

МЕТОДЫ КЛАССИФИКАЦИИ МНОГОМЕРНЫХ НАБЛЮДЕНИЙ И ИХ ПРИМЕНЕНИЕ

Е.А. Козак

*УО «Гродненский государственный университет имени Янки Купалы»,
факультет математики и информатики, специальность «Прикладная математика»,
кафедра фундаментальной и прикладной математики*

Научный руководитель – Т.В. Русилко, кандидат физико-математических наук, доцент, доцент кафедры фундаментальной и прикладной математики, Гродненский государственный университет имени Янки Купалы

Социально-экономические процессы и явления зависят от большого числа факторов их характеризующих, что обуславливает трудности, связанные с выявлением структуры взаимосвязей этих факторов. В подобных ситуациях использование методов классификации многомерных наблюдений является не только оправданным, но и существенно необходимым. Методы многомерной классификации предназначены для разделения рассматриваемой совокупности объектов, субъектов или явлений на однородные группы. Каждый из рассматриваемых объектов характеризуется большим количеством разных признаков. Задача заключается в выделении однородных групп объектов в многомерном пространстве признаков. Для решения столь сложных задач классификации применяют кластерный анализ. Спектр использования кластерного анализа весьма широк. Его применяют в медицине, экономике, государственном управлении и прочих областях.

Целью работы является изучение статистических методов разделения исходной совокупности объектов на кластеры или группы (классы) схожих между собой объектов на основе теории кластеризации; использование процедур кластеризации, реализованных в пакете STATISTICA.

Во введении описана суть задач, решаемых с помощью методов кластерного анализа.

В основной части рассмотрена задача классификации ВУЗов Беларуси. Кластеризация данных осуществлена иерархическим агломеративным методом и методом k -средних. Проведён анализ результатов классификации, сформулированы выводы. Результаты исследования представлены в наглядном графическом виде.

В заключении работы обобщаются полученные выводы.

Ключевые слова: классификация, кластерный анализ, иерархический агломеративный метод, метод k -средних, дисперсионный анализ

Введение. Кластерный анализ предназначен для разбиения множества объектов на однородные группы (классы) на основании некоторого математического критерия качества классификации [1; 2].

Основная цель кластерного анализа – выделить в исходных многомерных данных такие однородные подмножества, чтобы объекты внутри групп были похожи друг на друга, а объекты из разных групп – непохожи. Под «похожестью» понимается близость объектов в многомерном пространстве признаков, и тогда задача сводится к выделению в этом пространстве естественных скоплений объектов, которые и считаются однородными группами.

Решением задачи кластерного анализа является разбиение, удовлетворяющее некоторому критерию оптимальности. Этот критерий может представлять собой некоторый функционал, выражающий уровни желательности различных разбиений и группировок. Этот функционал часто называют целевой функцией. Задача считается решенной, если объекты i и j попадают в один кластер всякий раз, когда расстояние (отдалённость) между соответствующими точками x_i и x_j достаточно малая и, наоборот, попадают в разные кластеры, если расстояние между точками x_i и x_j достаточно большое.

Наиболее часто используются следующие функции расстояния: евклидово расстояние, квадрат евклидова расстояния, расстояние Чебышева, расстояние Хемминга (манхэттенское расстояние), расстояние Минковского.

Традиционно различают классификации иерархические и неиерархические. Соответственно, можно разделить алгоритмы получения этих классификаций. Принцип работы иерархических алгоритмов состоит в последовательном объединении в кластер сначала самых близких, а затем и все более отдалённых друг от друга элементов. Большинство из этих алгоритмов исходит из матрицы сходства (расстояний) и каждый отдельный элемент рассматривается вначале как отдельный кластер.

Помимо иерархических кластер-процедур, существуют неиерархические кластер-процедуры, которые иногда называют структурными. Здесь реализуется идея образования кластеров по принципу

выделения сгущений – мест наибольшей концентрации точек в рассматриваемом пространстве. Примером структурной кластер-процедуры является метод k -средних.

Алгоритмы кластерного анализа отличаются большим разнообразием. В прикладном пакете STATISTICA реализовано шесть видов иерархических процедур [3].

