

УДК 338.45

## Иерархическая кластеризация нефтеперерабатывающих заводов России

Пермякова Т.В., taaapv@gmail.com

Канд. экон. наук, доцент Файзуллин Р.В., rf85@mail.ru

«Ижевский государственный технический университет имени М.Т. Калашникова»  
426069, г. Ижевск, ул. Студенческая, д. 7

*Данное исследование направлено на изучение нефтеперерабатывающей промышленности России с помощью иерархической кластеризации. В процессе работы проанализированы данные первичной переработки жидких углеводородов (УВ) в России нефтеперерабатывающими заводами в 2009 и 2010 гг.*

*Исследование первичной переработки УВ нефтеперерабатывающих заводов (НПЗ) крупнейших вертикально-интегрированных компаний (ВИК) России проводилось следующим образом. Каждому признаку присвоена переменная, занесены исходные данные в программу «Statistica 10.0». Далее с помощью функций стандартизация данных и иерархическая кластеризация получаем дендрограмму. Для того, чтобы сделать какие либо выводы из исследования нужно определиться с выбором числа кластеров. Для этого воспользуемся методом Уорда. В результате работы выделены основные типы нефтеперерабатывающих предприятий и определены направления развития нефтеперерабатывающей промышленности.*

**Ключевые слова:** нефтеперерабатывающие заводы, иерархическая кластеризация, углеводород.

## Hierarchical clustering of the Russia's oil-refineries

Permyakova T.V. taaapv@gmail.com

Fayzullin R.V. rf85@mail.ru

Kalashnikov Izhevsk State Technical University  
Izhevsk 426069, Studencheskayast.7

*The research is focused on studying the oil-refining industry of Russia by means of the hierarchical clustering. In the process, analyzed data of primary processing of liquid hydrocarbons (HC) in Russian refineries in 2009 and 2010.*

*Investigation of primary processing for HC refining wires (refinery) the largest vertically -integrated companies (VIC) Russia carried out as follows . Each feature is assigned a variable input data entered into the program «Statistica 10.0». Further, using the functions of standardization of data and hierarchical clustering dendrogram obtain. In order to make any conclusions from the study need to choose the number of clusters. To do this, use the method of Ward .*

*The article considers basic types of the oil processing enterprises and a determination of development directions of the oil-refining industry.*

**Keywords:** oil refineries, hierarchical clustering, углеводород.

При анализе и прогнозировании социально-экономических явлений исследователь довольно часто сталкивается с многомерностью их описания. Это происходит при решении задачи сегментирования рынка, построении типологии стран по достаточно большо-

му числу показателей, прогнозирования конъюнктуры рынка отдельных товаров, изучении и прогнозировании экономической депрессии и многих других проблем. [1,2]

Методы многомерного анализа - наиболее действенный количественный инструмент исследования социально-экономических процессов, описываемых большим числом характеристик. К ним относятся кластерный анализ, таксономия, распознавание образов, факторный анализ. Кластерный анализ наиболее ярко отражает черты многомерного анализа в классификации, факторный анализ – в исследовании связи. Иногда подход кластерного анализа называют в литературе численной таксономией, численной классификацией, распознаванием с самообучением и т.д. [3,4]

В современном мире мы сталкиваемся с огромным количеством разнообразной информации, которую приходится адаптировать и интерпретировать под различные задачи стоящие перед нами в данный момент времени.

Исходная информация в социально-экономических исследованиях представляется чаще всего в виде набора объектов, каждый из которых характеризуется рядом признаков (показателей). Поскольку число таких объектов и признаков может достигать десятков и сотен, и визуальный анализ этих данных малоэффективен, то возникают задачи уменьшения, концентрации исходных данных, выявления структуры и взаимосвязи между ними на основе построения обобщенных характеристик множества признаков и множества объектов. Такие задачи могут решаться методом иерархической кластеризации.[5]

Кластерный анализ - задача разбиения заданной выборки объектов (ситуаций) на не-пересекающиеся подмножества, называемые кластерами, так, чтобы каждый кластер состоял из схожих объектов, а объекты разных кластеров существенно отличались.[6,7]

