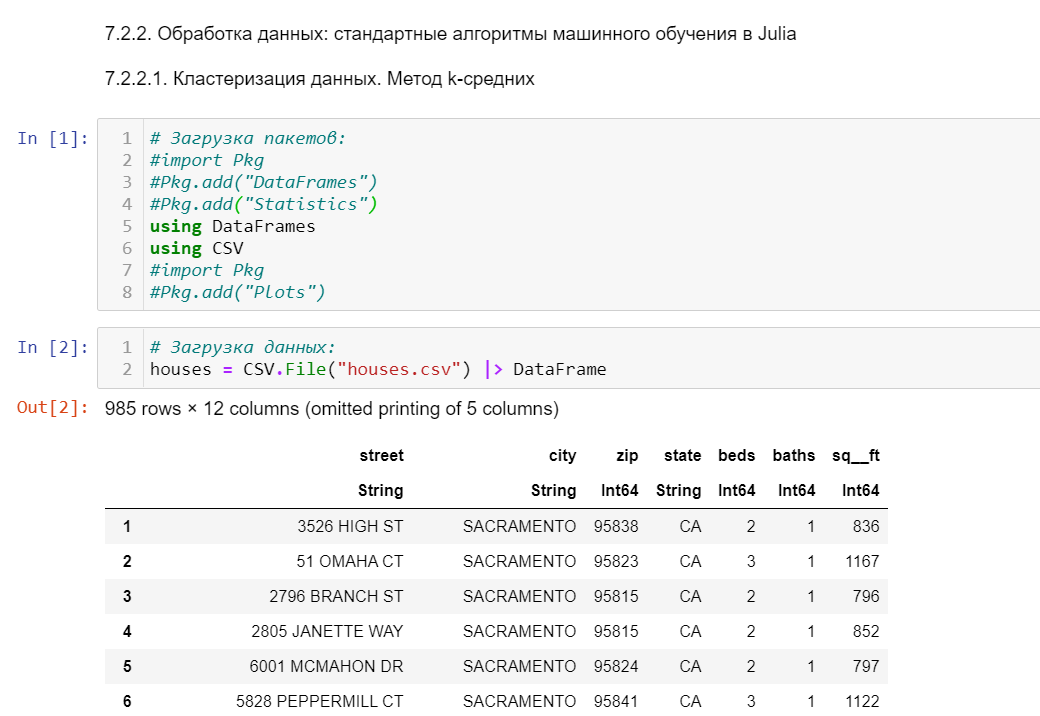


7.2.1.7. FileIO

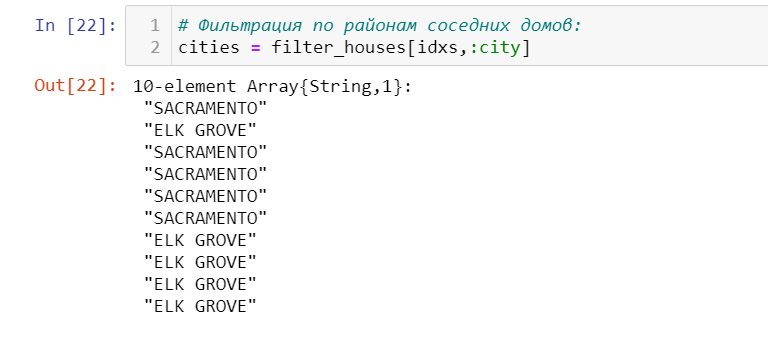
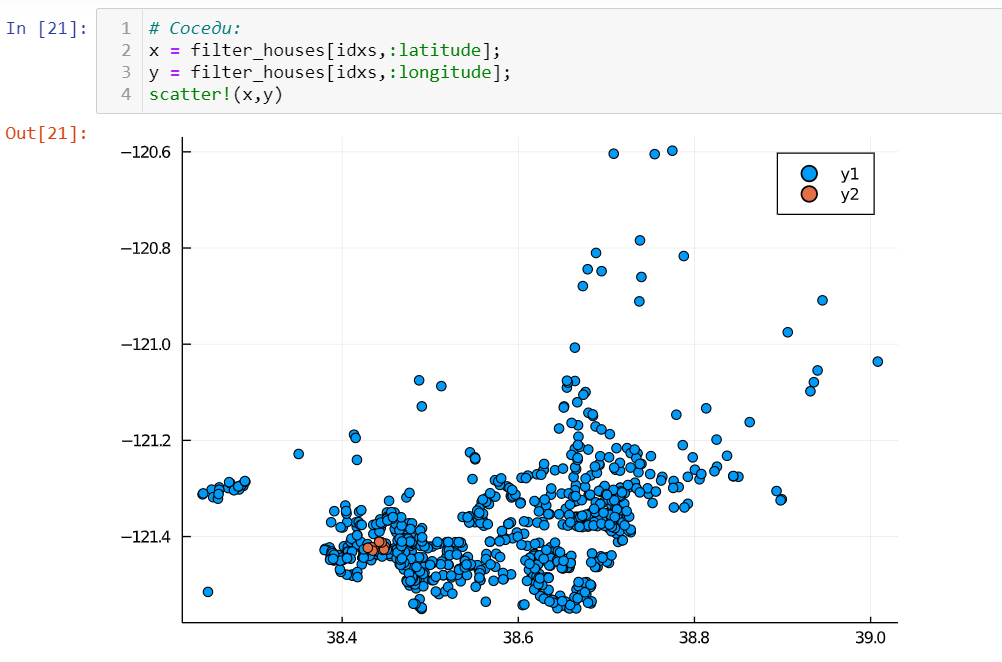
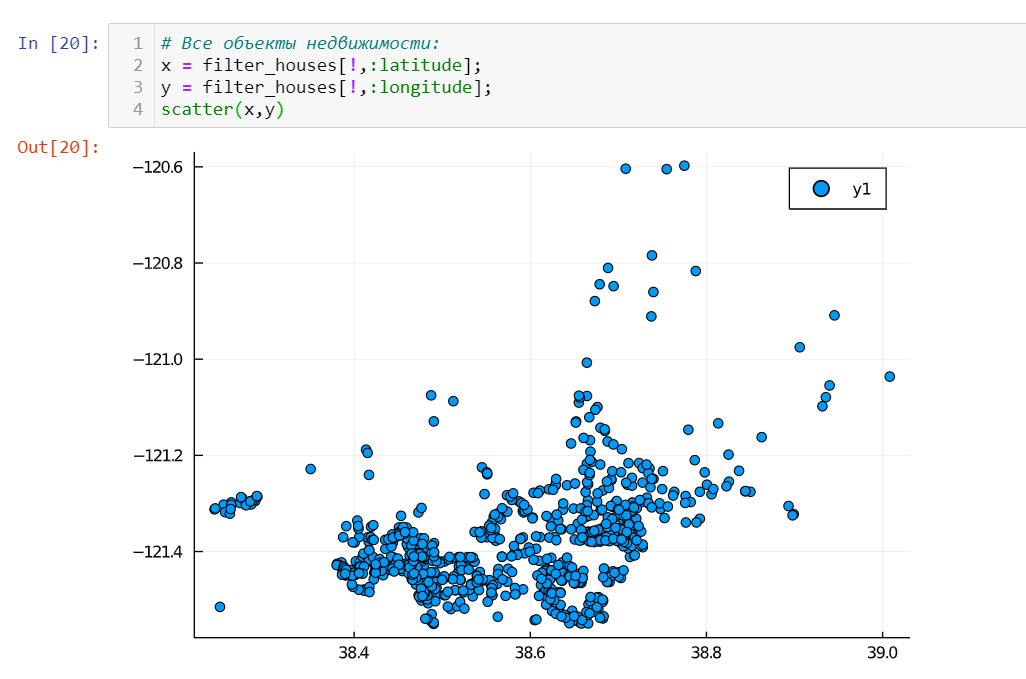
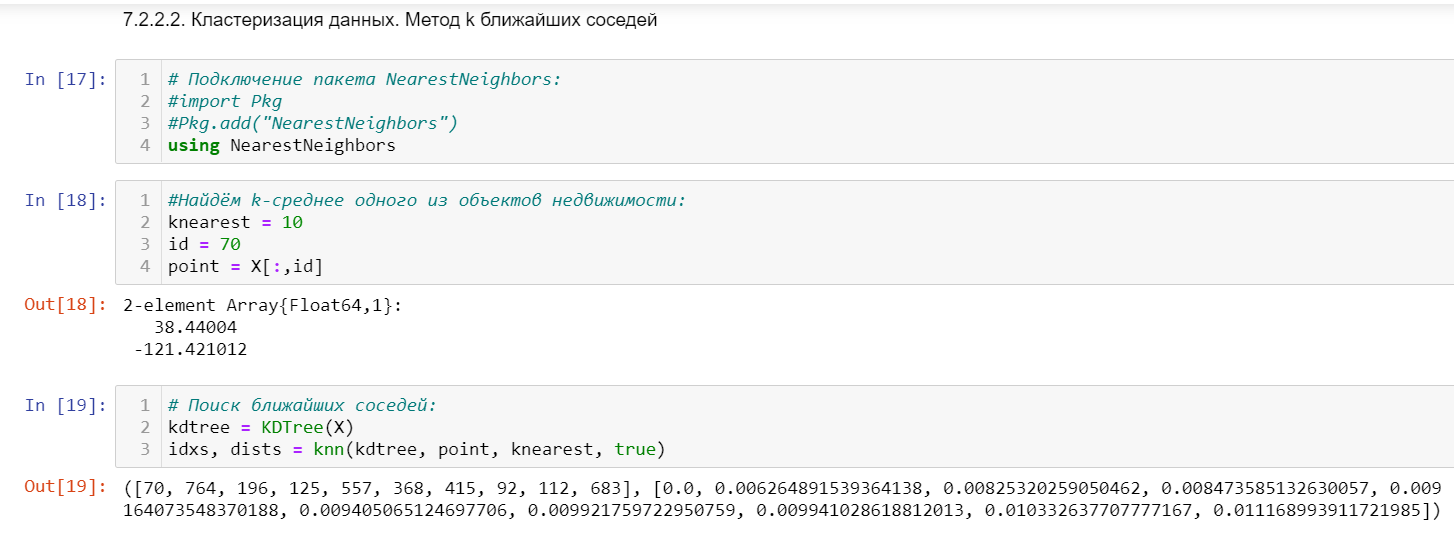


