

Министерство науки и высшего образования Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего образования
«Пермский национальный исследовательский политехнический
университет»
Электротехнический факультет
Кафедра: Информационные технологии и автоматизированные системы

Дисциплина: «Научно-исследовательский семинар»
Лабораторная работа №6
на тему: «Алгоритмы кластеризации в АП Loginom»

Выполнил: студент группы АСУ8-23-1м
Шеретов Марк Алексеевич
Проверил: к.т.н., доцент кафедры ИТАС
Суворов Александр Олегович

Пермь 2024

Постановка задачи

Цель работы: изучить алгоритмы кластеризации в АП Loginom.

Задачи проекта:

Используя материал учебного пособия «Анализ данных в АП Loginom» (автор А.Б. Яковлев), выполнить задания для самостоятельной работы п. 9.3 (стр. 152–153).

Выполнение работы

Создадим пакет для выполнения работы в LogiDom. На рисунке 1 представлен «пустой» редактор после создания пакета.

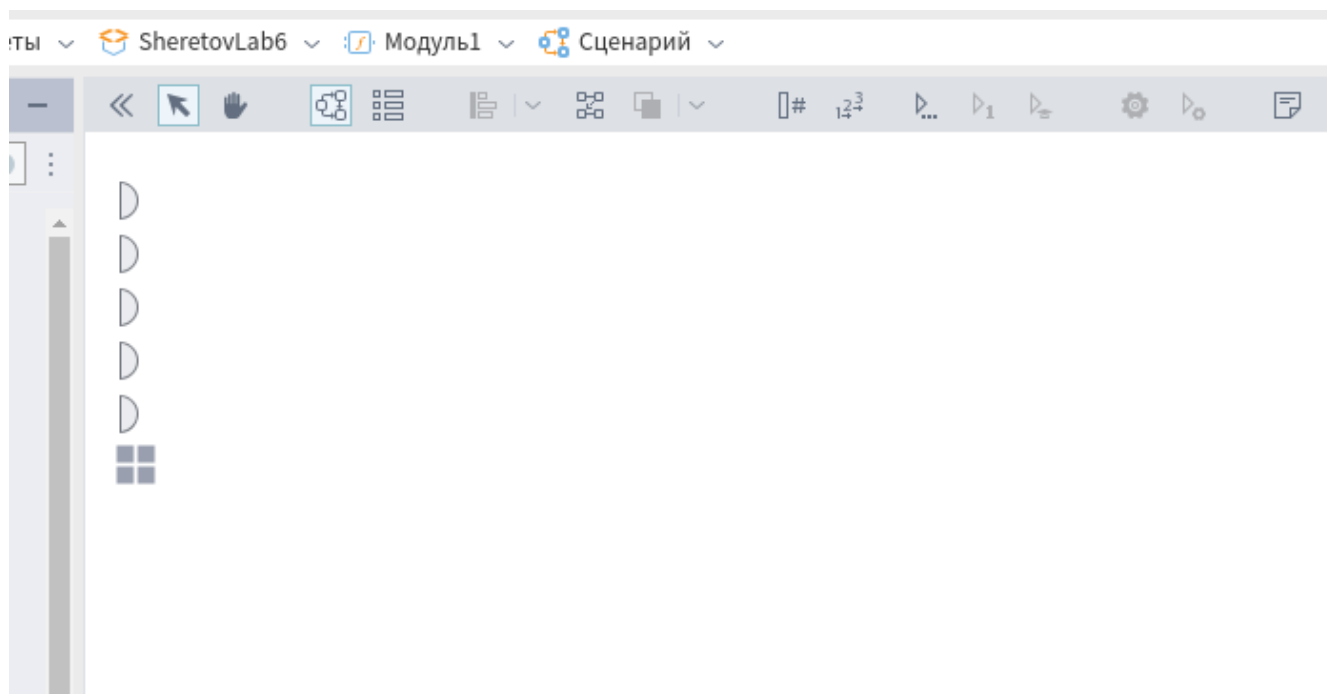


Рисунок 1 — пустой редактор

Задание 1.

Для импорта исходных данных добавим на схему компонент «Excel-файл» и настроим его.