Основное внимание иерархической кластеризации уделяется систематизации и обработке данных, которые направлены на выявление характера и структуры взаимосвязей между компонентами исследуемого многомерного признака и предназначены для получения научных и практических выводов. [8,9]

Исходным массивом многомерных данных обычно служат результаты измерения компонент многомерного признака для каждого из объектов исследуемой совокупности, т.е. последовательность многомерных наблюдений. Многомерный признак чаще всего интерпретируется как величина случайная, а последовательность наблюдений как выборка из генеральной совокупности. В этом случае выбор метода обработки исходных статистических данных производится на основе тех или иных допущений относительно природы закона распределения изучаемого многомерного признака. [10,11]

Прикладное значение иерархического кластерного анализа состоит в основном в решении следующих трех задач:

- задача статистического исследования зависимостей между рассматриваемыми показателями;
- задача классификации элементов (объектов или признаков);

- задача снижения размерности рассматриваемого признакового пространства и отбора наиболее информативных признаков. [12]

Кластерный анализ предназначен для разделения совокупностей объектов на классы, в каждый из которых должны входить объекты в определенном смысле однородные или близкие. При кластерном анализе заранее неизвестно, сколько получится групп объектов и какого они будут объема. [13]

Нефть играет ключевую роль в жизни современного общества. О важной роли энергоресурсов свидетельствует то обстоятельство, что более 70 % добываемых в мире полезных ископаемых относится к источникам энергии. Внутристрановое потребление нефти тесно связано с уровнем экономики страны [14]. Чтобы преобразовать добываемые ресурсы в энергию работают нефтеперерабатывающие заводы. Для того, чтобы провести их анализ и дать рекомендации к их более эффективной работе обратимся к методу иерархической кластеризации.

Компа- ния	НПЗ	Первичная переработка нефти, тыс.т		Загрузка устано- вок первичной пер- работки нефти, %		Мощность по сырью, тыс.т		Доля от объёма переработки нефти в РФ, %	
		20 09	20 10	200 9	2010	2	20 10	20 09	2010
Рос- нефть	Новокуйбы- шевский НПЗ	73 62	76 09	77	80	9 561	95 59	3, 1	3 ,0
	Сызранский НПЗ	63 86	65 20	72	73	8 856	88 83	2, 7	2 ,6
	Куйбышевский НПЗ	66 66	66 50	95	95	6 995	70 00	2, 8	2 ,7
	Ачинский НПЗ	71 02	74 62	100	100	7 102	74 62	3, 0	3 ,0
	Туапсинский НПЗ	52 21	45 15	100	86	5 232	52 32	2, 2	1 ,8
	Комсомоль- ский НПЗ	73 00	77 62	100	97	8 002	80 02	3, 1	3 ,1
Лукойл	"Пермнефтеорг- синтез"	1265 5	1302 6	100	100	126 42	1302 6	5,4	5,2
	"Волгограднеф- топереработка"	1125 9	1097 5	100	100	112 59	1097 5	4,8	4,4
	"Ухтанефтепере- работка"	4243	4098	100	100	424 3	4098	1,8	1,6
	"Нижегороднеф- теоргсинтез"	1579 7	1692 7	90	100	174 94	1699 5	6,7	6,8
ТНК-ВР	Саратовский НПЗ	5761	6717	89	96	650 2	6997	2,4	2,7
	Рязанская НПК	1437 3	1585 5	75	100	190 87	1585 5	6,1	6,3

Газ- пром	"Омский НПЗ"	1843 2	1984	95	95	195 04	2000 4	7,8	7,6
	"Нефтехим Салават"	5643	6607	88	86	642 7	7710	2,4	2,6
	"Сургутгазпром"	3010	3329	69	76	439 5	4397	1,3	1,3
	"Астраханьгаз-пром"	2059	2309	69	77	300 1	2999	0,9	0,9
Русс- нефть	"Орскнефтеорг-синтез"	5090	5134	77	100	662 7	5134	2,2	2,1
	"КраснодарЭко-Нефть"	2461	2518	79	84	309 9	3001	1,0	1,0
Баш- нефть	"Уфанафтехим"	7590	7817	80	82	950 0	9498	3,2	3,1
	Ново-Уфимский НПЗ	6641	6657	94	94	707 2	7074	2,8	2,7
	Уфимский НПЗ	6515	6719	83	85	787 8	7877	2,8	2,7