7.2.2. Обработка данных: стандартные алгоритмы машинного обучения в Julia

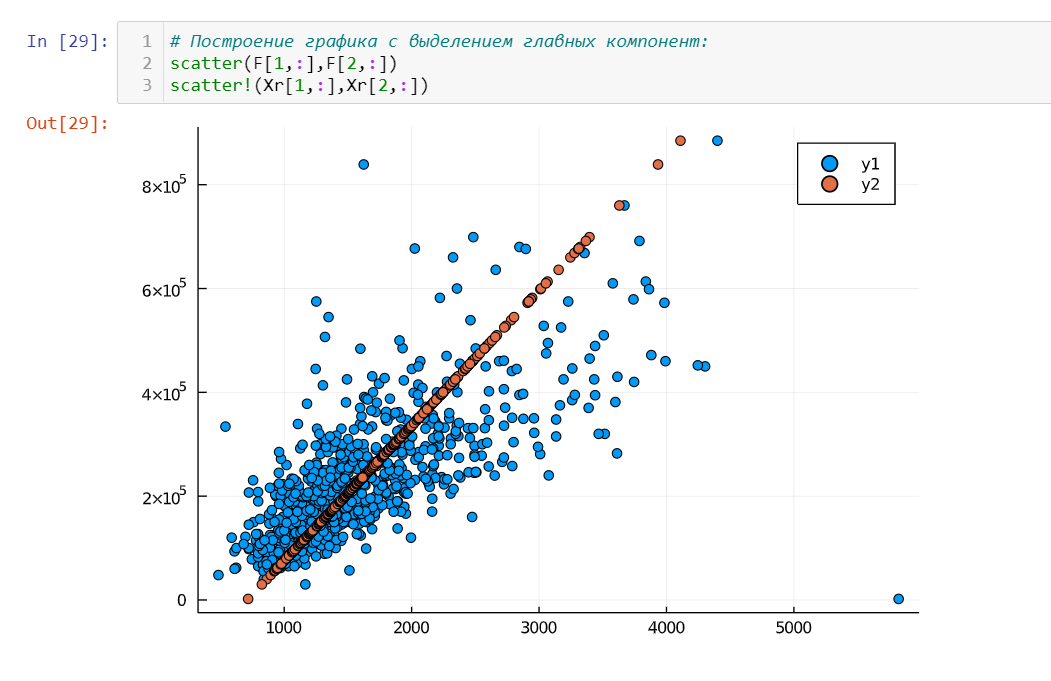
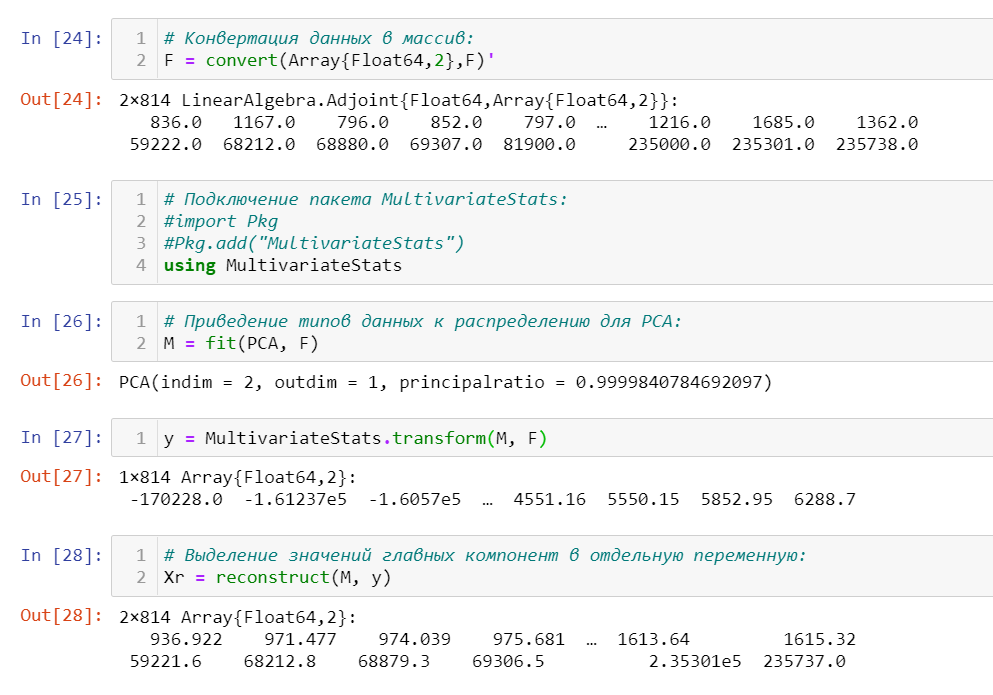
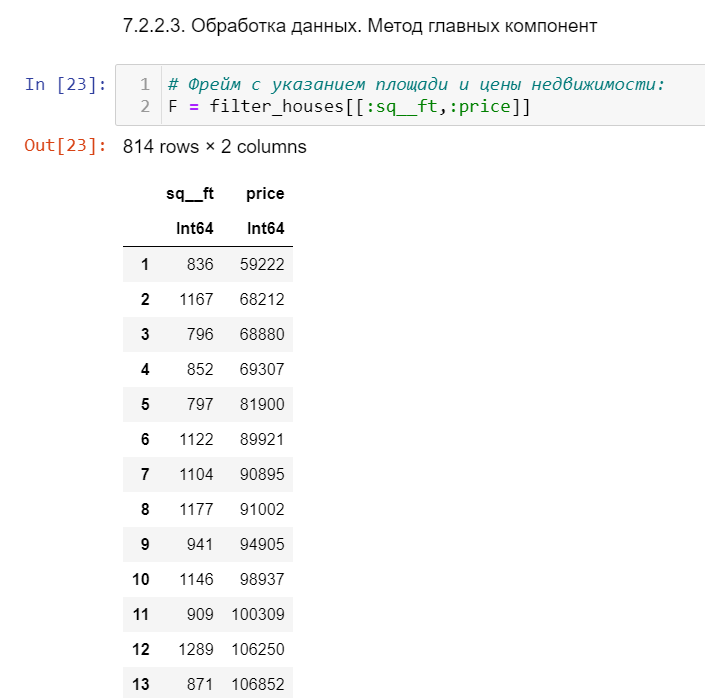
7.2.2.1. Кластеризация данных. Метод k-средних



