

## **Кластерный анализ стран на основе оценок ФАТФ**

Д.С. Налбандян

студент 5 курса НИЯУ МИФИ, Москва

Email: davanlb8@gmail.com

Д.В. Домашова

доцент кафедры финансового мониторинга НИЯУ МИФИ, Москва

Email: janedom@mail.ru

*Аннотация: В данной статье проводится кластерный анализ стран на основе оценок взаимных оценок ФАТФ с использованием метода иерархической кластеризации (метод Уорда), алгоритма К-средних и графика локтя для определения оптимального числа кластеров. Применение этих методов позволило выявить устойчивые группы стран с схожими профилями соответствия стандартам ПОД/ФТ. Результаты исследования подтверждают возможность структурирования стран по оценкам ФАТФ, что открывает перспективы для более точной группировки стран по уровню соответствия ПОД/ФТ и повышения эффективности международного мониторинга.*

*Ключевые слова: оценки ФАТФ, кластерный анализ, К-средних, метод Уорда, метод локтя, кластеризация*

## **Cluster analysis of countries based on FATF estimates**

D.S. Nalbandyan

4th year specialist's student at NRNU MEPhI, Moscow

Email: davanlb8@gmail.com

J.V. Domashova

Associate professor, Department of financial monitoring

NRNU MEPhI, Moscow

Email: janedom@mail.ru

*Abstract: This article provides a cluster analysis of countries based on FATF mutual estimates using the hierarchical clustering method (Ward's method), the K-means algorithm and the elbow graph to determine the optimal number of clusters. The use of these methods has made it possible to identify stable groups of countries with similar profiles of compliance with AML/CFT standards. The results of the study confirm the possibility of structuring countries according to FATF estimates, which opens up prospects for a more accurate grouping of*

*countries according to the level of AML/CFT compliance and improving the effectiveness of international monitoring.*

*Keywords: FATF estimates, cluster analysis, K-means, Word method, elbow method, clustering*

Современная система противодействия отмыванию доходов и финансированию терроризма (ПОД/ФТ) играет ключевую роль в обеспечении финансовой стабильности и экономической безопасности государств. Группа разработки финансовых мер борьбы с отмыванием денег (ФАТФ) разработала комплекс из 40 рекомендаций и 11 показателей эффективности, служащих международным стандартом оценки соответствия юрисдикций системе ПОД/ФТ. Результаты взаимных оценок ФАТФ оказывают значительное влияние на восприятие стран на международной арене, инвестиционную привлекательность, а также финансово-экономическое взаимодействие с другими государствами и институтами. Однако остается открытым вопрос: существуют ли устойчивые группы стран по уровню соответствия стандартам ФАТФ, и каково их распределение? Кластеризация на основе оценок ФАТФ позволяет выявить типологии стран, обладающих схожими характеристиками в области ПОД/ФТ, и тем самым содействует более точной оценке странового риска, а также оптимизации международных политик в сфере финансового мониторинга. Кроме того, типологизация может быть полезна для дальнейшего совершенствования методологии ФАТФ, в том числе при распределении технической помощи и планировании взаимных оценок. Настоящее исследование направлено на выявление таких групп с использованием методов кластерного анализа. В работе применяются алгоритмы *k*-средних (*k*-means) и иерархической кластеризации (метод Уорда), а также метод локтя для определения оптимального количества кластеров и метод главных компонент (РСА) для визуализации результатов. Выбор данных ограничен исключительно оценками ФАТФ, что позволяет сосредоточиться на качественной структуре международной ПОД/ФТ-системы и выявить закономерности без влияния макроэкономических факторов.

Кластерный анализ представляет собой метод многомерного статистического анализа, направленный на выявление в исследуемых данных естественных групп (кластеров), внутри которых объекты обладают максимальной схожестью, а между кластерами наблюдаются существенные различия. В рамках настоящего исследования проводится кластеризация стран исключительно на основе оценок ФАТФ. Для кластеризации данных применяется алгоритм *k*-means (*k*-средних), один из наиболее широко используемых методов, принцип которого основан на минимизации внутрикластерной дисперсии. На начальном этапе алгоритм случайным образом выбирает *k*-центроидов (центральных точек