Рис 1 – Исходные данные объёма первичной переработки жидких УВ в России нефтеперерабатывающими заводами в 2009 и 2010 гг. [15]

Рис 1 содержит следующую информацию о нефтеперерабатывающих предприятиях России:

Первичная переработка нефти, тыс. т за 2009 г. –  $x_1$ ;

Первичная переработка нефти, тыс. т за 2010 г. –  $x_2$ ;

Загрузка установок первичной переработки нефти, % за 2009 г. –  $x_3$ ;

Загрузка установок первичной переработки нефти, % за 2010 г. –  $x_4$ ;

Мощность по сырью, тыс.т за 2009 г. –  $x_5$ ;

Мощность по сырью, тыс.т за 2010 г. –  $x_6$ ;

Доля от объёма первичной переработки нефти в РФ, % за 2009 г. –  $x_7$ ;

Доля от объёма первичной переработки нефти в РФ, % за 2009 г. –  $x_8$ .

Использование кластерного анализа для решения данной задачи наиболее эффективно. В общем случае кластерный анализ предназначен для объединения некоторых объектов в кластеры таким образом, чтобы в один кластер попадали максимально схожие, а объекты различных классов максимально отличались друг от друга. Количественный показатель сходства рассчитывается заданным способом на основании данных, характеризующих объекты [16,17].

Все кластерные алгоритмы нуждаются в оценках расстояний между кластерами или объектами, и ясно, что при вычислении расстояния необходимо задать масштаб измерений. [18]

Поскольку различные измерения используют абсолютно различные типы шкал, данные необходимо стандартизовать, так что каждая переменная будет иметь среднее 0 и стандартное отклонение 1. [19]

С помощью программы «Statistica 10.0» проводим кластерный анализ. Получаем следующие результаты. [20]