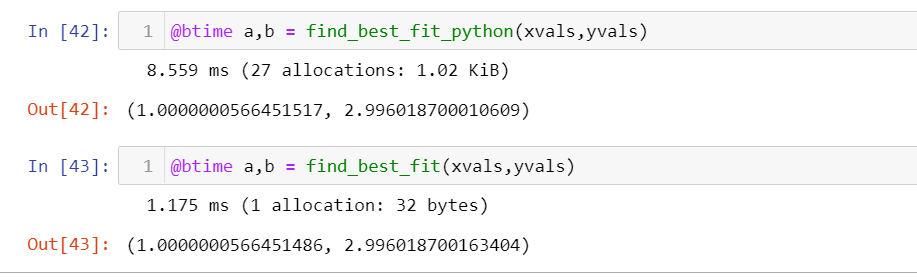
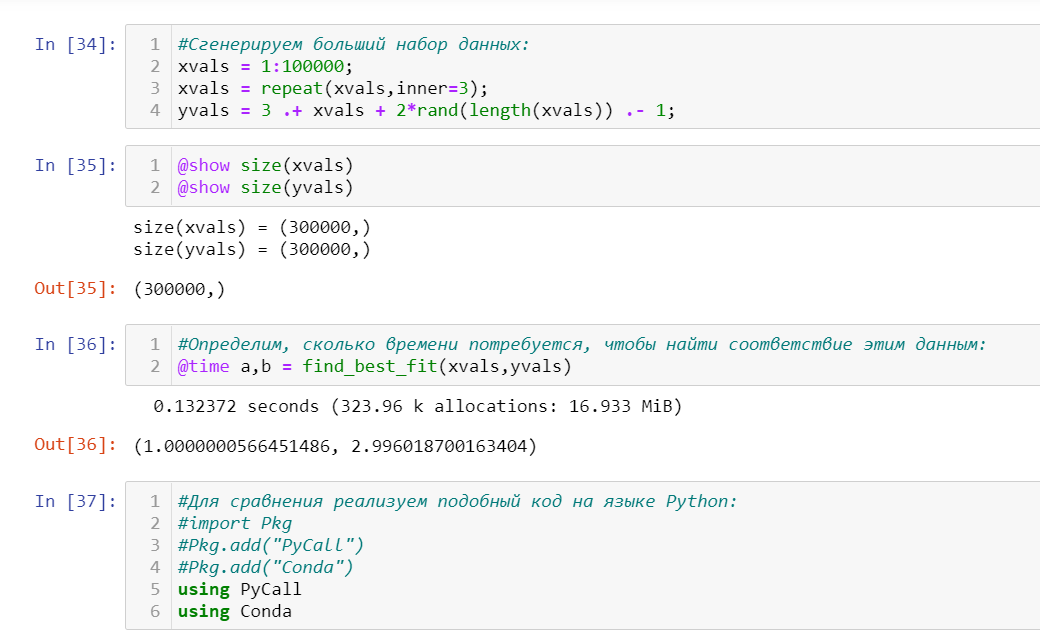
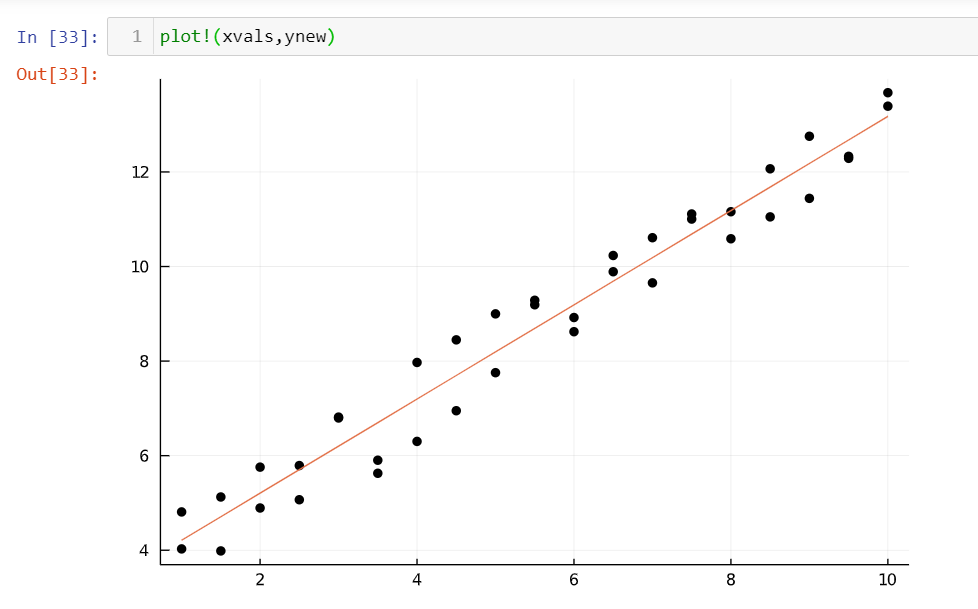
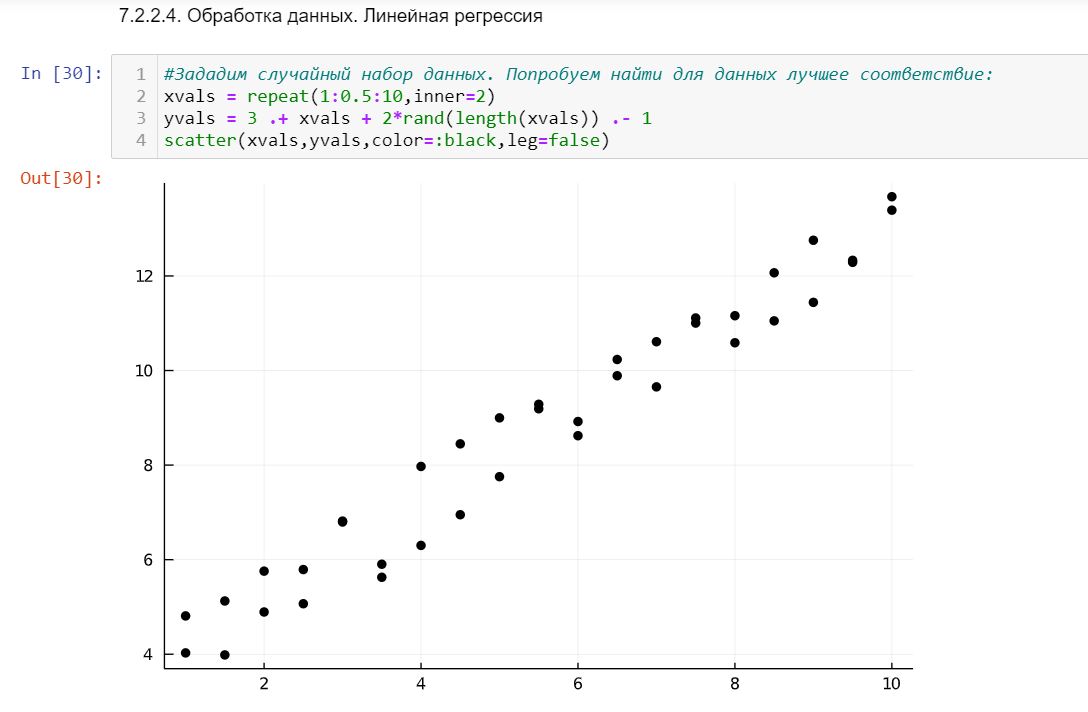
7.2.2.2. Кластеризация данных. Метод k ближайших соседей



