

Липецкий государственный технический университет

Факультет Автоматизации и Информатики

Кафедра прикладной математики

ЛАБОРАТОРНАЯ РАБОТА №5

по дисциплине «Статистические методы в прикладных задачах»

Кластерный анализ

Вариант 8

Студенты

Стюфляев А.Р. Коретников Н.И.

Группа АС-19-1

Руководитель

доцент

Рыжкова Д.В.

Липецк, 2021 г

Оглавление

Цель работы.....	3
Задание кафедры	4
Ход выполнения лабораторной работы	6

Цель работы

Научиться проводить кластерный анализ по имеющимся данным.

Задание кафедры

Задание 1. По данным Таблицы 1 и алгоритму кластерного анализа провести классификацию объектов иерархическим методом (древовидная кластеризация).

Таблица 1.

№ п/п.	Страны	Число врачей на 10000 населения	Смертность на 100000 населения	ВВП по паритету покупательной способности, в % к США	Расходы на здравоохранение, в % к США
		X1	X2	X3	X4
1	Россия	44.5	84.98	20.4	3.2
2	Австралия	32.5	30.58	71.4	8.5
3	Австрия	33.9	38.42	78.7	9.2
4	Азербайджан	38.8	60.34	12.1	3.3
5	Армения	34.4	60.22	10.9	3.2
6	Беларусь	43.6	60.79	20.4	5.4
7	Бельгия	41	29.82	79.7	8.3
8	Болгария	36.4	70.57	17.3	5.4
9	Великобритания	17.9	34.51	69.7	7.1
10	Венгрия	32.1	64.73	24.5	6
11	Германия	38.1	36.63	76.2	8.6
12	Греция	41.5	32.84	44.4	5.7
13	Грузия	55	62.64	11.3	3.5
14	Дания	36.7	34.07	79.2	6.7
15	Ирландия	15.8	39.27	57	6.7
16	Испания	40.9	28.46	54.8	7.3
17	Италия	49.4	30.27	72.1	8.5
18	Казахстан	38.1	69.04	13.4	3.3
19	Канада	27.6	25.42	79.9	10.2
20	Киргизия	33.2	53.13	11.2	3.4

Задание 2. Решить Задание 1, предварительно нормировав исходные данные.

Задание 3. Решить Задание 1 при условии, что расстояния между кластерами измеряются по принципу “дальнего соседа”, предварительно нормируя исходные данные.

Задание 4. Решить Задание 1, но в качестве расстояния между объектами принять “расстояние городских кварталов (Манхэттенское расстояние)”, а расстояния между кластерами измерять по методу Варда. Не нормируя предварительно исходные данные.

Задание 5. Решить Задание 1 методом К-средних. Предварительно нормируя исходные данные.

Ход выполнения лабораторной работы

Задание 1. По данным таблицы 1, за исключением 8-ой строки, и алгоритму кластерного анализа проведем классификацию объектов иерархическим методом (древовидная кластеризация). В качестве расстояния между объектами принимаем “обычное евклидово расстояние”, а расстояния между кластерами измеряем по принципу “ближайшего соседа”.

Исходные данные не нормируем.