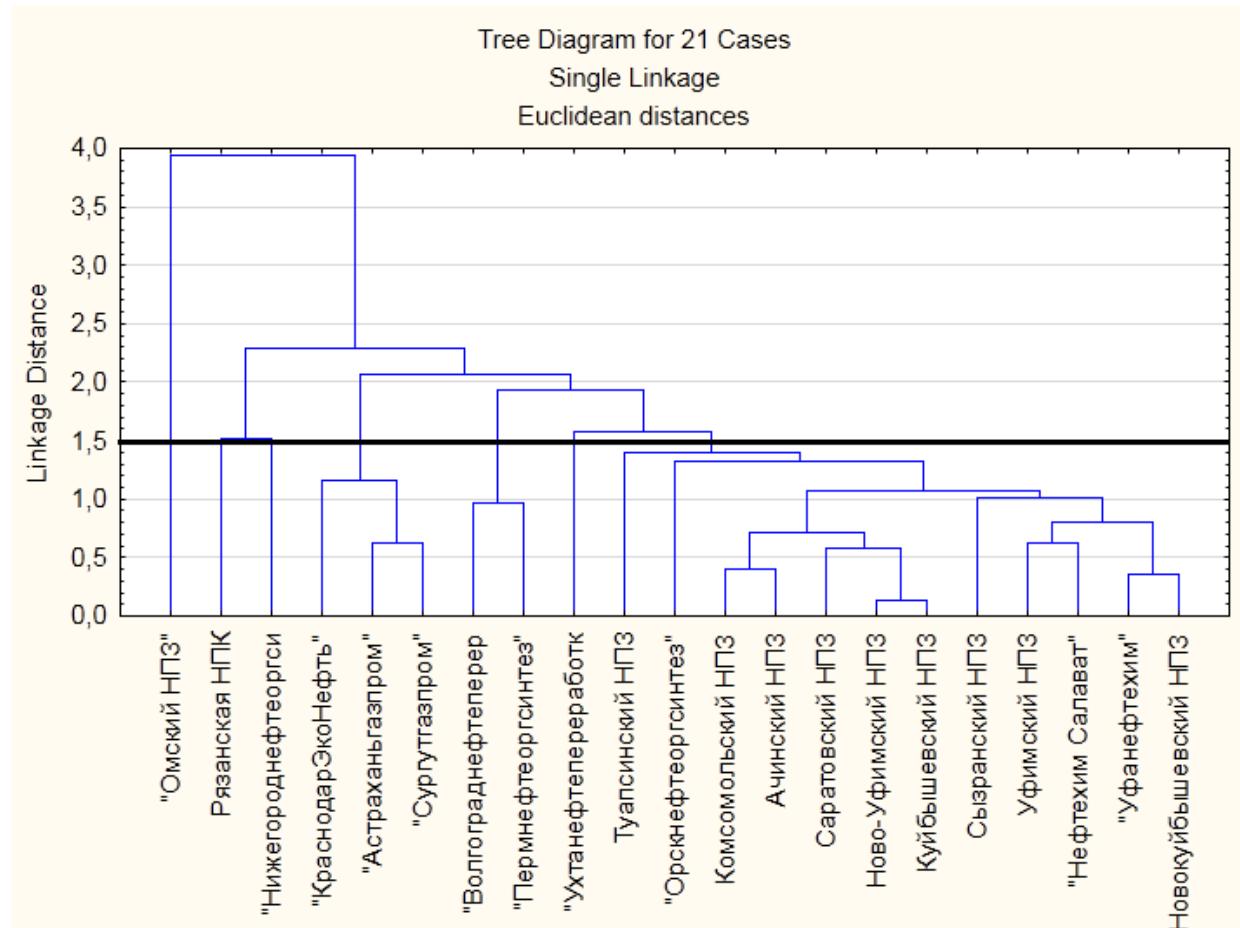


Рис 2 - дендрограмма

Дендрограмма, полученная на основе данных объёма первичной переработки жидких УВ в России предприятиями, входящими в структуру ВИНКОв, независимыми переработчиками и МНПЗ в 2009 и 2010 гг.

Двигаясь вниз, НПЗ, которые "теснее соприкасаются друг с другом", объединяются и формируют кластеры. Каждый узел диаграммы, приведенной выше, представляет объединение двух или более кластеров, положение узлов на вертикальной оси определяет расстояние, на котором были объединены соответствующие кластеры.

При построении кластеризации важной является проблема выбора числа кластеров. Предполагается, что кластеризация должна выявить естественные локальные сгущения объектов. Поэтому число кластеров  $K$  является параметром, часто существенно усложняющим вид алгоритма, если предполагается неизвестным, и существенно влияющим на качество результата, если предполагается известным.

Для выбора количества кластеров воспользуемся специальным инструментом программы «Statistica 10.0».