7.2.2.3. Обработка данных. Метод главных компонент



7.2.2.4. Обработка данных. Линейная регрессия



**Задания для самостоятельного выполнения**

7.4.1. Кластеризация

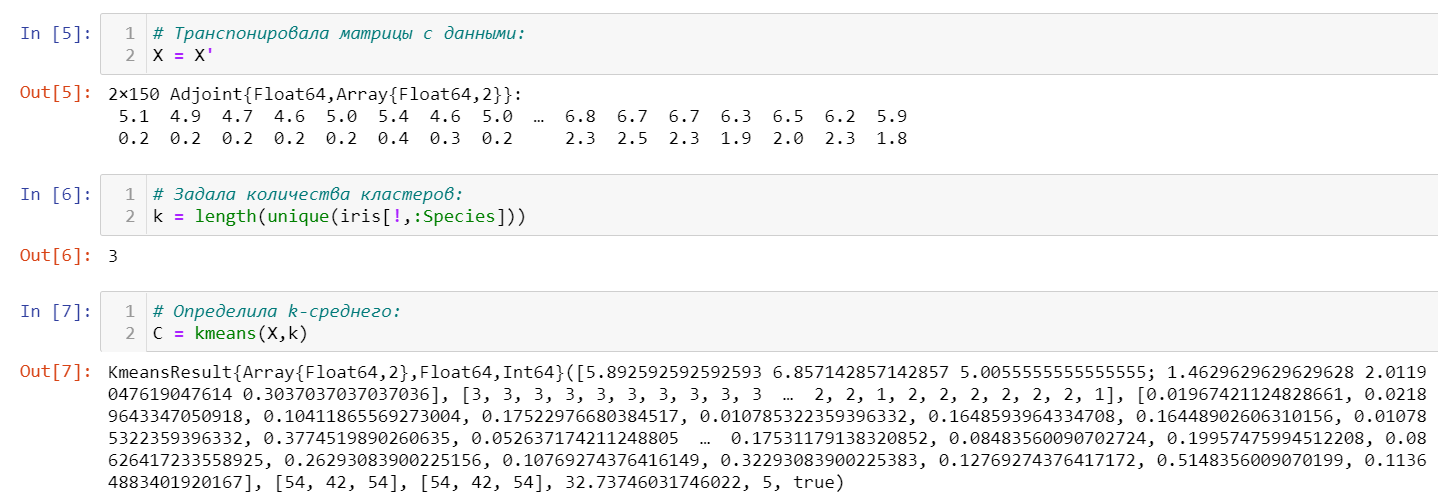
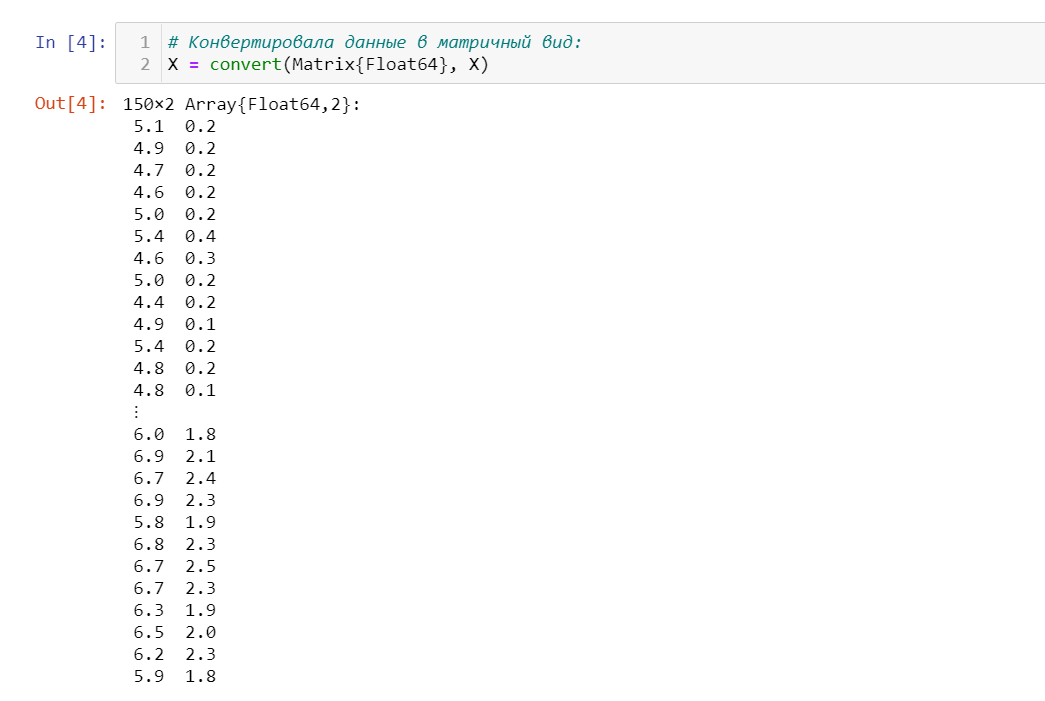
Загрузила