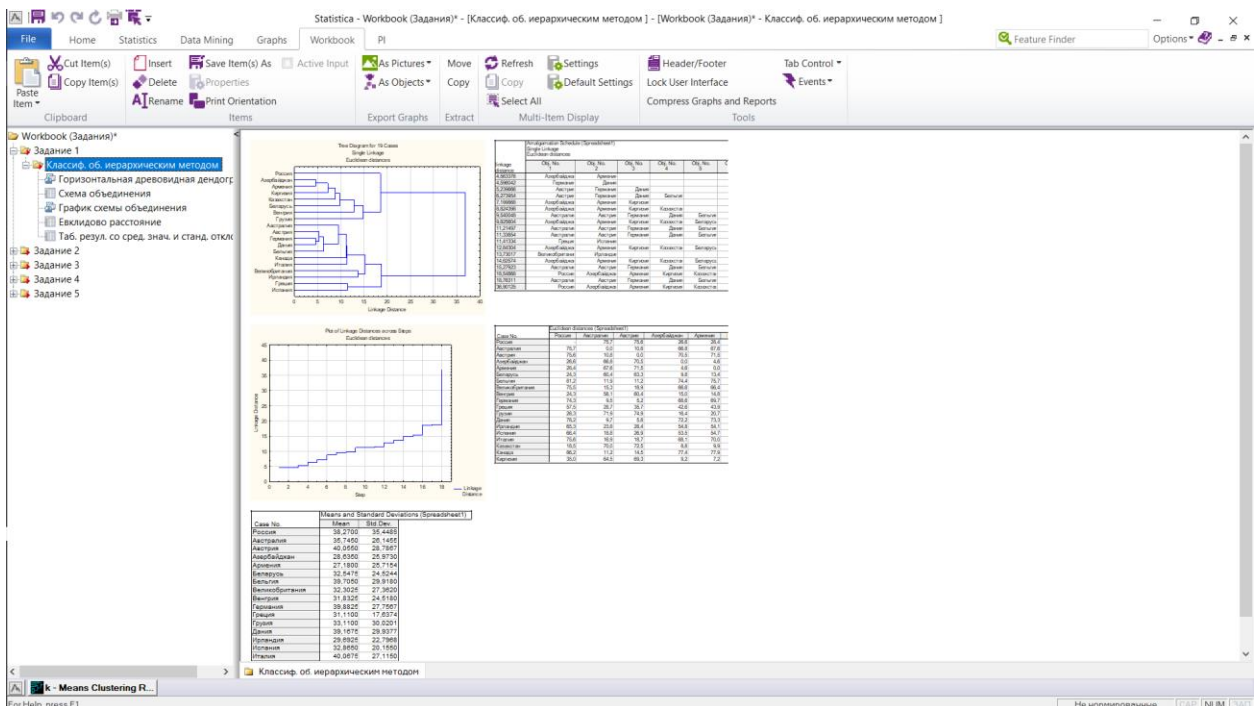


Рисунок 1 - Классификация объектов иерархическим методом (не нормированные данные)

Задание 2. По данным таблицы 1, за исключением 8-ой строки, и алгоритму кластерного анализа проведем классификацию объектов иерархическим методом (древовидная кластеризация). В качестве расстояния между объектами принимаем “обычное евклидово расстояние”, а расстояния между кластерами измеряем по принципу “ближайшего соседа”.

Исходные данные нормируем.

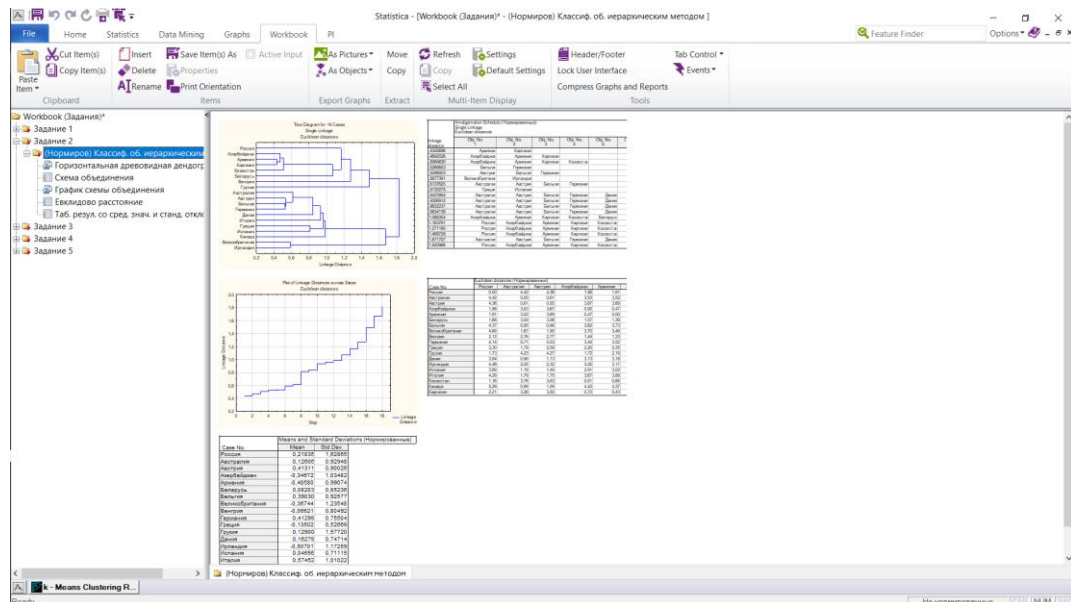


Рисунок 2 - Классификация объектов иерархическим методом (нормированные данные)

Задание 3. По данным таблицы 1, за исключением 8-ой строки, и алгоритму кластерного анализа проведем классификацию объектов иерархическим методом (древовидная кластеризация). В качестве расстояния между объектами принимаем “обычное евклидово расстояние”, а расстояния между кластерами измеряем по принципу “дальнего соседа”.