кластеров), после чего каждый объект данных присваивается ближайшему центроиду [1]. Затем центроиды пересчитываются как среднее значение всех объектов в кластере, и процесс повторяется до достижения их устойчивости (сходимости). Так как результат метода чувствителен к выбору числа кластеров, в исследовании применяется метод локтя (elbow method), позволяющий определить оптимальное количество кластеров. График локтя строится путем нанесения значений суммы квадратов расстояний между объектами и центроидами (within-cluster sum of squares, WCSS) для различных значений  $k$ . Оптимальное число кластеров соответствует точке, где темпы снижения WCSS существенно замедляются, что визуально напоминает форму «локтя» [2]. Дополнительно используется иерархическая кластеризация с применением метода Уорда, который также направлен на минимизацию внутрикластерной дисперсии. Метод позволяет построить дендрограмму, отражающую последовательное объединение стран в группы. В отличие от  $k$ -means, метод Уорда не требует предварительного задания количества кластеров и дает возможность визуально определить их оптимальное число [3]. В данном исследовании дендрограмма используется как первый шаг для определения структуры данных и подтверждения устойчивости кластеризации, выполненной методом  $k$ -means. Для визуализации кластеров в низкоразмерном пространстве применяется метод главных компонент (Principal Component Analysis, PCA), позволяющий отразить структуру многомерных данных на плоскости при сохранении наиболее значимой информации. Метод преобразует исходные переменные (оценки по 40 рекомендациям ФАТФ и 11 показателям эффективности) в новые координаты, главные компоненты, которые максимально объясняют дисперсию данных [4]. Визуализация в пространстве первых двух главных компонент облегчает интерпретацию результатов кластерного анализа.

Дендрограмма, построенная с использованием метода Уорда, помогает определить оптимальное количество кластеров, анализируя моменты, где расстояния между объединяемыми кластерами резко возрастают. В данном исследовании дендрограмма будет использована как первый шаг кластерного анализа для визуализации структуры данных и выбора подходящего количества кластеров для дальнейшей кластеризации. Методом Уорда при пороговом значении расстояния  $\rho_{\text{пор}} = 18$  все субъекты разбиваются на 4 класса, каждый класс был помечен разным цветом, результат показан на рисунке 1, а состав кластеров приведён в таблице 1.