Основная часть. Цель исследования данной работы – применить методику проведения кластерного анализа для классификации ВУЗов Беларуси по семи факторам: количество факультетов, кафедр, студентов, преподавателей, зарубежных партнёров; позиции в мировом рейтинге и рейтинге среди белорусских ВУЗов. Данные приведены в таблице 1. Источник данных – официальные сайты ВУЗов.

Таблица 1 – Исходные данные

ВУЗы	Показатели						
	Кол-во факультетов	Кол-во кафедр	Кол-во студентов	Кол-во преподавателей	Партнёры	Мировой рейтинг	Рейтинг среди бел. ВУЗов
ГрГУ	15	62	12914	723	154	3382	3
БГТУ	11	52	15800	592	200	4970	8
БГУ	27	191	25500	3800	220	684	1
БГУИР	7	32	16000	900	59	3814	4
БГЭУ	11	49	20000	1100	119	4887	7
БГУТ	9	26	10000	300	6	11464	35
БарГУ	5	18	3273	404	144	12539	41
МГЛУ	6	16	7000	1500	156	11918	37
БГМУ	9	72	7046	888	60	4368	6
ВГМУ	8	63	7000	283	53	7221	17
ГГУ	13	43	8500	586	135	3910	5
БрГТУ	6	29	11900	437	80	4982	9
ГГАУ	7	30	8000	340	57	7911	21
МИТ СО	5	9	7000	172	56	8283	22
БрГУ	10	34	7000	560	50	8486	25
БГПУ	10	32	12000	1000	60	5279	10
БНТУ	16	119	34387	2195	130	2715	2

Предварительно нормируем исходные данные в целях устранения различий в единицах измерения показателей. Стандартизованные данные представлены в таблице 2.

Таблица 2 – Стандартизованные данные

ВУЗы	Показатели						
	Кол-во факультетов	Кол-во кафедр	Кол-во студентов	Кол-во преподавателей	Партнёры	Мировой рейтинг	Рейтинг среди бел. ВУЗов
ГрГУ	0,872	0,234	0,046	-0,228	0,863	-0,854	-0,909
БГТУ	0,130	0,009	0,409	-0,374	1,631	-0,386	-0,522
БГУ	3,097	3,135	1,632	3,202	1,961	-1,649	-1,062
БГУИР	-0,610	-0,44	0,435	-0,031	-0,722	-0,727	-0,832
БГЭУ	0,130	-0,058	0,939	0,191	0,278	-0,411	-0,603
БГУТ	-0,239	-0,575	-0,321	-0,700	-1,607	1,526	1,539
БарГУ	-0,981	-0,755	-1,169	-0,584	0,696	1,842	1,998
МГЛУ	-0,796	-0,800	-0,699	0,637	0,896	1,659	1,692
БГМУ	-0,239	0,459	-0,693	-0,044	-0,706	-0,564	-0,679
ВГМУ	-0,425	0,256	-0,699	-0,719	-0,823	0,276	0,162
ГГУ	0,501	-0,193	-0,510	-0,381	0,546	-0,699	-0,756
БрГТУ	-0,796	-0,508	-0,081	-0,547	-0,372	-0,383	-0,450
ГГАУ	-0,610	-0,485	-0,573	-0,655	-0,756	0,479	0,468
МИТ СО	-0,981	-0,957	-0,699	-0,843	-0,772	0,589	0,544
БрГУ	-0,054	-0,395	-0,699	-0,410	-0,873	0,648	0,774
БГПУ	-0,054	-0,440	-0,069	0,080	-0,706	-0,295	-0,373
БНТУ	1,057	1,516	2,753	1,412	0,462	-1,051	-0,985

В первую очередь используем **иерархический агломеративный** метод. Кластеризацию проведём методом полной связи с использованием евклидова расстояния. Результатом данного метода является

вертикальная древовидная диаграмма, представленная на рисунке 1, на которой по оси абсцисс находятся наблюдения, а по оси ординат – расстояния.