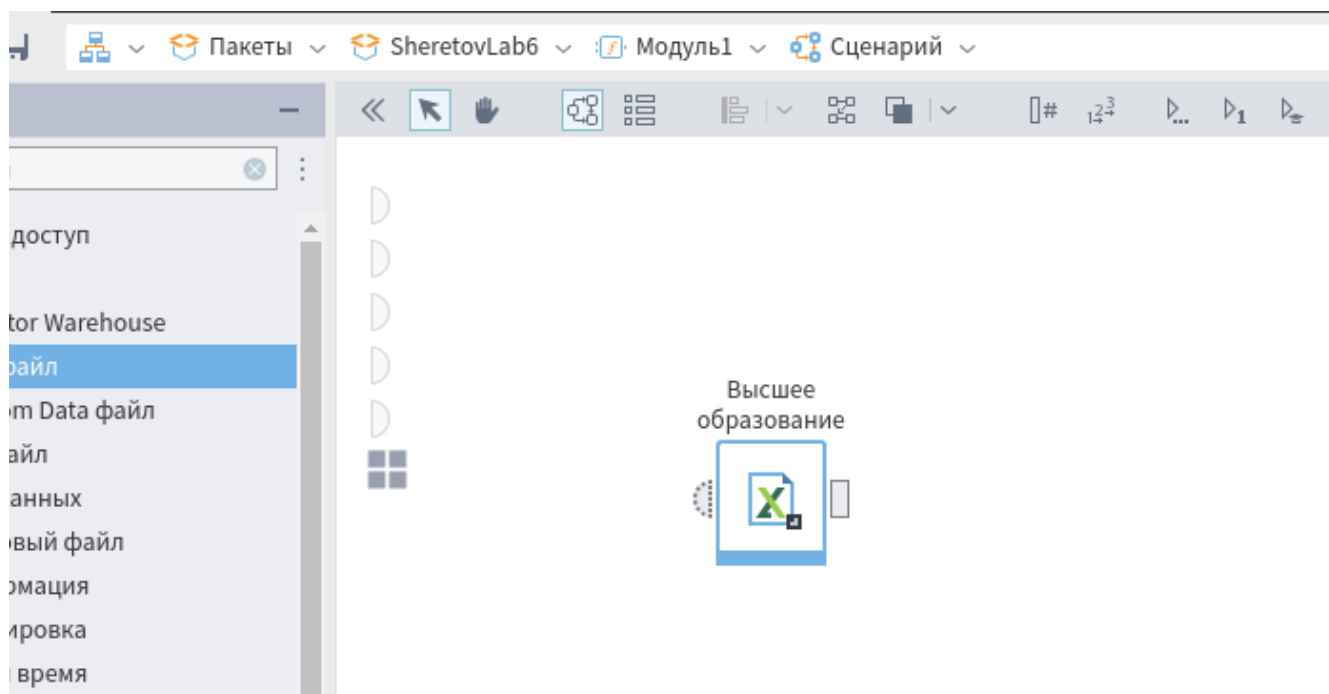


Рисунок 2 — Компонент «Excel-файл» на схеме сценария

Настроим компонент для импорта файла «Высшее образование» как ранее для остальных excel-файлов. Результат импорта представлен на рисунке 3