Исходные данные нормируем.

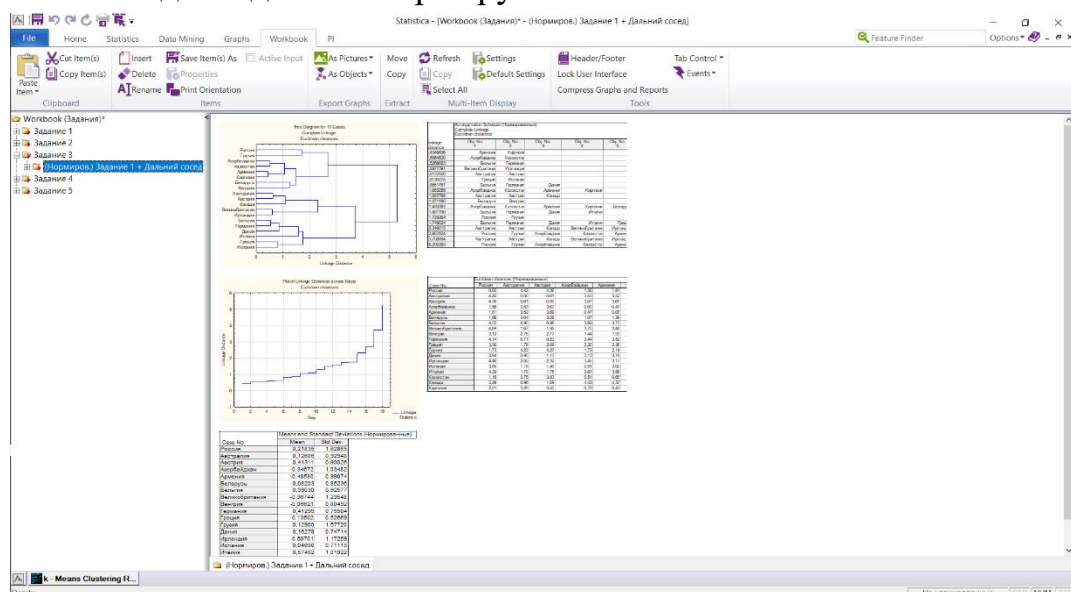


Рисунок 3 - Классификация объектов иерархическим методом с изменениями (нормированные данные)

Задание 4. По данным таблицы 1, за исключением 8-ой строки, и алгоритму кластерного анализа проведем классификацию объектов иерархическим методом (древовидная кластеризация). В качестве расстояния между объектами принимаем “расстояние городских кварталов (Манхэттенское расстояние)”, а расстояния между кластерами измеряем по методу Варда.

Исходные данные не нормируем.

