Министерство науки и высшего образования РФ

Севастопольский государственный университет

Кафедра информационных систем

ОТЧЕТ

по лабораторной работе №4

Кластерный анализ. Основные этапы и задачи кластерного анализа данных

по дисциплине «Интеллектуальный анализ данных»

Выполнил:

Студент группы ИС/б 17-2-о

Черняев Н.Г.

Проверил:

Сырых О.А.

г. Севастополь 2020

1.Цель работы

Закрепить теоретические знания и приобрести практические навыки в проведении кластерного анализа по экспериментальным данным. Исследовать возможности языка R для проведения кластерного анализа.

2. Ход работы

1. Проведем кластерный анализ методом k-средних в R.

Для этого разобьем данные на 2 кластера.

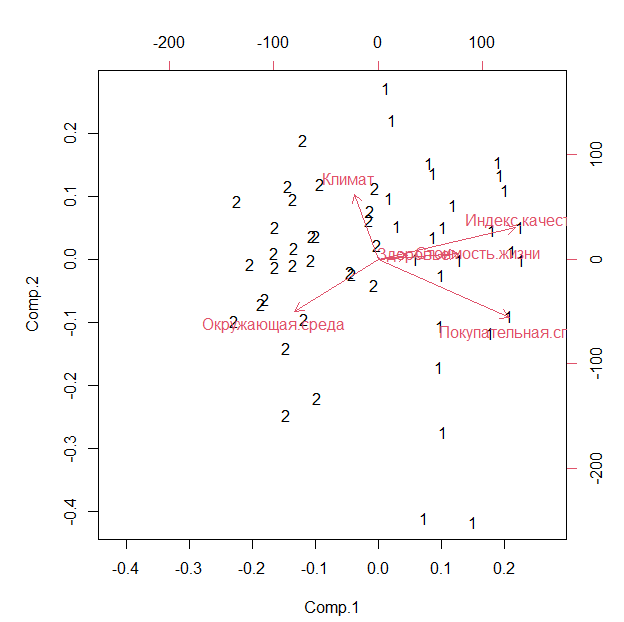


Рисунок 2.1 – Разбиение данных на 2 кластера

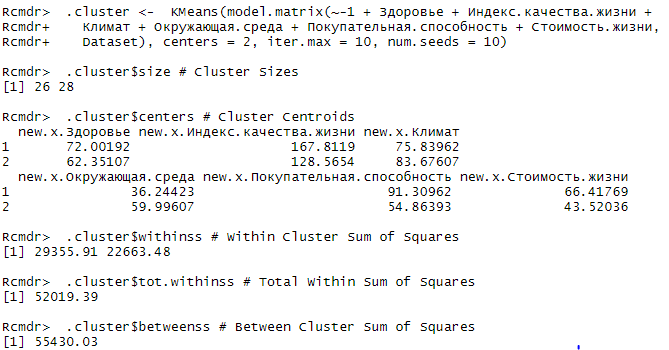
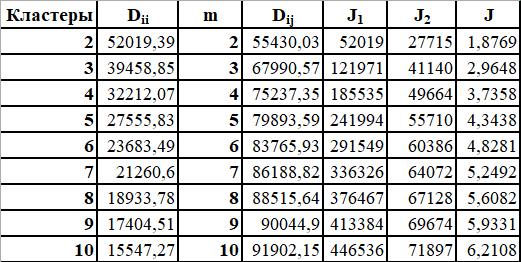


Рисунок 2.2 – Результат выполненного анализа

Для выбора лучшей группировки в смысле критерия минимума отношений средних внутри кластерных и меж кластерных расстояний было проведено деление на 2 – 10 кластеров и заполнена таблица

Таблица 1 – Расчет численного показателя меры качества классификации



Графически значения данного показателя качества классификации представлено на рис 2.3. Для этого построена диаграмма, на которой по оси Х – количество кластеров, по оси Y – значения показателя J.

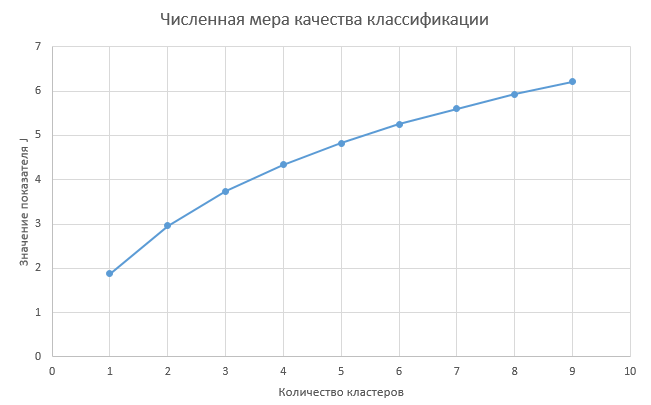


Рисунок 2.3 – Диаграмма численной меры качества классификации

1. Проведем иерархический кластерный анализ в среде R.