#	ab Регион	9.0 1. Число образо...	9.0 2. Численнос...	9.0 3. Прием на ...	9.0 4. Выпуск ба...	9.0 5. Численность...
1	Белгородская область	9.677419355	342.5806452	81.29032258	95.48387097	50.6779661
2	Брянская область	15.49755302	281.4029364	59.54323002	83.19738989	41.68115942
3	Владимирская область	12.16893343	239.7995705	62.27630637	55.83392985	47.28358209
4	Воронежская область	13.28761252	393.4847835	100.3000429	111.4444921	59.09586057
5	Ивановская область	10.67961165	303.8834951	72.81553398	87.37864078	65.62300319
6	Калужская область	20.79207921	217.8217822	51.48514851	61.38613861	57.90909091
7	Костромская область	9.216589862	208.9093702	49.15514593	72.19662058	67.5
8	Курская область	14.28571429	487.5	123.2142857	127.6785714	44.24908425
9	Липецкая область	14.70588235	215.3979239	51.90311419	65.74394464	57.22891566
10	Московская область	15.98579041	153.0263697	34.15767181	41.12583686	43.90178571
11	Орловская область	14.47368421	419.7368421	117.1052632	113.1578947	59.90595611
12	Рязанская область	15.04424779	315.0442478	75.22123894	83.18584071	55.05617978
13	Смоленская область	28.15432742	273.2012513	63.60792492	78.20646507	63.09160305
14	Тамбовская область	11.42857143	278.0952381	67.61904762	87.61904762	53.66438356
15	Тверская область	17.62452107	223.7547893	54.40613027	59.77011494	63.49315068
16	Тульская область	13.28021248	232.4037185	57.7689243	67.06507304	51.8
17	Ярославская область	18.08176101	292.4528302	71.5408805	74.68553459	57.06989247
18	г. Москва	16.95052717	616.1394972	171.9383617	182.3195458	67.21205739
19	Республика Карелия	12.6984127	226.984127	50.79365079	53.96825397	71.88811189
20	Республика Коми	12.83547258	239.2065344	56.00933489	60.67677946	42.43902439
21	Архангельская область	8.517887564	193.3560477	51.10732538	51.95911414	76.47577093
22	Вологодская область	10.94276094	204.5454545	46.2962963	63.97306397	49.83539095
23	Калининградская область	16.39344262	280.7377049	72.74590164	69.67213115	53.75912409
24	Ленинградская область	8.431703204	55.64924115	11.24227094	12.36649803	35.45454545
25	Мурманская область	18.37270341	202.0997375	41.99475066	57.74278215	46.23376623
26	Новгородская область	14.61038961	202.9220779	40.58441558	47.07792208	56.72
27	Псковская область	17.02786378	222.9102167	51.08359133	58.82352941	60.625
28	г. Санкт-Петербург	15.88212782	580.3673938	164.5618064	157.2904707	80.00659413
29	Республика Адыгея	15.5210643	279.3791574	66.51884701	82.03991131	57.93650794
30	Республика Калмыкия	14.33691756	304.6594982	75.2688172	71.68458781	49.17647059
31	Краснодарский край	12.51360174	236.6702938	60.39173014	72.72397534	48.97318008
32	Астраханская область	17.66437684	370.9519136	96.17271835	81.45240432	51.4021164
33	Волгоградская область	12.96150825	281.2254517	71.48468185	75.01963865	63.0027933
34	Ростовская область	13.45609065	352.2190746	88.76298395	106.2322946	60.2613941
35	Республика Дагестан	13.26699834	235.4892206	44.11276949	64.67661692	56.43661972
36	Республика Ингушетия	8.456659619	188.1606765	35.94080338	61.31078224	54.71910112
82	Кабардино-Балкарская Республика	6.960556845	198.3758701	45.24361949	54.52436195	74.03508772

Рисунок 3 — результат импорта файла «Высшее образование».

Проведем кластеризацию регионов, используя алгоритм k-means. Для этого переместим компонент *Кластеризация* в рабочую область сценария. Последовательность обработки данных задаётся соединением выходного порта узла импорта с входным портом кластеризации, как показано на рисунке 4.