using RDatasets

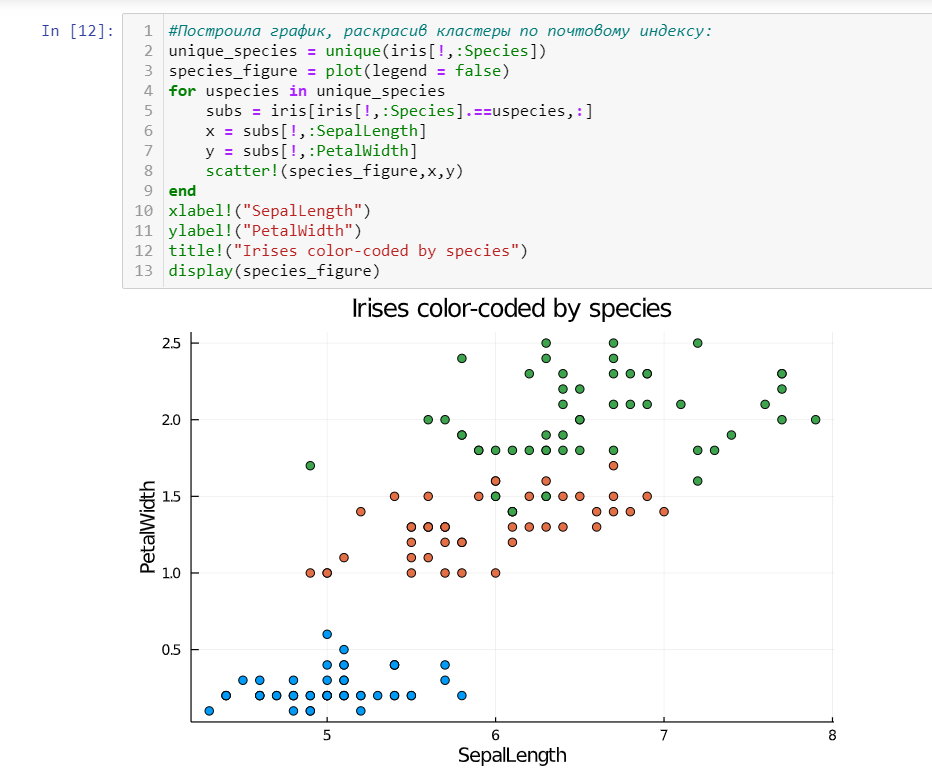
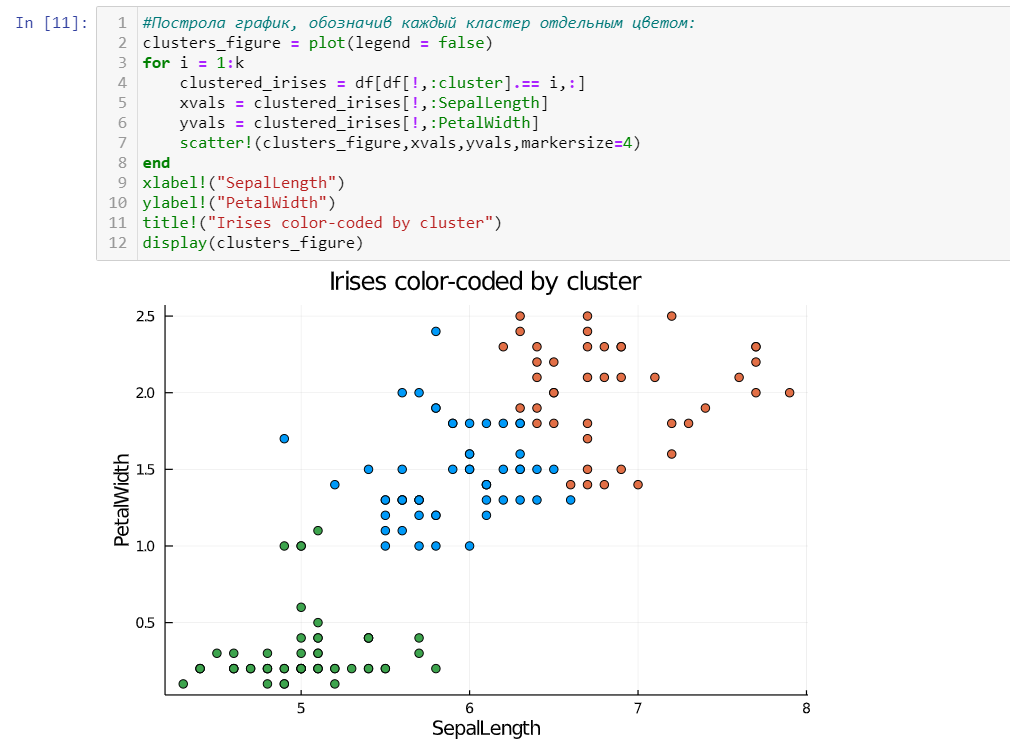
iris = dataset("datasets", "iris")



Использовала Clustering.jl для кластеризации на основе k-средних.