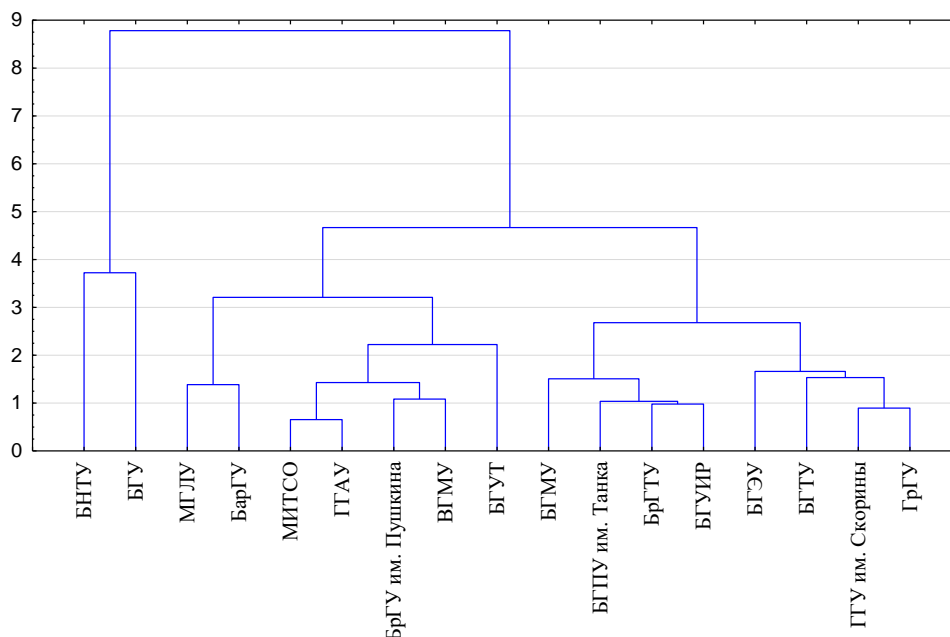


Рисунок 1 - Вертикальная древовидная диаграмма (дендограмма)

Дендограмма очевидно иллюстрирует наличие трехкластеров. Элементы кластеров перечислены в таблице 3.

Таблица 3 - Результат кластеризации агломеративным иерархическим методом

Кластер 1	Кластер 2	Кластер 3
БНТУ БГУ	МГЛУ БарГУ МИТСО ГГАУ БрГУ им. Пушкина БГМУ БГУТ	ГрГУ им. Янки Купалы ГГУ им. Франциска Скорины БГТУ БГЭУ БГУИР БрГТУ БГПУ им. Максима Танка БГМУ

Таким образом, ГрГУ оказался в одном кластере с такими ВУЗами, как БГТУ, БГУИР, БГЭУ, БГМУ, ГГУ им. Франциска Скорины, БрГТУ, БГПУ им. Максима Танка. В первом кластере находятся ведущие ВУЗы Беларуси.

На рисунке 2 изображена матрица расстояний между объектами. Рассмотрим расстояния между ГрГУ и ведущими ВУЗами страны из кластера 1: расстояние между ГрГУ и БГУ – 5,43, между ГрГУ и БНТУ – 3,45. Как видно из таблицы, по исследуемым факторам ГрГУ имеет наименьшее расстояние среди ВУЗов со второго и третьего кластеров к ведущему ВУЗу страны – БГУ. К БНТУ наименьшее расстояние имеет БГЭУ, который входит в третий кластер, следом за ним следует наш университет.