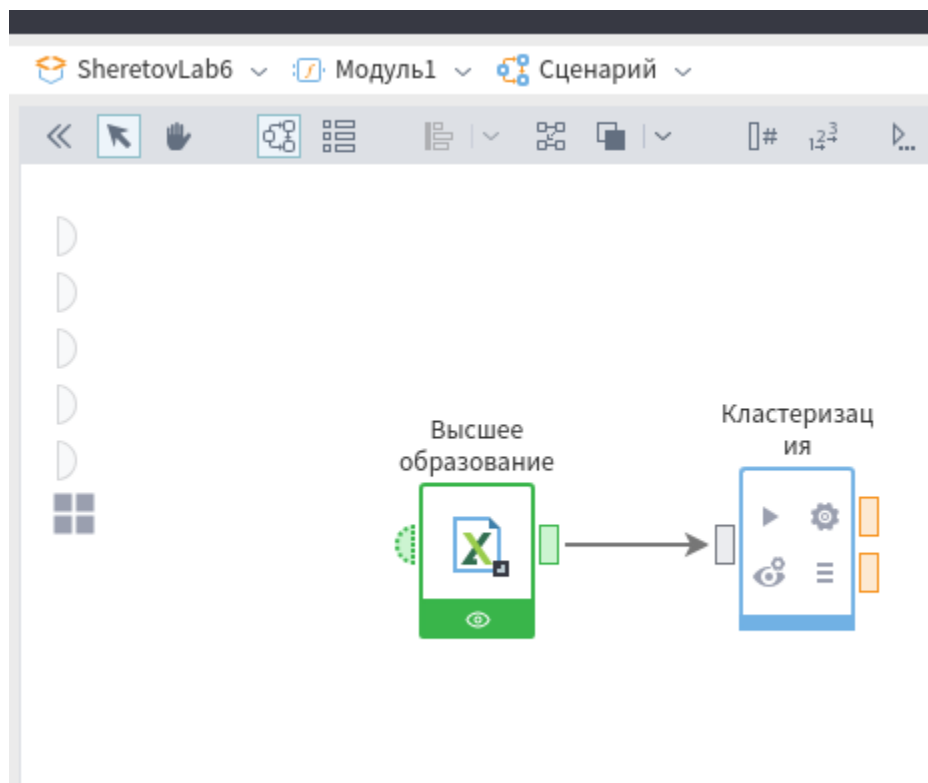


Рисунок 4 — размещение компонента «Кластеризация» на схеме сценария

Пройдём шаги настройки добавленного компонента, для этого на шаге *Настройка входных столбцов* настроим назначение исходных столбцов данных. Для столбца «Регион» выберем назначение «Не задано», а для всех остальных столбцов выберем назначение «Используемое», как показано на рисунке 5.

Настройка входных столбцов

Метка	Имя	Вид данных	Назначение
ab Регион	Region	Дискретный	Не задано
9.0 1. Число образовательных организа...	_1__CHislo_obrazovatelnykh_organizats...	Непрерывн...	Используемое
9.0 2. Численность студентов, обучающ...	_2__CHislennost_studentov_obuchayu...	Непрерывн...	Используемое
9.0 3. Прием на обучение по программа...	_3__Priem_na_obuchenie_po_program...	Непрерывн...	Используемое
9.0 4. Выпуск бакалавров, специалисто...	_4__Vypusk_bakalavrov_spetsialistov_...	Непрерывн...	Используемое
9.0 5. Численность профессорско-преп...	_5__CHislennost_professorsko_prepoda...	Непрерывн...	Используемое

Рисунок 5 — настройка входных столбцов

На шаге *Кластеризация* снимем флажок с параметра *Автоматическая настройка* и зададим число кластеров равным 5, исходя из предположения, что регионы могут быть отнесены к группам с лучшей, выше средней, средней, ниже средней и худшей обеспеченностью учреждениями. Настройки продемонстрированы на рисунке 6.