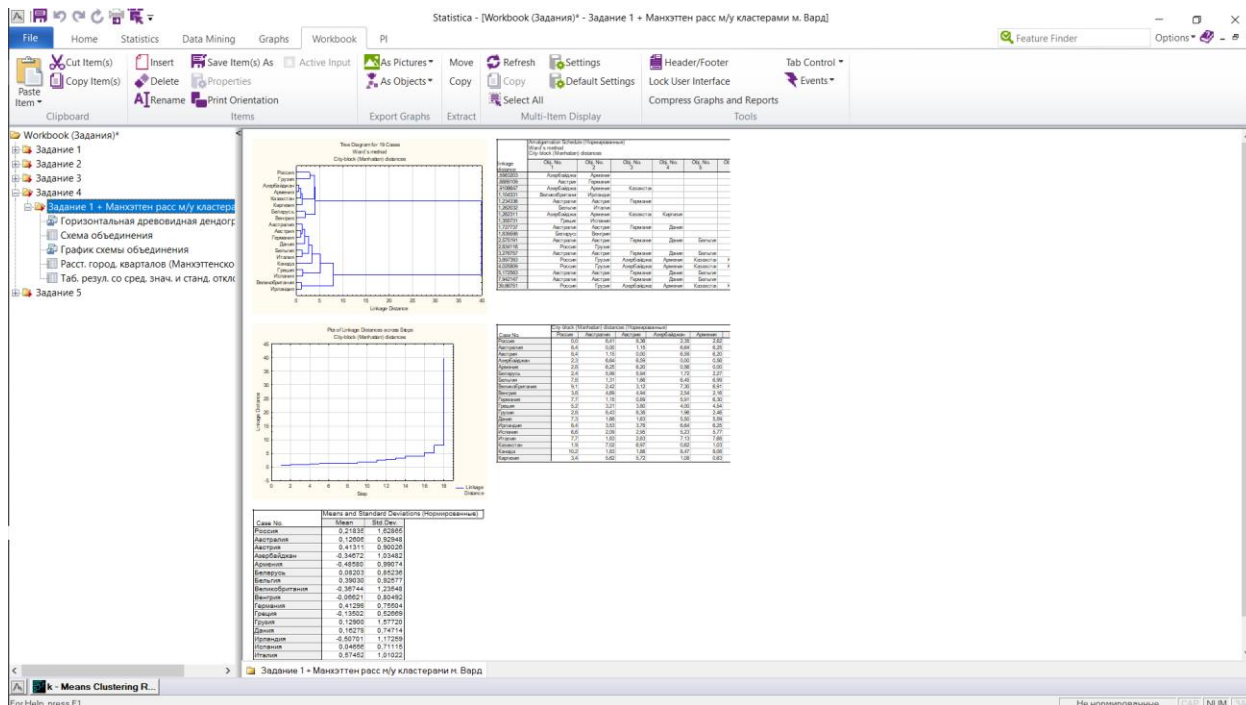


Рисунок 4 - Классификация объектов иерархическим методом с изменениями (не нормированные данные)

Задание 5. По данным таблицы 1, за исключением 8-ой строки, и алгоритму кластерного анализа проведем классификацию объектов методом К-средних.

Исходные данные нормируем.

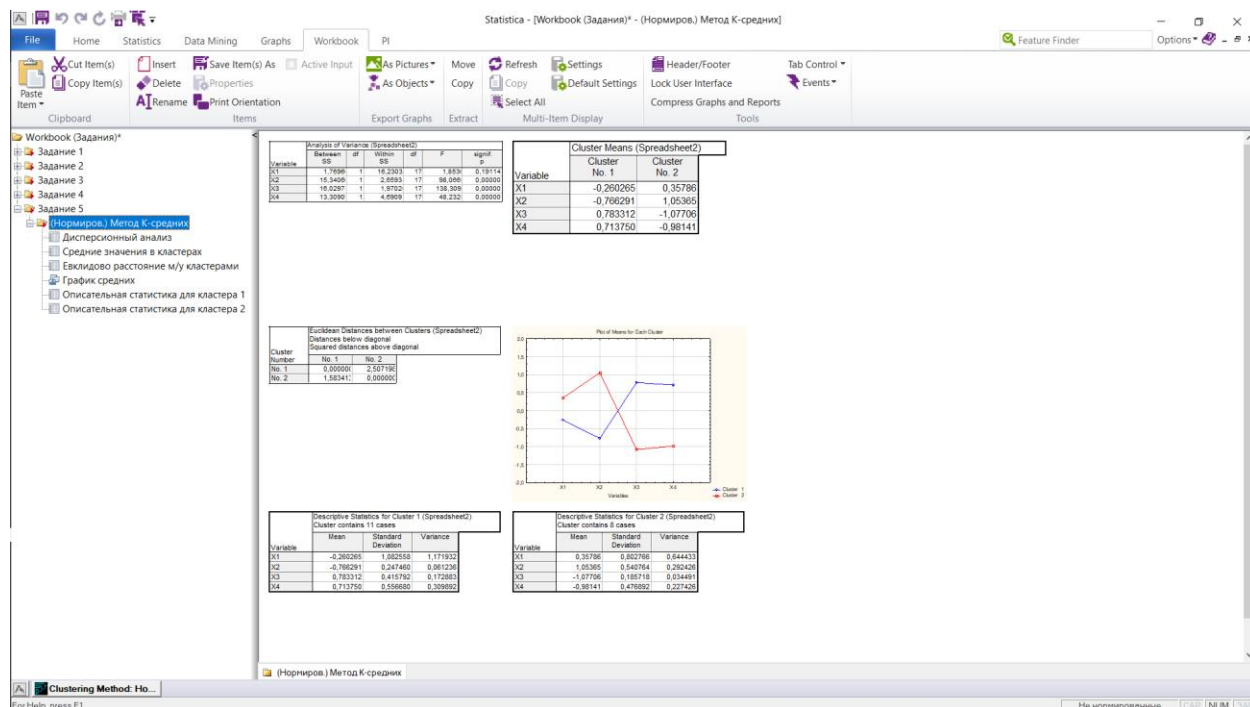


Рисунок 5 - Классификация объектов методом К-средних (нормированные данные)