Case No.	Euclidean distances (ВУЗы)														
	ГрГУ	БГТУ	БГУ	БГУИР	БГЭУ	БГУТ	БарГУ	МГЛУ	БГМУ	БГМУ	ГГУ им. Скорины	БрГТУ	ГГАУ	МИТСО	БрГУ им. Пушкина
ГрГУ	0,00	1,31	5,43	2,32	1,50	4,47	4,67	4,27	2,11	2,78	0,90	2,33	3,10	3,56	3,15
БГТУ	1,31	0,00	5,90	2,57	1,56	4,42	4,07	3,66	2,68	2,94	1,53	2,33	3,05	3,39	3,24
БГУ	5,43	5,90	0,00	6,82	5,76	8,58	8,78	7,90	6,54	7,38	6,19	7,27	7,80	8,38	7,59
БГУИР	2,32	2,57	6,82	0,00	1,47	3,56	4,44	4,07	1,51	2,07	1,98	0,98	2,14	2,45	2,50
БГЭУ	1,50	1,56	5,76	1,47	0,00	3,83	4,34	3,77	2,03	2,49	1,66	1,76	2,60	3,04	2,75
БГУТ	4,47	4,42	8,58	3,56	3,83	0,00	2,63	2,93	3,43	2,22	3,96	3,09	1,78	1,85	1,48
БарГУ	4,67	4,07	8,78	4,44	4,34	2,63	0,00	1,38	4,18	3,12	4,13	3,66	2,62	2,48	2,57
МГЛУ	4,27	3,66	7,90	4,07	3,77	2,93	1,38	0,00	3,94	3,21	3,85	3,50	2,73	2,74	2,61
БГМУ	2,11	2,68	6,54	1,51	2,03	3,43	4,18	3,94	0,00	1,40	1,65	1,44	1,96	2,46	2,12
БГМУ	2,78	2,94	7,38	2,07	2,49	2,22	3,12	3,21	1,40	0,00	2,21	1,46	0,86	1,43	1,08
ГГУ им. Скорины	0,90	1,53	6,19	1,98	1,66	3,96	4,13	3,85	1,65	2,21	0,00	1,74	2,45	2,85	2,56
БрГТУ	2,33	2,33	7,27	0,98	1,76	3,09	3,66	3,50	1,44	1,46	1,74	0,00	1,42	1,67	1,94
ГГАУ	3,10	3,05	7,80	2,14	2,60	1,78	2,62	2,73	1,96	0,86	2,45	1,42	0,00	0,66	0,73
МИТСО	3,56	3,39	8,38	2,45	3,04	1,85	2,48	2,74	2,46	1,43	2,85	1,67	0,66	0,00	1,20
БрГУ им. Пушкина	3,15	3,24	7,59	2,50	2,75	1,48	2,57	2,61	2,12	1,08	2,56	1,94	0,73	1,20	0,00
БГПУ им. Танка	2,12	2,48	6,69	0,99	1,50	2,92	3,84	3,47	1,19	1,51	1,63	1,04	1,56	2,00	1,70
БНТУ	3,45	3,72	3,72	3,95	2,95	6,11	6,78	6,00	4,30	5,00	4,16	4,57	5,30	5,82	5,28

Рисунок 2 - Матрица расстояний

Для проведения сравнительного анализа нас интересует сопоставление описательных статистических характеристик кластеров. Определим их для каждого из кластеров, используя исходные ненормированные данные. С помощью данного анализа можно увидеть средние значения переменных в каждом кластере, а также минимальное и максимальное значения по каждой переменной в соответствующем кластере. Результаты представлены на рисунках 3-5.

Variable	Descriptive Statistics (ВУЗы)		
	Include cases: 16:17		
	Mean	Minimum	Maximum
Кол-во факультетов	21,50	16,00	27,00
Кол-во кафедр	155,00	119,00	191,00
Кол-во студентов	29943,50	25500,00	34387,00
Кол-во преподавателей	2997,50	2195,00	3800,00
Партнёры	175,00	130,00	220,00
Позиция в мировом рейтинге	1699,50	684,00	2715,00
Позиция в рейтинге среди бел. ВУЗов	1,50	1,00	2,00

Рисунок 3 - Описательные статистики для 1-го кластера

Variable	Descriptive Statistics (ВУЗы)		
	Include cases: 1:7		
	Mean	Minimum	Maximum
Кол-во факультетов	7,143	5,000	10,00
Кол-во кафедр	28,000	9,000	63,00
Кол-во студентов	7039,000	3273,000	10000,00
Кол-во преподавателей	508,429	172,000	1500,00
Партнёры	74,571	6,000	156,00
Позиция в мировом рейтинге	9688,857	7221,000	12539,00
Позиция в рейтинге среди бел. ВУЗов	28,286	17,000	41,00

Рисунок 4 - Описательные статистики для 2-го кластера

Variable	Descriptive Statistics (ВУЗы)		
	Include cases: 8:15		
	Mean	Minimum	Maximum
Кол-во факультетов	10,25	6,000	15,00
Кол-во кафедр	46,38	29,000	72,00
Кол-во студентов	13020,00	7046,000	20000,00
Кол-во преподавателей	778,25	437,000	1100,00
Партнёры	108,38	59,000	200,00
Позиция в мировом рейтинге	4449,00	3382,000	5279,00
Позиция в рейтинге среди бел. ВУЗов	6,50	3,000	10,00

Рисунок 5 - Описательные статистики для 3-го кластера

Нас интересует третий кластер, включающий ГрГУ. В целом, кластер 3 по всем показателям, кроме позиции в рейтинге среди белорусских ВУЗов, существенно превышает в среднем кластер 2. Очевидно, что кластер 1 содержит сильнейшие ВУЗы страны по исследуемым показателям.