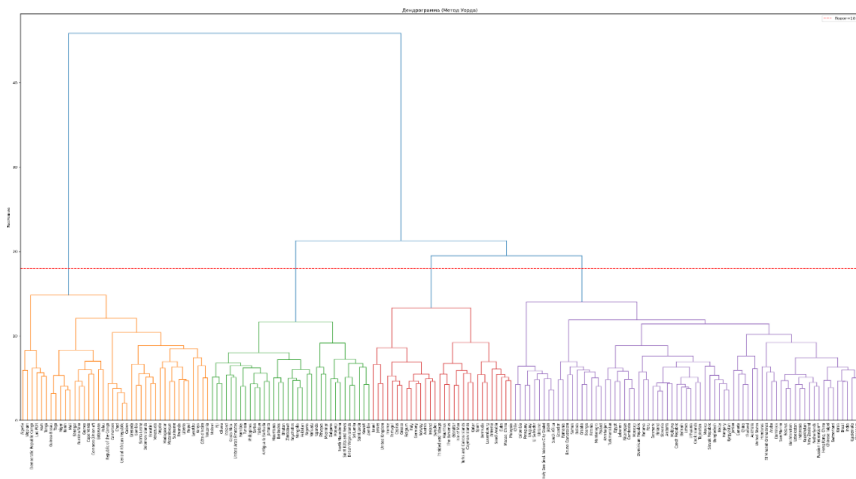


Рисунок 1 – Дендрограмма Уорда при пороговом значении 18

Таблица 1 – Разбиение на 4 кластера с помощью метода Уорда

Номер кластера	Количество наблюдений в кластере	Состав кластера
0	40	Algeria, Angola, Benin, Botswana, Burkina Faso, Cameroon, Cape Verde, Central African Republic, Chad, Comores (Union of), Côte d'Ivoire, Democratic Republic Congo, Eswatini, Gabon, Gambia, Grenada, Guinea, Guinea-Bissau, Haiti, Kenya, Lao PDR, Lesotho, Liberia, Madagascar, Mali, Mozambique, Nepal, Niger, Palau, Republic of the Congo, Rwanda, Senegal, Sierra Leone, Solomon Islands, Suriname, Tanzania, Togo, Tonga, Venezuela, Vietnam.
1	73	Andorra, Aruba, Australia, Azerbaijan, Bahrain, Bangladesh, Belarus, Bolivia, Brazil, Brunei Darussalam, Cambodia, Canada, Chile, China, Chinese Taipei, Colombia, Cook Islands, Croatia, Czech Republic, Denmark, Dominica, Dominican Republic, Ecuador, Egypt, El Salvador, Estonia, Finland, Georgia, Guatemala, Holy See (incl. Vatican City State), Honduras, Hong Kong, China, Hungary, India, Indonesia, Japan, Jordan, Kazakhstan, Korea, Kyrgyzstan, Latvia, Lebanon, Liechtenstein, Lithuania, Malta, Mexico, Moldova, Montenegro, Morocco, Netherlands, New Zealand, Nicaragua, Panama, Paraguay, Peru, Poland, Romania, Russian Federation**, Samoa, San Marino, Singapore, Slovak Republic, Slovenia, South Africa, St Vincent-Grenadines, Switzerland, Tajikistan, Thailand, Turkmenistan, Ukraine, United States, Uruguay, Uzbekistan
2	34	Antigua & Barbuda, Barbados, Bhutan, British Virgin Islands, Bulgaria, Costa Rica, Ethiopia, Fiji, Ghana, Iceland, Jamaica, Kuwait, Malawi, Mauritania, Monaco, Mongolia, Myanmar, Namibia, Nigeria, North Macedonia, Pakistan, Philippines, Saint Kitts and Nevis, Saint Lucia, Serbia, Seychelles, Sri Lanka,

		Tunisia, Türkiye, Uganda, United Arab Emirates, Vanuatu, Zambia, Zimbabwe
3	30	Armenia, Austria, Belgium, Bermuda, Cayman Islands, Cuba, Cyprus, France, Germany, Gibraltar, Greece, Ireland, Isle of Man, Israel, Italy, Jersey, Luxembourg, Macao, China, Malaysia, Mauritius, Norway, Portugal, Qatar, Saudi Arabia, Spain, Sweden, The Bahamas, Trinidad and Tobago, Turks and Caicos Islands, United Kingdom

Далее мы проведем кластерный анализ, используя метод **K-means**, который позволяет разделить данные на заданное количество кластеров, минимизируя вариацию внутри кластеров и обеспечивая четкую интерпретацию результатов. Для определения оптимального числа кластеров был применен метод локтя. Этот метод основан на анализе изменения инерции (суммы квадратов расстояний от точек до центров кластеров) при увеличении количества кластеров. Как видно на рисунке 2, точка «излома» наблюдается при значении числа кластеров, равном 2 и 4. Для дальнейшего анализа было решено разделить данные на 4 кластера, как и по методу Уорда.

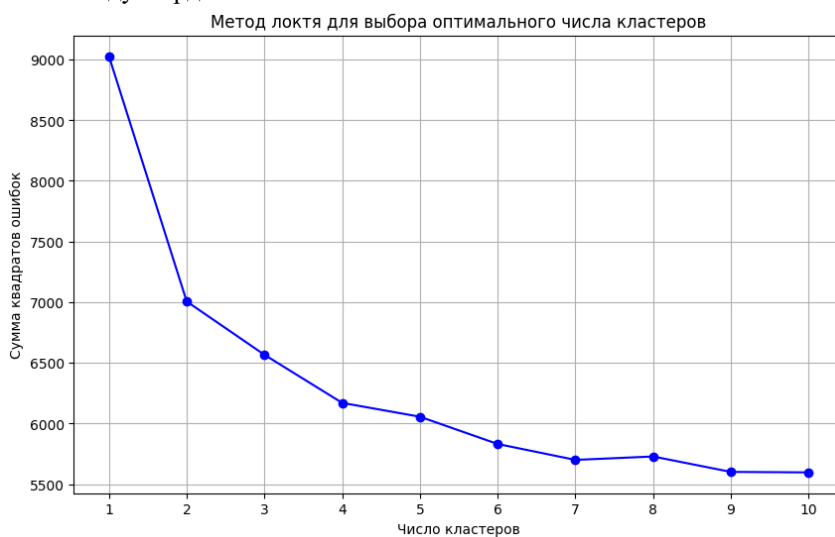


Рисунок 2 – Метод локтя

Для разбиения стран на группы была использована модель кластеризации K-means с четырьмя кластерами. В результате каждая страна была отнесена к одному из четырех кластеров, что представлено в таблице 2. Такое разбиение позволяет группировать страны по схожести их характеристик, связанных с показателями эффективности и технического соответствия.