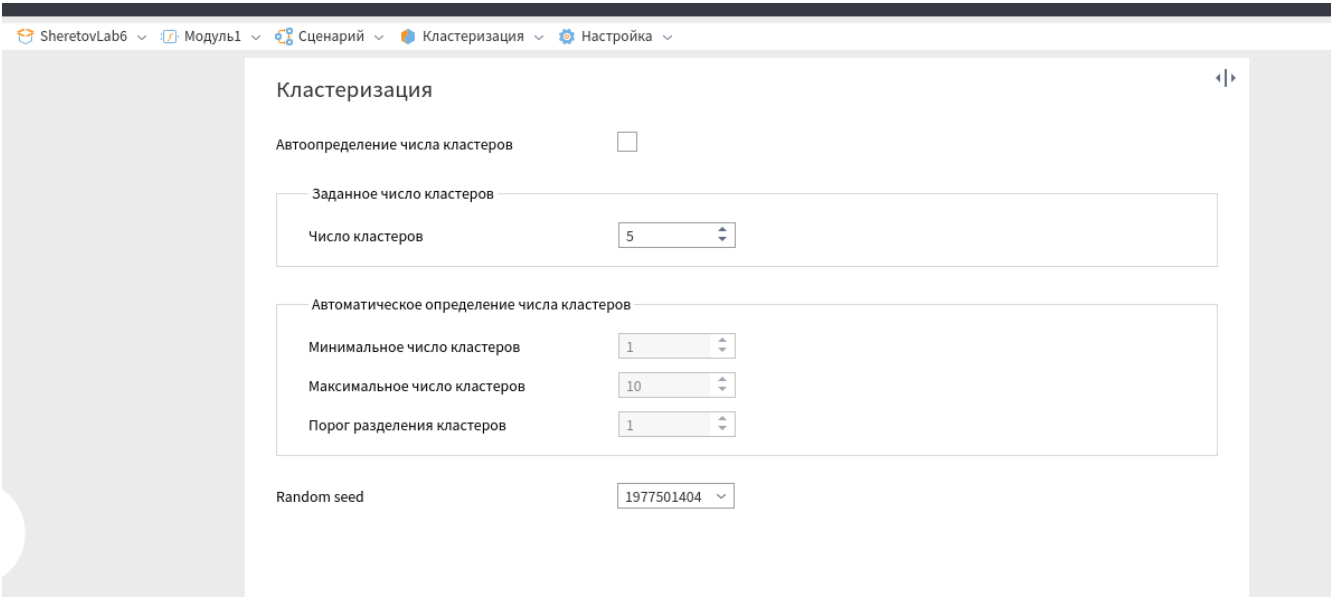


Рисунок 6 — настройки кластеризации

Переобучим узел *Кластеризация* и перейдём к настройкам визуализатора и сначала откроем визуализатор *Центры кластеров*, вид которого показан на рисунке 7.

#	12 Номер кластера	9.0 1. Число образова...	9.0 2. Числен...	9.0 3. Прием н...	9.0 4. Выпус...	9.0 5. Численн...
1	0	14.86992665	509.2171662	141.7746725	132.1199424	60.16986409
2	1	8.186643283	196.7731289	51.0722366	50.34878163	71.50103149
3	2	31.78840393	226.4412334	48.27747157	59.94637271	39.84534917
4	3	13.90505745	323.8881989	80.99267397	85.85171207	57.33909682
5	4	12.42069732	222.1371629	51.48903598	59.21181721	49.7219542

Рисунок 7 — таблица «Центры кластеров»

Видно, что регионы, входящие в кластер 0, относятся к группе лучших по обеспеченности образовательными учреждениями, в кластер 1 — к группе выше

средних, в кластер 2 — к группе средних, в кластер 3 — в группе ниже средних, в кластер 4 — в группе худших. Об этом можно судить по средним значениям всех факторов.

В визуализаторе *Профили кластеров* можно посмотреть общую структуру сформированных кластеров, вид визуализатора представлен на рисунке 8.

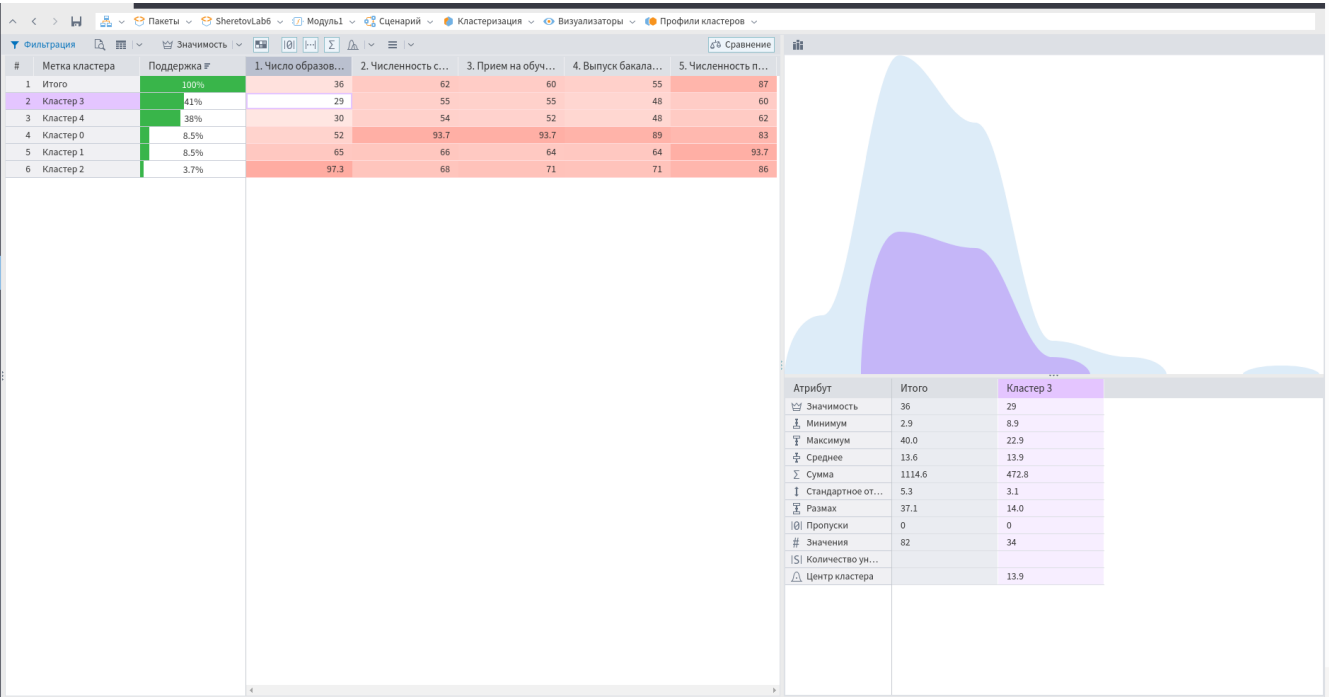


Рисунок 8 — визуализатор «профили кластеров»