Анализируя исходные данные и максимальные значения по каждой переменной для третьего кластера, можно сделать вывод, что в своем кластере ГрГУ имеет наилучшие показатели по количеству факультетов, позиции в мировом рейтинге и рейтинге среди белорусских ВУЗов. По количеству кафедр в данном кластере лидирует БГМУ, по количеству студентов и преподавателей – БГЭУ, по количеству партнёров – БГТУ.

Кластеризация методом *k*-средних. Данный метод существенно отличается от иерархических агломеративных методов. Метод *k*-средних является итерационной процедурой. Он строит ровно *k* различных кластеров расположенных на возможно больших расстояниях друг от друга.

Зададим число кластеров равное трем. Для начального определения центров кластеров выберем метод сортировки расстояний и выбора наблюдений на постоянных интервалах. В итоге получаем кластеры, представленные на рисунках 6-8. Столбец «Distance» содержит расстояние от объекта до центра кластера.

	Distance
БГУТ	0,555751
БарГУ	0,701242
МГЛУ	0,761977
ВГМУ	0,549253
ГГАУ	0,321537
МИТСО	0,374178
БрГУ им. Пушкина	0,306863

Рисунок 6 - Кластер 1

	Distance
ГрГУ	0,486050
БГТУ	0,606904
БГУИР	0,443737
БГЭУ	0,373318
БГМУ	0,481442
ГГУ им. Скорины	0,353025
БрГТУ	0,418143
БГПУ им. Танка	0,371902

Рисунок 7 - Кластер 2

	Distance
БГУ	0,703587
БНТУ	0,703587

Рисунок 8 - Кластер 3

Метод *k*-средних иллюстрирует, что ГрГУ оказался в одном кластере с такими ВУЗами как БГТУ, БГУИР, БГЭУ, БГМУ, ГГУ им. Франциска Скорины, БрГТУ, БГПУ им. Максима Танка. В третьем кластере оказались два ведущих ВУЗа Беларуси: БГУ, БНТУ.

В таблице, изображённой на рисунке 9, приведены расстояния между кластерами. Можно сделать вывод, что наибольшее расстояние между первым и третьим кластером (евклидово расстояние равно 2,502421), то есть ВУЗы, входящие в эти кластеры, значительно отличаются. Наименьшее расстояние между первым и вторым кластером (евклидово расстояние равно 0,973995), то есть ВУЗы, входящие в эти кластеры, наиболее схожи. Второй кластер, в который входит ГрГУ, ближе по показателям к ведущим ВУЗам страны, нежели чем ВУЗы из первого кластера.

Cluster Number	Euclidean Distances between Clusters (ВУЗы)		
	Distances below diagonal		
	Squared distances above diagonal		
	No. 1	No. 2	No. 3
No. 1	0,000000	0,948667	6,262112
No. 2	0,973995	0,000000	3,290581
No. 3	2,502421	1,813995	0,000000

Рисунок 9 - Таблица расстояний между кластерами

Индикатором того насколько хорошо проведена классификация является метод дисперсионного анализа, который тестирует нулевую гипотезу о равенстве средних значений показателей по кластерам.

На рисунке 10 представлена таблица результатов дисперсионного анализа. Анализируя *F*-статистики, делаем заключение, что проведенная методом *k*-средних классификация эффективна по всем переменным, кроме фактора «партнеры». На уровне значимости 0,05 данные являются статистически однородными по переменной «партнеры» и не могут быть разделены на разные группы. Следовательно ВУЗы Беларуси по переменной «партнеры» существенно не различаются. Тем не менее на уровне значимости 0,1 различие по «партнерам» должно быть признано статистически значимым.

Variable	Analysis of Variance (ВУЗы)					
	Between SS	df	Within SS	df	F	signif. p
Кол-во факультетов	11,02134	2	4,97866	14	15,49600	0,000282
Кол-во кафедр	12,89841	2	3,10159	14	29,11046	0,000010
Кол-во студентов	13,02552	2	2,97448	14	30,65369	0,000008
Кол-во преподавателей	12,40513	2	3,59487	14	24,15553	0,000029
Партнёры	4,52940	2	11,47060	14	2,76409	0,097334
Позиция в мировом рейтинге	13,02740	2	2,97260	14	30,67747	0,000008
Позиция в рейтинге среди бел. ВУЗов	12,74630	2	3,25370	14	27,42233	0,000014

Рисунок 10 - Дисперсионный анализ

На рисунке 11 изображён график средних значений переменных для каждого кластера. На оси абсцисс отмечены переменные, по которым проводилась классификация, на оси ординат – средние значения переменных в разрезе получаемых кластеров. Исходя из графика, можно сделать вывод, что третий кластер имеет наилучшие показатели по всем переменным. ВУЗы, входящие в первый кластер, имеют худшие показатели по всем переменным.