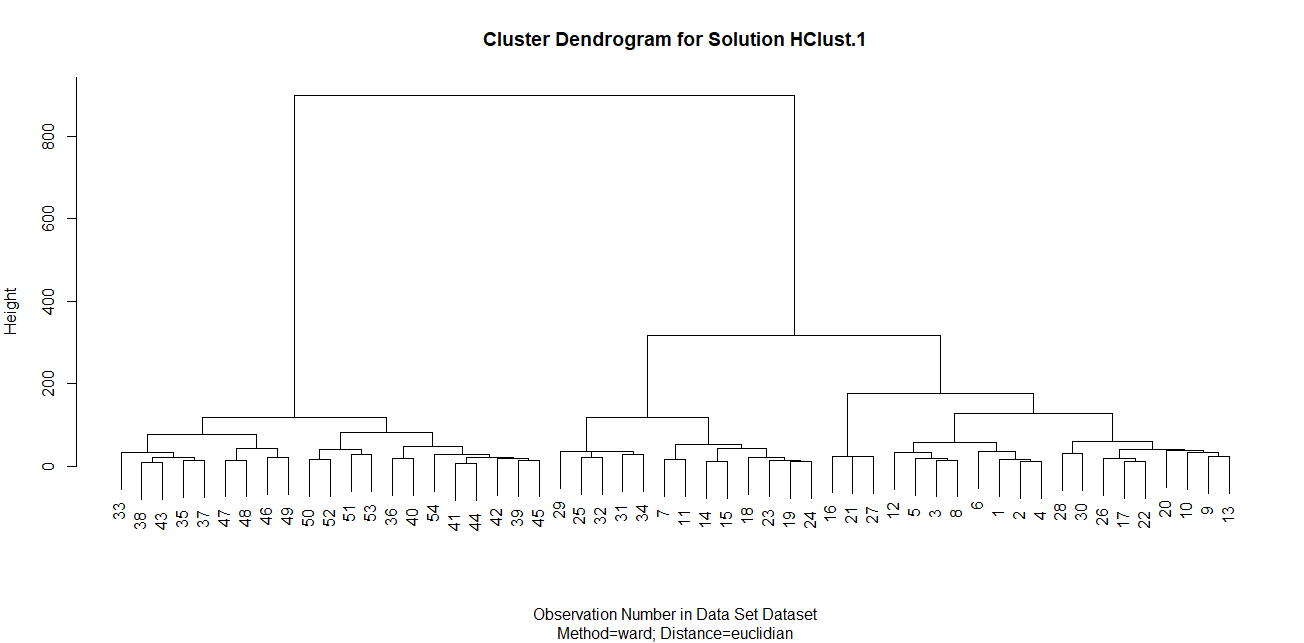


Рисунок 2.4 – Метод Уорда

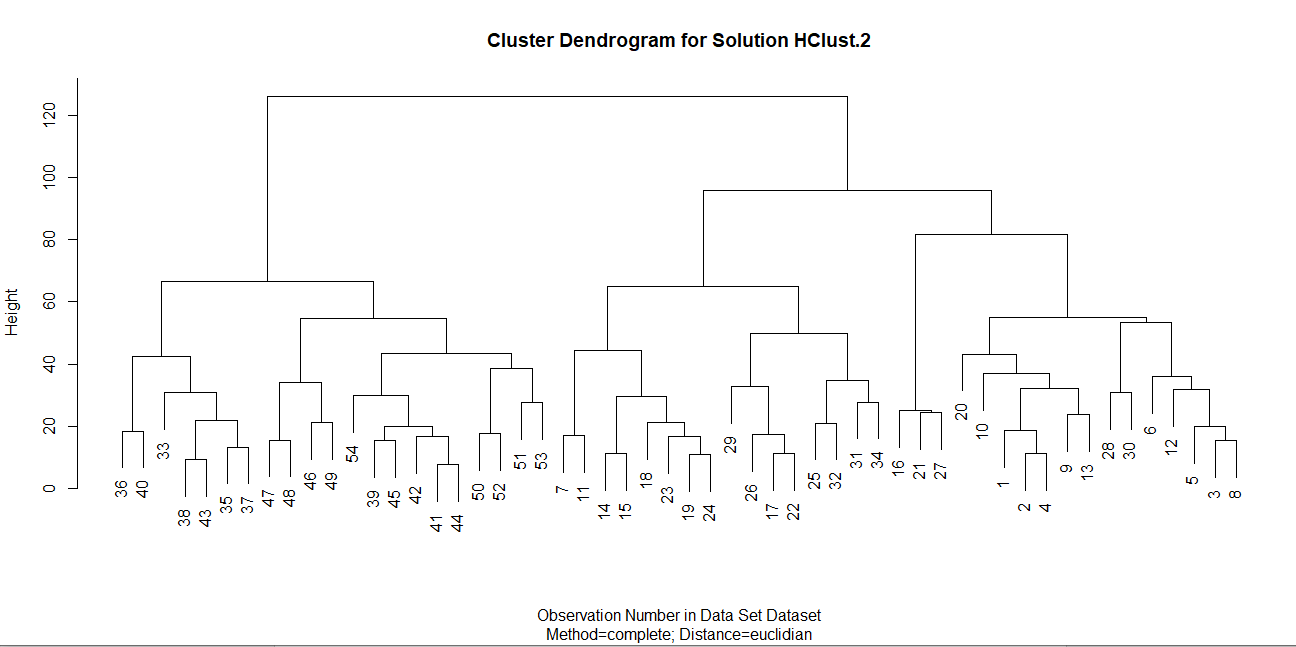


Рисунок 2.5 – Метод полной связи

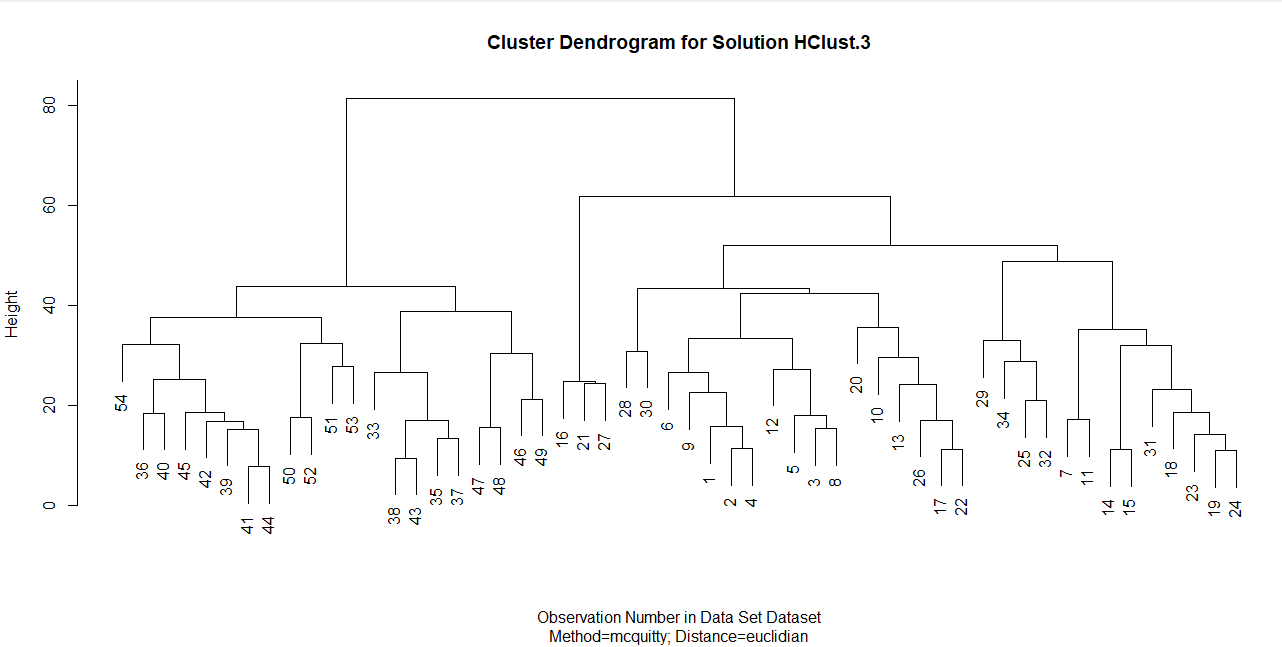


Рисунок 2.6 – Метод Мак-Квитти

3. Вывод

В ходе выполнения лабораторной работы были закреплены теоретические знания и приобретены практические навыки в проведении кластерного анализа по экспериментальным данным, а также были исследованы возможности языка R для проведения кластерного анализа.