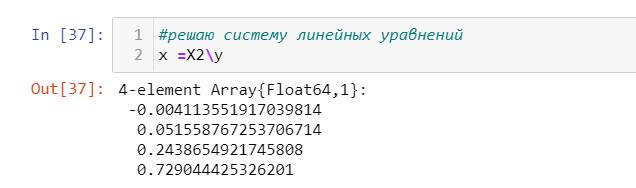
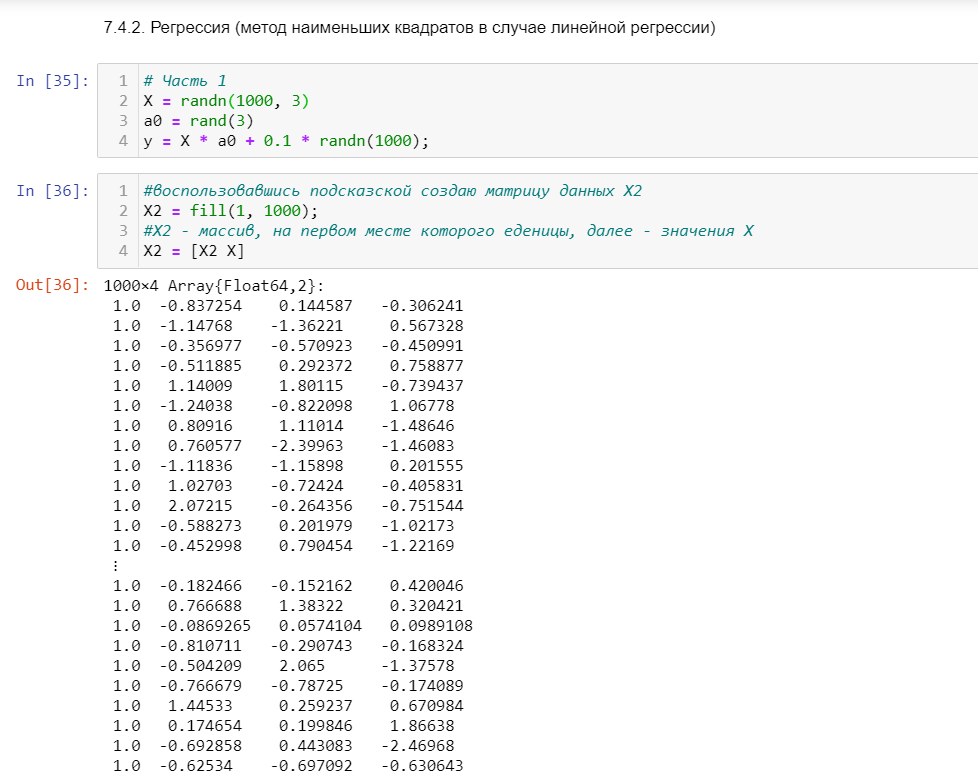
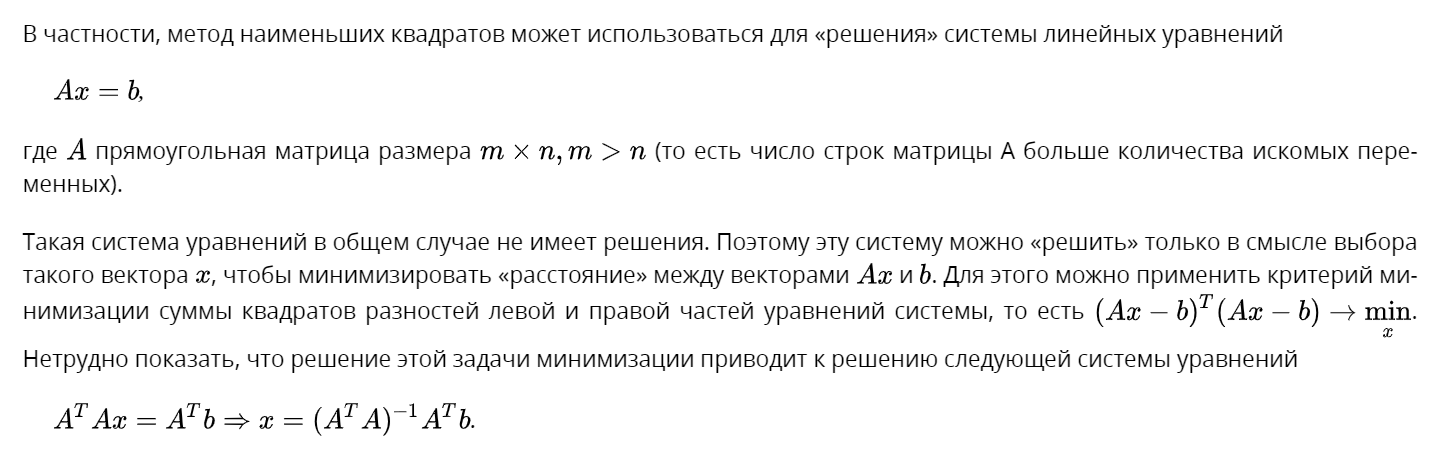
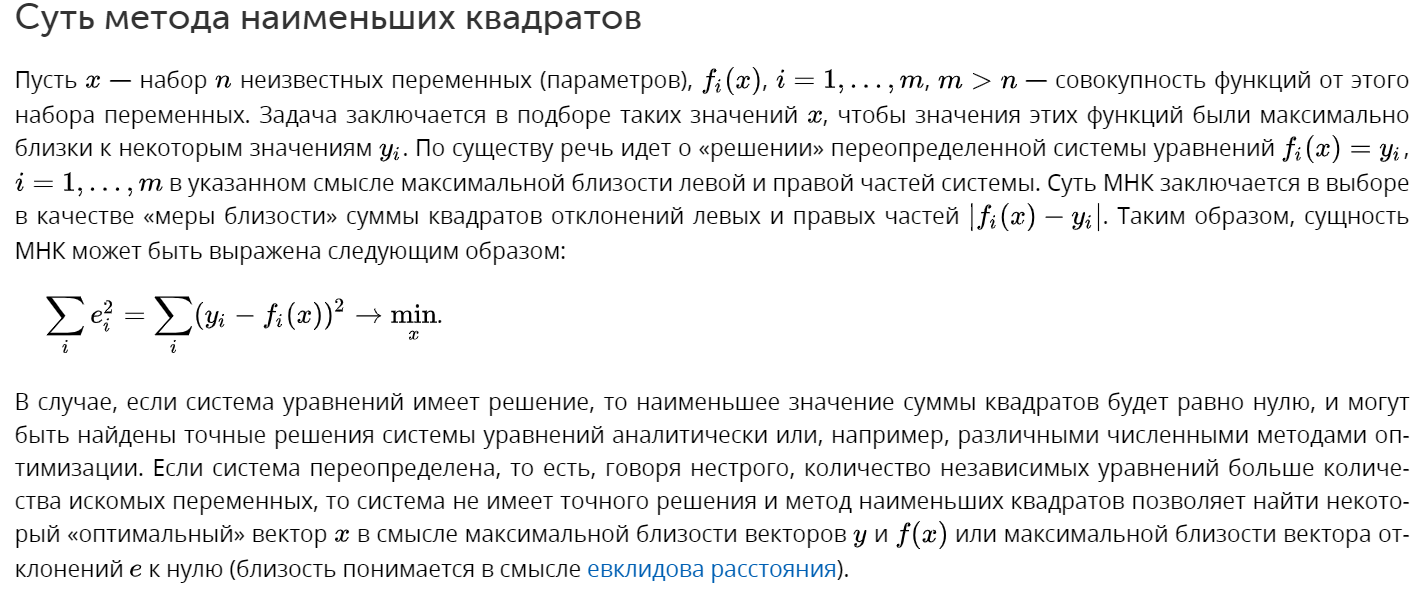
Сделала точечную диаграмму полученных кластеров.