Второй кластер, в который входит ГрГУ, уступает по количеству факультетов, кафедр, студентов, преподавателей третьему кластеру. По позиции в мировом рейтинге и рейтинге среди белорусских ВУЗов второй кластер близок к ведущим ВУЗам страны, в то время как первый кластер значительно уступает по данным признакам (чем меньше значение данного признака, тем лучше). Второй кластер наиболее схож с ВУЗами входящих в первый кластер по количеству преподавателей.

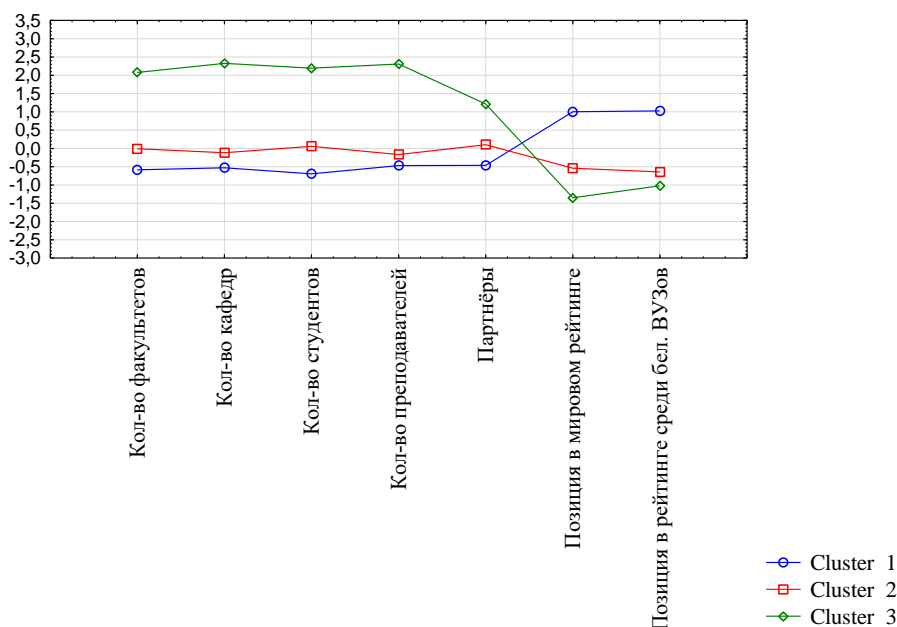


Рисунок 11 - Линейный график средних значений переменных для каждого кластера

На рисунках 12-18 рассмотрим описательные статистики для каждого из показателей: медиана, нижний и верхний квартили.

Breakdown Table of Descriptive Statistics (ВУЗы) N=17 (No missing data in dep. var. list)			
Cluster	Кол-во факультетов Q25	Кол-во факультетов Median	Кол-во факультетов Q75
1	5,00000	7,00000	9,00000
2	8,00000	10,50000	12,00000
3	16,00000	21,50000	27,00000
All Grps	7,00000	9,00000	11,00000

Рисунок 12 – Описательные статистики для признака «количество факультетов»

Breakdown Table of Descriptive Statistics (ВУЗы) N=17 (No missing data in dep. var. list)			
Cluster	Кол-во кафедр Q25	Кол-во кафедр Median	Кол-во кафедр Q75
1	16,0000	26,0000	34,0000
2	32,0000	46,0000	57,0000
3	119,0000	155,0000	191,0000
All Grps	29,0000	34,0000	62,0000

Рисунок 13 – Описательные статистики для признака «количество кафедр»

Breakdown Table of Descriptive Statistics (ВУЗы) N=17 (No missing data in dep. var. list)			
Cluster	Кол-во студентов Q25	Кол-во студентов Median	Кол-во студентов Q75
1	7000,00	7000,00	8000,00
2	10200,00	12457,00	15900,00
3	25500,00	29943,50	34387,00
All Grps	7000,00	10000,00	15800,00