При помощи визуализатора *Разбиение на кластеры* можно определить принадлежность регионов к соответствующим кластерам. Так например, как показано на рисунке 9, Пермский край относится к кластеру 4 — регионам с худшей обеспеченностью образовательными учреждениями.

Пакеты SheretovLab6 Модуль1 Сценарий Кластеризация Визуализаторы Таблица																	
#	12	Номер класте...	90	Расстояние до центра кластера	ab	Регион	90	1. Число обра...	90	2. Численност...	90	3. Прием на...	90	4. Выпуск б...	90	5. Численно...	
46		3		0.7170922912		Иркутская область		11.1893908		328.2221301		81.64111065		93.24492333		52.95454545	
47		3		1.280671861		Новосибирская область		11.58580739		399.3482983		106.0825489		94.85879797		57.05349041	
48		3		2.119226223		Республика Саха (Якутия)		22.91666667		280.2083333		72.91666667		77.08333333		68.92193309	
49		3		1.386524403		Приморский край		12.44167963		280.9745982		74.13167444		78.27890098		70.40590406	
50		3		1.493668823		Хабаровский край		14.99250375		417.5412294		93.70314843		110.9445277		52.53141831	
51		3		1.971156436		г. Севастополь		16.82692308		365.3846154		98.55769231		55.28846154		71.31578947	
52		4		1.469333272		Брянская область		15.49755302		281.4029364		59.54323002		83.19738989		41.68115942	
53		4		0.4899512939		Владимирская область		12.16893343		239.7995705		62.27630637		55.83392985		47.28358209	
54		4		1.766431003		Калужская область		20.79207921		217.8217822		51.48514851		61.38613861		57.90909091	
55		4		0.8676671568		Липецкая область		14.70588235		215.3979239		51.90311419		65.74394464		57.22891566	
56		4		1.436794805		Московская область		15.98579041		153.0263697		34.15767181		41.12583686		43.90178571	
57		4		1.633673266		Тверская область		17.62452107		223.7547893		54.40613027		59.77011494		63.49315068	
58		4		0.4561010414		Тульская область		13.28021248		232.4037185		57.7689243		67.06507304		51.8	
59		4		0.7315384402		Республика Коми		12.83547258		239.2065344		56.00933489		60.67677946		42.43902439	
60		4		0.4172029254		Вологодская область		10.94276094		204.5454545		46.2962963		63.97306397		49.83539095	
61		4		3.198955073		Ленинградская область		8.431703204		55.64924115		11.24227094		12.36649803		35.45454545	
62		4		1.237179289		Мурманская область		18.37270341		202.0997375		41.99475066		57.74278215		46.23376623	
63		4		0.9952403943		Новгородская область		14.61038961		202.9220779		40.58441558		47.07792208		56.72	
64		4		1.34882991		Псковская область		17.02786378		222.9102167		51.08359133		58.82352941		60.625	
65		4		0.6119476089		Краснодарский край		12.51360174		236.6702938		60.39173014		72.72397534		48.97318008	
66		4		0.7419972883		Республика Дагестан		13.26699834		235.4892206		44.11276949		64.67661692		56.43661972	
67		4		1.093778643		Республика Ингушетия		8.456659619		188.1606765		35.94080338		61.31078224		54.71910112	
68		4		2.016279921		Чеченская Республика		2.869440459		237.446198		52.36728838		47.34576758		41.72205438	
69		4		1.45429977		Республика Марий Эл		7.288629738		285.7142857		75.80174927		64.13994169		51.2244898	
70		4		1.04937727		Пермский край		12.14882308		252.088079		61.12376614		64.92027335		59.50301205	
71		4		1.196208581		Кировская область		10.79414032		279.1056284		71.70393215		67.84888204		42.90055249	

Рисунок 9 — разбиение регионов на кластеры