Таблица 2 – Разбиение на 4 кластера с помощью K-means

Номер кластера	Количество наблюдений в кластере	Состав кластера
0	41	Algeria, Angola, Benin, Cameroon, Cape Verde, Central African Republic, Chad, Comores (Union of), Côte d'Ivoire, Democratic Republic Congo, Eswatini, Gabon, Gambia, Grenada, Guinea, Guinea-Bissau, Haiti, Kenya, Lao PDR, Lesotho, Liberia, Madagascar, Mali, Mozambique, Myanmar, Nepal, Niger, Palau, Republic of the Congo, Rwanda, Samoa, Senegal, Sierra Leone, Solomon Islands, Suriname, Tanzania, Togo, Tonga, Uganda, Venezuela, Vietnam
1	39	Slovenia, Türkiye, Gibraltar, Malta, Belgium, Greece, Germany, Luxembourg, Isle of Man, Cyprus, Austria, Qatar, Norway, Cayman, Islands, Macao, China, Malaysia, Cuba, United Arab Emirates, Italy, The Bahamas, Panama, Singapore, Saudi Arabia, Ireland, Mauritius, Aruba, Jersey, Turks and Caicos Islands, Uruguay, Bermuda, Iceland, Armenia, Paraguay, Spain, Zimbabwe, Trinidad and Tobago, Malawi, Vanuatu, United Kingdom
2	35	Andorra, Australia, Bahrain, Belarus, Brunei Darussalam, Canada, Chile, China, Chinese Taipei, Cook Islands, Egypt, Estonia, Finland, France, Guatemala, Hong Kong, China, India, Indonesia, Israel, Japan, Kazakhstan, Korea, Liechtenstein, Mexico, Netherlands, New Zealand, Peru, Portugal, <b>Russian Federation</b> , San Marino, Sweden, Switzerland, Thailand, United States, Uzbekistan
3	62	Antigua & Barbuda, Azerbaijan, Bangladesh, Barbados, Bhutan, Bolivia, Botswana, Brazil, British Virgin Islands, Bulgaria, Burkina Faso, Cambodia, Colombia, Costa Rica, Croatia, Czech Republic, Denmark, Dominica, Dominican Republic, Ecuador, El Salvador, Ethiopia, Fiji, Georgia, Ghana, Holy See (incl. Vatican City State), Honduras, Hungary, Jamaica, Jordan, Kuwait, Kyrgyzstan, Latvia, Lebanon, Lithuania, Mauritania, Moldova, Monaco, Mongolia, Montenegro, Morocco, Namibia, Nicaragua, Nigeria, North Macedonia, Pakistan, Philippines, Poland, Romania, Saint Kitts and Nevis, Saint Lucia, Serbia, Seychelles, Slovak Republic, South Africa, Sri Lanka, St Vincent-Grenadines, Tajikistan, Tunisia, Turkmenistan, Ukraine, Zambia

После того, как мы получили результаты классификации по двум методам кластерного анализа перед нами, стоит задача выбора наилучшего результата. Сделать это можно с помощью т. н. функционала разбиения. Мы будем опираться на следующий функционал: разбиение считается самым эффективным, если сумма квадратов расстояний от каждого значения кластера до среднего значения по всем кластерам минимальна:

$$Q_1(S) = \sum_{l=1}^p \sum_{O_i \in S_l} d^2(O_i, \bar{X}(l)) \rightarrow \min$$

3. Результат вычисления функционалов разбиения представлен в таблице

Таблица 3 – Функционал качества по двум методам

№ Кластера	Уорд	К-средних
0 Кластер	822,975	862,2927
1 Кластер	1253,589	675,0256
2 Кластер	573,6176	541,7714
3 Кластер	524,3	1069,484
Сумма	3174,482	3148,574

По результатам видно, что метод К-средних показывает наилучшие значения. Для интерпретации кластеров построим график средних значений признаков в кластерах по К-средних, который показан на рисунке 3.