Рисунок 14 – Описательные статистики для признака «количество студентов»

Breakdown Table of Descriptive Statistics (ВУЗы) N=17 (No missing data in dep. var. list)			
Cluster	Кол-во преподавателей Q25	Кол-во преподавателей Median	Кол-во преподавателей Q75
1	283,000	340,000	560,000
2	589,000	805,500	950,000
3	2195,000	2997,500	3800,000
All Grps	404,000	592,000	1000,000

Рисунок 15 – Описательные статистики для признака «количество преподавателей»

Breakdown Table of Descriptive Statistics (ВУЗы) N=17 (No missing data in dep. var. list)			
Cluster	Партнёры Q25	Партнёры Median	Партнёры Q75
1	50,000€	56,000€	144,000€
2	60,000€	99,500€	144,500€
3	130,000€	175,000€	220,000€
All Grps	57,000€	80,000€	144,000€

Рисунок 16 – Описательные статистики для признака «партнёры»

Breakdown Table of Descriptive Statistics (ВУЗы) N=17 (No missing data in dep. var. list)			
Cluster	Позиция в мировом рейтинге Q25	Позиция в мировом рейтинге Median	Позиция в мировом рейтинге Q75
1	7911,00€	8486,00€	11918,00€
2	3862,00€	4627,50€	4976,00€
3	684,00€	1699,50€	2715,00€
All Grps	3910,00€	4982,00€	8283,00€

Рисунок 17 – Описательные статистики для признака «позиция в мировом рейтинге»

Breakdown Table of Descriptive Statistics (ВУЗы) N=17 (No missing data in dep. var. list)			
Cluster	Позиция в рейтинге среди бел. ВУЗов Q25	Позиция в рейтинге среди бел. ВУЗов Median	Позиция в рейтинге среди бел. ВУЗов Q75
1	21,0000€	25,0000€	37,0000€
2	4,5000€	6,5000€	8,5000€
3	1,0000€	1,5000€	2,0000€
All Grps	5,0000€	9,0000€	22,0000€

**Рисунок 18 – Описательные статистики для признака
«позиция в рейтинге среди белорусских ВУЗов»**

Анализируя результаты, делаем вывод о заметном различии между кластерами по всем анализируемым переменным. Наилучшее значение медианы по каждому показателю имеет третий кластер. Второй кластер, в который входит ГрГУ, близок по показателю «позиция в рейтинге среди белорусских ВУЗов» с третьим кластером (разница равна 5). Между первым и вторым кластером разница по данной переменной равна 18,5. По всем остальным показателям второй кластер превышает первый приблизительно в два раза.

Заключение. В данной работе был рассмотрен такой метод классификации многомерных наблюдений как кластерный анализ. В частности, были применены два метода: иерархический агломеративный метод и метод k -средних. Была проведена классификация ВУЗов Беларуси по семи факторам: количество факультетов, кафедр, студентов, преподавателей, зарубежных партнёров; позиции в мировом рейтинге и рейтинге среди белорусских ВУЗов. В результате проведенного исследования оптимальной была признана классификация, выделяющая три однородных кластера. Исходя из таблицы расстояний выяснилось, что кластер, в который входит ГрГУ, имеет наименьшее расстояние до кластера, в котором расположены лидирующие ВУЗы страны (БГУ, БНТУ). В своем кластере ГрГУ является одним из лидеров (имеет наилучшие показатели по количеству факультетов, позициям в мировом рейтинге и рейтинге среди белорусских ВУЗов). Дисперсионный анализ показал, что классификация эффективна по всем переменным, кроме фактора «партнёры».

Список литературы

1. Айвазян, С.А. Классификация многомерных наблюдений / С.А. Айвазян, З. И. Бежаева, О.В. Староверов. – М.: Статистика, 1974. – 200 с.
2. Берестнева, О.Г., Прикладная математическая статистика / О.Г. Берестнева, О.В. Марухина, Г.Е. Шевелёв. – Томск ТПУ, 2012. – 188 с.
3. Тюрин, Ю.Н. Анализ данных на компьютере / Ю.Н. Тюрин, А.А. Макаров. – М.: ИНФРА-М: Финансы и статистика, 1995. – 384 с.