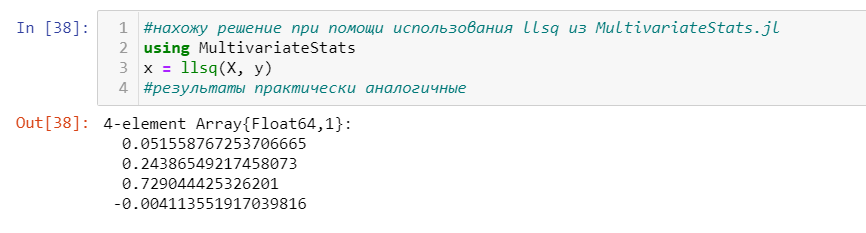
7.4.2. Регрессия (метод наименьших квадратов в случае линейной

регрессии)  
Часть 1:

Пусть регрессионная зависимость является линейной. Матрица наблюдений факторов 𝑋 имеет размерность 𝑁 × 3 randn (N, 3), массив результатов 𝑁 × 1, регрессионная зависимость является линейной. Найдите МНК-оценку для линейной модели.

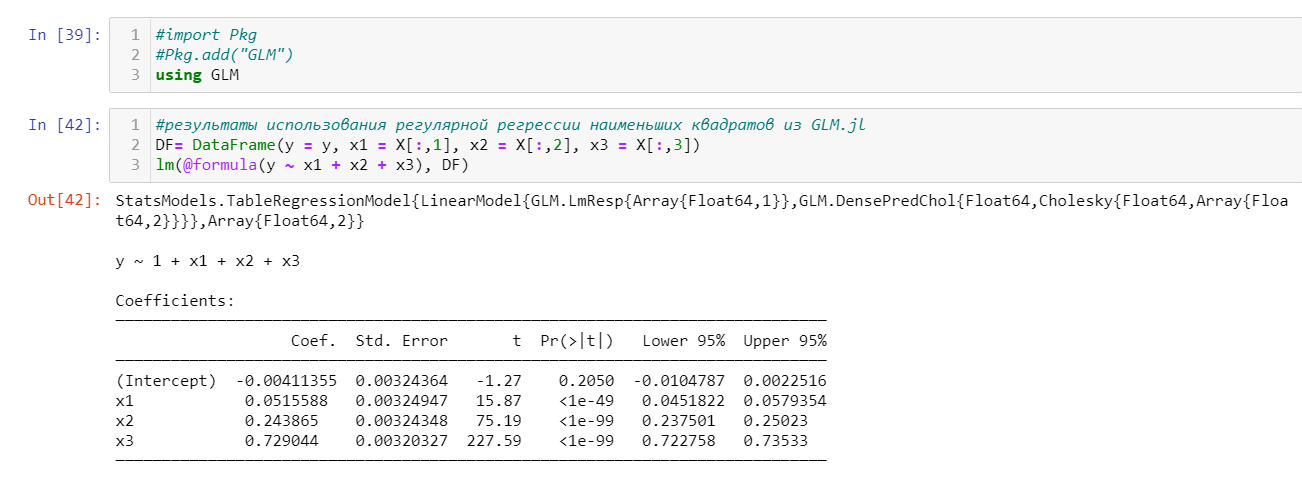


Сравнила свои результаты с результатами использования llsq из MultivariateStats.jl



Погрешность минимальна.

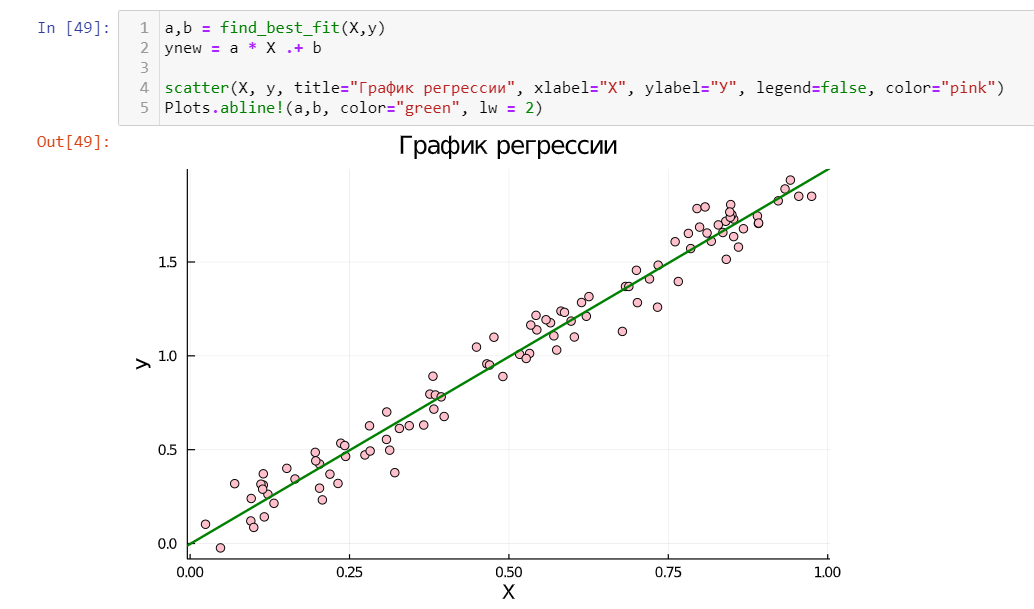
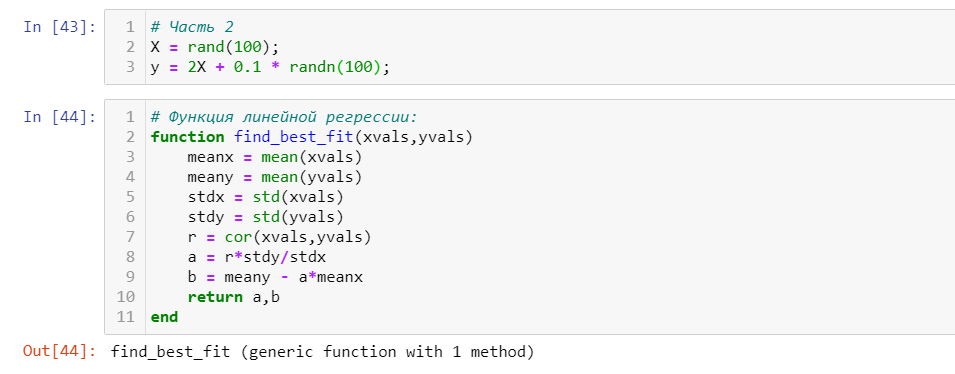
Сравнила свои результаты с результатами использования регулярной регрессии наименьших квадратов из GLM.jl.