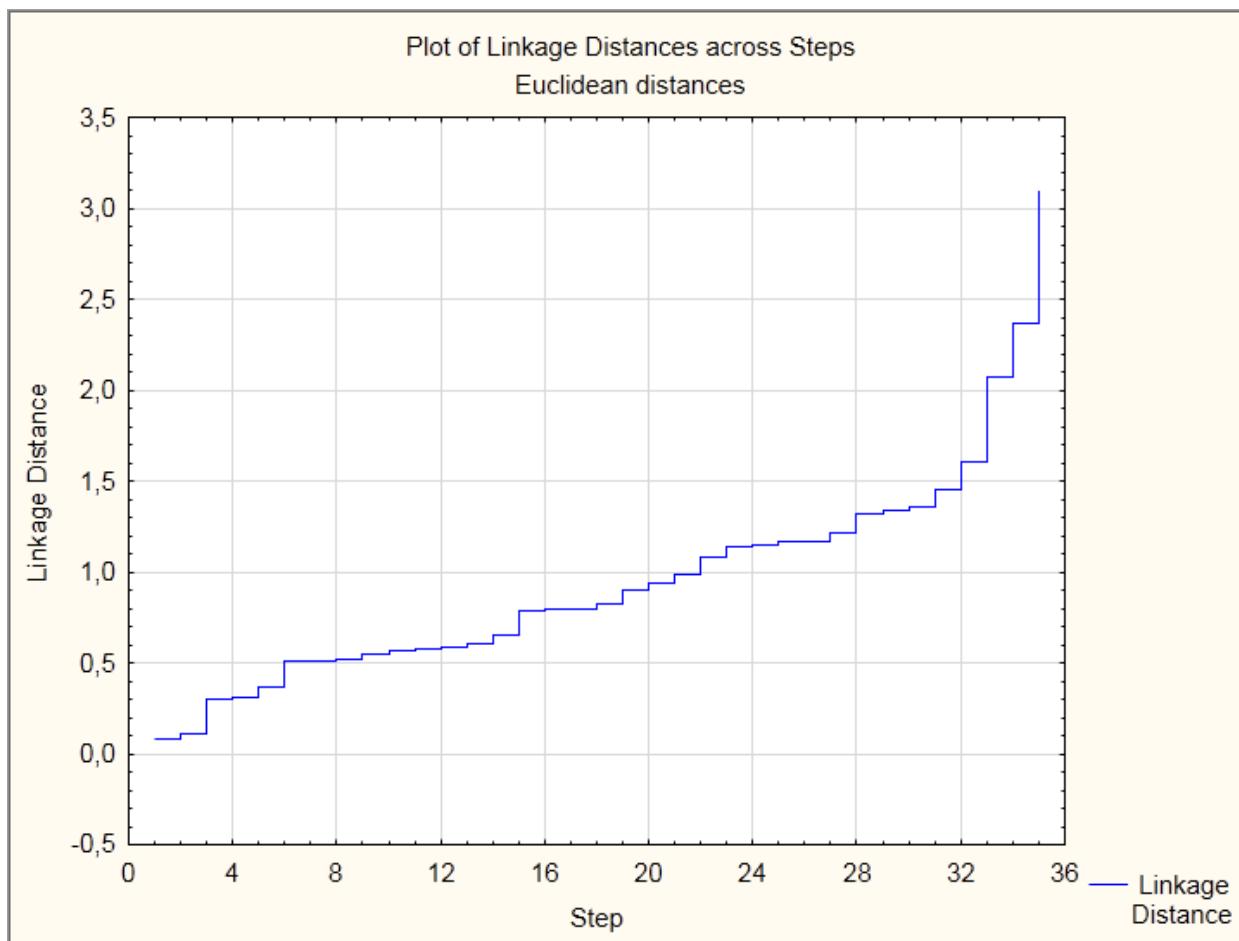


Рис 3 - график объединения объектов в классы

Мы получили график объединения объектов в классы методом Уорда.

На графике находится точка «перелома» и номер шага  $m$ , на котором произошел «перелом». В данном случае точкой перелома является шаг под номером 30. Количество классов равно  $n-m$ , где  $n$  - количество объектов в выборке. Откуда получаем,  $36-30=6$ .

Следовательно, наше предположение, основанное на аналитическом анализе, оказалось правильным, верное количество кластеров равно шести.

В результате получились кластеры, каждый из которых содержит те НПЗ, значения которых однородны. В первом кластере Омский нефтеперерабатывающий завод — один из крупнейших нефтеперерабатывающих предприятий России. Принадлежит компании «Газпром нефть». Является одним из самых современных нефтеперерабатывающих заводов России и одним из крупнейших в мире. Поэтому его характеристики существенно отличаются от других НПЗ, представленных для анализа. Можно также отметить, что показатели ОНПЗ(Омского нефтеперерабатывающего завода) растут из года в год. В настоящее время на ОНПЗ продолжается реализация среднесрочной инвестиционной программы, которая предусматривает ряд высокобюджетных проектов.