Результаты схожи с теми, что получались выше.

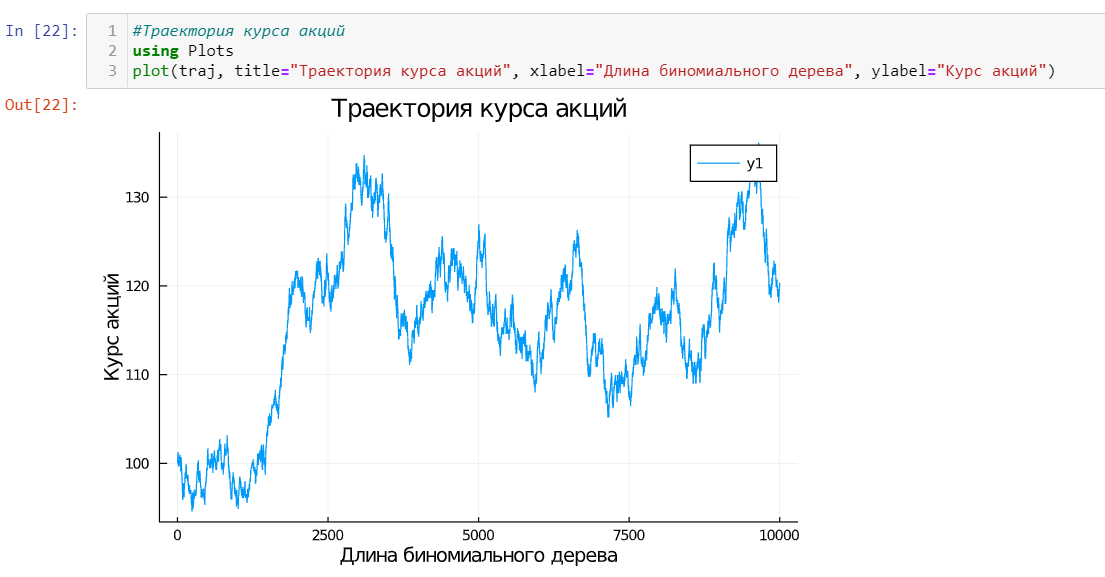
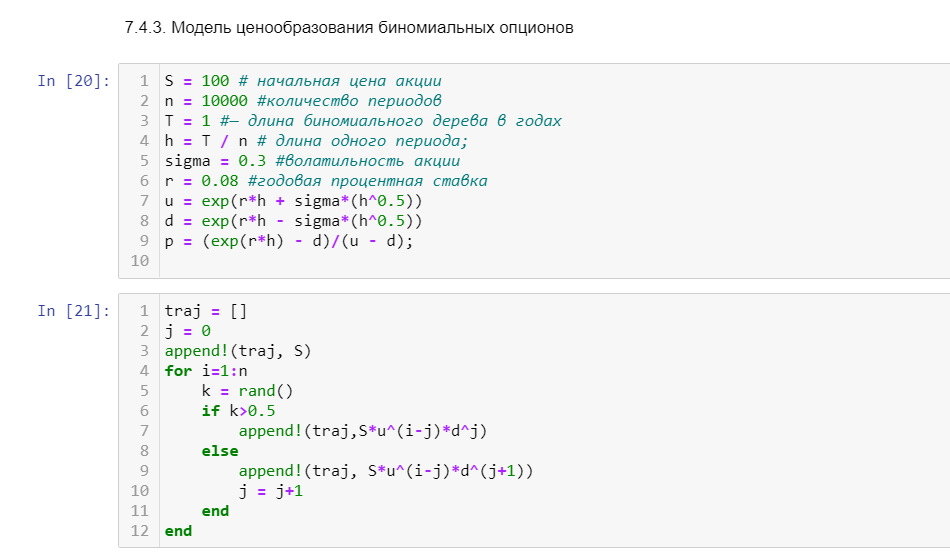
Часть 2

Нашла линию регрессии, используя данные (𝑋, 𝑦). Построила график (𝑋, 𝑦), используя точечный график. Добавила линию регрессии, используя abline!. Добавила заголовок «График регрессии» и подписала оси 𝑥 и 𝑦.



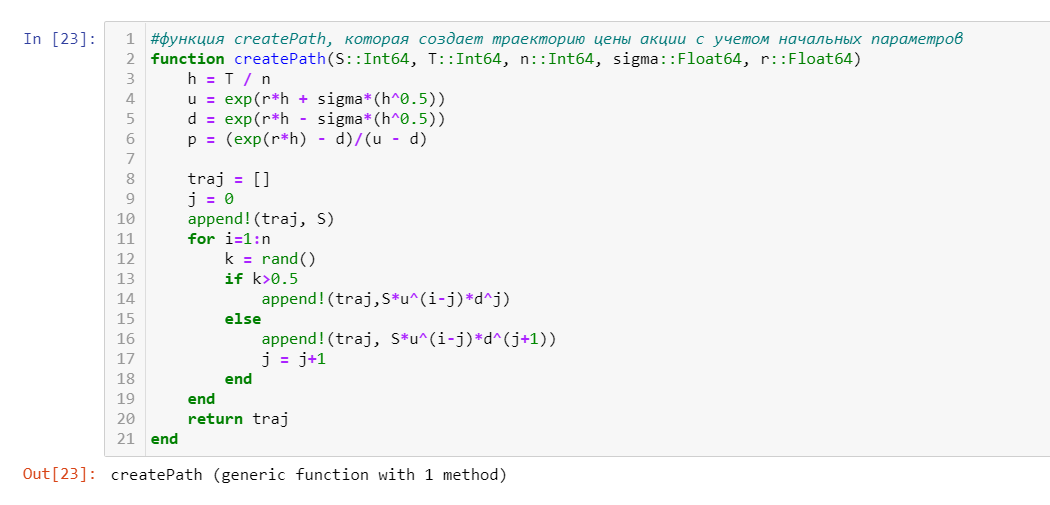
7.4.3. Модель ценообразования биномиальных опционов

Пусть 𝑆 = 100, 𝑇 = 1, 𝑛 = 10000, 𝜎 = 0.3 и 𝑟 = 0.08. Построила траекторию курса акций. Функция rand () генерирует случайное число от 0 до 1.



Создала функцию createPath (S :: Float64, r :: Float64, sigma :: Float64, T :: Float64, n :: Int64), которая создает траекторию цены акции с учетом начальных параметров. Использовала createPath, чтобы создать 10 разных траекторий и построить их все на одном графике.

Функция была создана на основе предыдущего пункта.



Распараллила генерацию траектории. Использовала Threads.@threads, pmap и @parallel



**Вывод:**

Использовала и изучила пакеты Julia для обработки данных.