Во второй кластер вошли Рязанская НПК и "Нижегороднефтесинтез", которые уступают по средним показателям ОНПЗ, но показывают высокие результаты переработки по сравнению с остальными НПЗ России. Рязанская нефтеперерабатывающая компа-

ния (РНПК) — основное и самое мощное предприятие международной компании ТНК-ВР. Его проектная мощность — 17 млн тонн нефти в год. В результате проведенной модернизации производственные мощности увеличились почти в три раза, по производству пара — в два раза. ООО «ЛУКОЙЛ-Нижегороднефтеоргсинтез» - одно из ведущих предприятий в России мощностью по первичной переработке нефти 17 млн тонн в год. Это третье место в стране и первое в ОАО «ЛУКОЙЛ». Эти два НПЗ по своим характеристикам однородны, поэтому мы их объединяя в один кластер. Рязанская НПК и «Нижегороднефтеоргсинтез» развиваются в одном направлении, хотя и принадлежат разным компаниям, ТНК-ВР и Лукойл соответственно.

В следующий третий кластер вошли "КраснодарЭкоНефть", "Астраханьгазпром", "Сургутгазпром". Эти компании показали однородные результаты, они также принадлежат к разным компаниям "КраснодарЭкоНефть" - к РуссНефти, а "Астраханьгазпром" и "Сургутгазпром" - к Газпрому. "КраснодарЭкоНефть" - одно из старейших предприятий отрасли выпускающее высококачественную продукцию удовлетворяющее требование российских стандартов. Действующие технологические установки позволяют заводу выпускать широкий ассортимент продукции, в том числе: прямогонные бензиновые фракции, летнее дизельное топливо, печное топливо. Проектная мощность завода – 3,0 млн. тонн нефти в год. Завод перерабатывает малосернистые, малопарафинистые нефти различных месторождений. ООО «Газпром добыча Астрахань» — крупнейший газохимический комплекс, не имеющий аналогов в России. Для обеспечения устойчивой и эффективной деятельности разработана «Комплексная программа развития нефтегазодобывающего комплекса Астраханского региона до 2030 года», предусматривающая инвестиции в проектирование и строительство новых объектов, в реконструкцию и техническое перевооружение действующих производств. "Сургутгазпром" — одно из крупных газотранспортных предприятий Западной Сибири.

В четвертый кластер вошли два НПЗ компаний Лукойл - "Волгограднефтепереработка" и "Пермнефтеоргсинтез". Волгограднефтепереработка дважды (в 2004 и 2008 г.) признавалось лучшим предприятием Компании, неоднократно становилось победителем в конкурсах, проводимых областной администрацией, Российским союзом промышленников и предпринимателей. Это крупнейший производитель горюче-смазочных материалов в Южном федеральном округе. Объем переработки в 2010 г. составил 10975 тыс. тонн нефти. Глубина переработки 83,05% - один из крупнейших нефтеперерабатывающих заводов России. "Пермнефтеоргсинтез" - предприятие топливно-масляного направления, расположено в пяти километрах от миллионного города Перми, административного центра Пермского края. "Пермнефтеоргсинтез" по многим показателям уверенно занимает лидирующие позиции в отрасли. Ежегодно завод перерабатывает порядка 13 млн тонн нефти. Глубина переработки нефти достигает 93%. Более половины выпускаемых предприятием нефтепродуктов отгружается на экспорт.

В пятый кластер вошел НПЗ "Ухтанефтепереработка". Это предприятие компании Лукойл значительно отличается от других НПЗ. Ухтинский НПЗ основан в 1939 году.

Численность работающих в ОАО "ЛУКОЙЛ-Ухтанефтепереработка" 1360 человек. Годовая мощность завода по переработке сырья составляет 5.8 млн. тонн. Продукция завода находит широкий спрос в России, более половины выпускаемой продукции отгружается за рубеж. [20]

Шестой кластер включил в себя самое большое количество нефтеперерабатывающих заводов России. Это Туапсинский НПЗ, "Орскнефтеоргсинтез", Комсомольский НПЗ Ачинский НПЗ, Саратовский НПЗ, Ново-Уфимский НПЗ, Куйбышевский НПЗ, Сызранский НПЗ, Уфимский НПЗ, "Нефтехим Салават", "Уфанефтехим", Новокуйбышевский НПЗ. В этом кластере четыре НПЗ Роснефти - Туапсинский, Комсомольский, Ачинский, Сызранский и три НПЗ Башнефти - Ново-Уфимский, Уфимский, "Уфанефтехим". Все предприятия, представленные в этом кластере являются ведущими НПЗ России и самыми рентабельными и эффективными по размерам и добыче сырья, их средняя мощность - 7500 тыс.т.

Основные производства размещены преимущественно вблизи районов концентрированного потребления нефтепродуктов: в Рязанской, Ярославской, Горьковской, Ленинградской областях, Краснодарском крае, Сибири. Здесь построен крупнейший НПЗ - в Омске. Этот нефтеперерабатывающий завод выделился в отдельный кластер, его анализировать нужно индивидуально. Первый кластер составляет Омский нефтеперерабатывающий завод.

Нефтеперерабатывающая промышленность России - организационно высококонцентрированная и территориально диверсифицированная отрасль нефтегазового комплекса, обеспечивающая переработку около 50% объема жидких углеводородов, добываемых в стране. Технологический уровень большинства заводов, несмотря на проведенную в последние годы модернизацию, значительно уступает показателям развитых стран. Особенno рекомендуется провести модернизацию производства заводам, включенными в третий, пятый и шестой кластеры.

У большинства российских компаний добыча нефти значительно превышает объем перерабатывающих мощностей, а основные финансовые потоки формируются за счет добычи экспорта и сырья. Наибольший коэффициент обеспеченности перерабатывающими мощностями у «Башнефти» за счет интеграции в структуру Башкирской группы заводов. Эти заводы входят в шестой кластер, который занимает одно из лидирующих позиций в рейтинге.

В результате проведения кластерного анализа нефтеперерабатывающих предприятий России было выделено шесть кластеров. Каждый кластер включил в себя те НПЗ, которые однородны по своим параметрам первичной переработки, мощности по сырью и загрузке оборудования. Был проведен анализ каждого из выбранных кластеров. На основе этого анализа можно принимать различные экономические решения. Они могут быть связаны с инвестированием в эти нефтеперерабатывающие заводы, покупкой или продажей, проведением статистических исследований, модернизацией производства и т.д.

Рекомендации для первого кластера:

- Реализовать проекты по наращиванию экспорта нефтепродуктов;
- Модернизация части имеющихся мощностей

Рекомендации для второго кластера:

- Увеличить уровень загрузки НПЗ
- Повысить объемы финансирования НПЗ
- Ориентация на экспорт нефтепродуктов;

Рекомендации для третьего кластера:

- Сменить собственников НПЗ, что могло бы стать хорошим катализатором роста котировок
- Повысить качество продуктов, а значит и их стоимости

Рекомендации для четвертого кластера:

- Увеличить глубину переработки, так как показатель глубины переработки напрямую сказывается на стоимости НПЗ
- Модернизация оборудования и производства

Рекомендации для пятого кластера:

- Повысить качество продуктов, а значит и их стоимости
- Увеличить уровень загрузки НПЗ
- Модернизация части имеющихся мощностей.

Рекомендации для шестого кластера:

- Перейти под контроль государственных компаний, что дало бы дополнительную премию к справедливой цене - около 20%
- Развитие нефтехимии
- Развить систему нефтепродуктопроводов, что позволит нарастить экспорт нефтепродуктов

Рекомендации для нефтеперерабатывающей промышленности в целом:

- Строительство новых НПЗ;
- Продолжить модернизацию существующих НПЗ практически во всех регионах страны (европейская часть, Сибирь, Дальний Восток), а в случае наличия технических возможностей расширить их мощности по сырью;
- Построить НПЗ в Сибири и на Дальнем Востоке - территория наиболее активной нефтедобычи;
- Построить новые высокотехнологические НПЗ в европейской части страны;

Таким образом, в результате кластерного анализа нефтеперерабатывающих заводов России сформированы рекомендации для всей нефтеперерабатывающей промышленности России. Это строительство новых НПЗ, модернизация существующих НПЗ, развитие системы трубопроводов, увеличение уровня загрузки НПЗ, смена собственников, а также немаловажное значение в развитии нефтегазового комплекса является строительство и работа современных малых НПЗ.

**Список литературы:**

1. Берестнева О. Г., Воловоденко В. А. Визуализация экспериментальных многомерных данных на основе обобщенных графических образов // Вестник науки Сибири. – 2011. – № 1 (1). – С. 363-369.
2. Хорошилов В.А., Шевченко Да. Новые информационные технологии в социологических исследованиях // Социол. исслед. 1991. № 11. С. 115—120.
3. Воловоденко В. А., Берестнева О. Г. Методы визуализации в оценке состояния человека-оператора // Известия Южного федерального университета. Технические науки. – 2012. – Т. 134. – № 9. – С. 9–13.
4. Сошникова Л.А., Тамашевич В.Н., Уебе Г., Шефер М. Многомерный статистический анализ в экономике. М.: ЮНИТИ-ДАНА, 1999. - 598 с.
5. Фишер Р. А. Статистические методы для исследователей. М.: Гостехиздат, 1954, 267 с.
6. Дюран Б., Оделл П. Кластерный анализ. - М.: Статистика, - 1977, - 128 с.
7. Классификация и кластер. /Под ред. Дж. Вэн Райзина. - М.: Мир, 1980, -390 с.
8. Мандель И.Д. Кластерный анализ. - М.: Финансы и статистика. 1988. - 176 с.
9. Мостеллер Ф., Тьюки Дж. Анализ данных и регрессии. М.: Финансы и статистика, 1982.
- 10.Калинина В. Н., Соловьев В. И. Введение в многомерный статистический анализ: Учебное пособие / ГУУ. – М., 2003. – 66 с.
- 11.Жамбю М. Иерархический кластер-анализ и соответствия: Пер. с фр. М.: Финансы и статистика, 1988. - 342с.
- 12.Факторный, дискриминантный и кластерный анализ: Пер. с англ./Дж.-О. Ким, Ч. У. Миоллер, У. Р. Клекка и др.; Под ред. И. С. Енюкова. — М.: Финансы и статистика, 1989.— 215 с.
- 13.Колмогоров А.Н. Основные понятия теории вероятностей. Изд. 2-ое. - М.: Наука, 1974.- 118 с.
- 14.Глушкова А.С., Файзуллин Р.В. Методика оценки внутристранового объема потребления нефти // Проблемы экономики и управления нефтегазовым комплексом. 2013. - №10. С.36-40
- 15.Современное состояние нефтеперерабатывающей промышленности России // А. Г. Коржубаев, И. А. Соколова, А. С. Ивашин // Минеральные ресурсы России. – 2011. – № 04. – С. 50-57.
- 16.Кузнецов Д.Ю., Трошина Т.Л. Кластерный анализ и его применение // Ярославский педагогический вестник. – 2006. – Вып. 4. – С. 103-107.
- 17.Асадуллин А.Ф., Файзуллин Р.В. Анализ бюджетной обеспеченности районов и городов Удмуртской республики // Приволжский научный вестник. 2013. - №8(24). С.70-78.
- 18.Прикладная статистика: Классификации и снижение размерности: Справ. изд. / С. А. Айвазян, В. М. Бухштабер, И. С. Енюков, Л. Д. Мешалкин; Под ред. С. А. Айвазяна.— М.: Финансы и статистика, 1989,— 607 с.

- 19.Перегудов Ф.И., Тарасенко Ф.П. Основы системного анализа: Учеб. 2-е изд., доп. - Томск: Изд-во НТЛ, 1997. - 396с.
- 20.Наследов А.Д. SPSS – Компьютерный анализ данных в психологии и социальных науках. – 2-е изд. – СПб.: Питер, 2005. – 416 с.