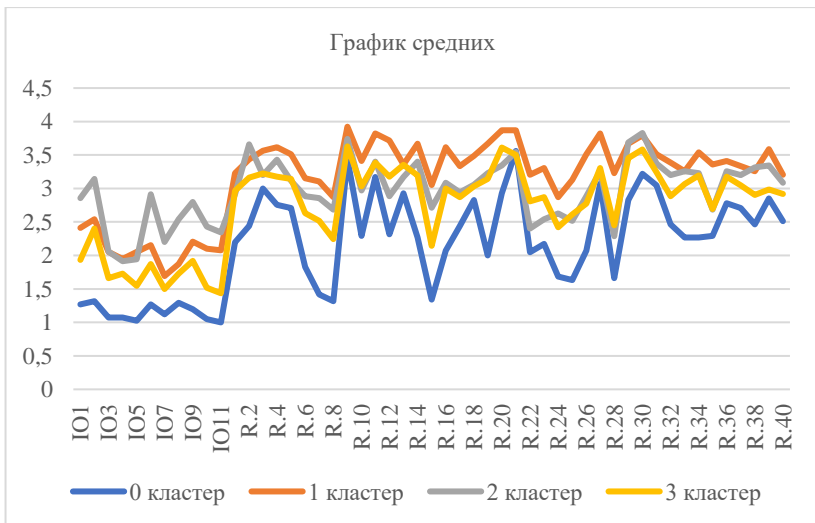


Рисунок 3 – График средних значений в кластерах

**Кластер 0** объединяет 41 страну, среди которых типичными представителями являются Гвинея-Бисау, Венесуэла и Того. Страны этого кластера демонстрируют самые низкие результаты по ключевым

индикаторам эффективности и рекомендациям в области борьбы с отмыванием денег (AML) и финансированием терроризма (CFT). Основными проблемами для стран этого кластера являются недостаточное понимание рисков и слабая национальная координация действий, что затрудняет разработку эффективных мер противодействия. Международное сотрудничество также находится на низком уровне, что выражается в ограниченном обмене информацией и трудностях при запросе данных для международных расследований. Надзор за финансовыми учреждениями и другими субъектами, подверженными рискам отмывания денег, является недостаточным, а применение риск-ориентированного подхода не реализовано в полной мере. Кроме того, существуют значительные проблемы с прозрачностью юридических лиц и доступностью данных о бенефициарной собственности, что затрудняет выявление конечных выгодоприобретателей. Правоохранительные органы сталкиваются с трудностями в проведении расследований, а вынесенные санкции часто оказываются недостаточно эффективными, чтобы предотвратить дальнейшие нарушения. Превентивные меры, такие как идентификация клиентов и отчетность о подозрительных транзакциях, также реализуются на недостаточно высоком уровне. Для стран данного кластера первоочередными мерами являются улучшение оценки национальных рисков и их внедрение в политику, укрепление координации между различными государственными органами и повышение прозрачности юридических структур. Важным шагом является налаживание международного взаимодействия, включая участие в программах и инициативах FATF. Необходимы реформы в сфере надзора за финансовыми и нефинансовыми субъектами с внедрением жестких требований и санкций за их несоблюдение. Улучшение работы правоохранительных органов, в том числе усиление их ресурсной базы, должно сопровождаться активным внедрением превентивных мер, таких как более строгая проверка клиентов и повышение осведомленности о рисках. Таким образом, решение этих проблем позволит значительно повысить стандарты AML/CFT в странах кластера 0, снизить риски и улучшить их международную репутацию.

**Кластер 1** включает 31 страну, среди которых типичными представителями являются Словения, Турция, Гибралтар. Эти страны демонстрируют самые высокие оценки среди всех кластеров по рекомендациям FATF (Recommendations, R), что свидетельствует о развитой нормативной базе и политических механизмах в сфере противодействия отмыванию денег (AML) и финансированию терроризма (CFT). Высокий уровень соответствия международным стандартам проявляется в таких направлениях, как оценка рисков и применение риск-ориентированного подхода, прозрачность бенефициарной собственности, строгий надзор за финансовыми институтами и нефинансовыми

субъектами, а также применение целевых финансовых санкций. Эти страны активно участвуют в международном сотрудничестве и демонстрируют значительные успехи в укреплении своих механизмов по борьбе с финансовыми преступлениями. Однако, несмотря на развитую нормативную базу, показатели стран по Immediate Outcomes (IO) занимают не самое высокое место среди четырех кластеров, что указывает на определенные проблемы в реализации установленных стандартов. Хотя в целом страны хорошо понимают риски, связанные с отмыванием денег и финансированием терроризма, и предпринимают значимые действия для их минимизации, уровень координации между национальными органами и эффективность практической реализации остаются неравномерными. Превентивные меры, такие как применение механизмов финансового контроля и выявление подозрительных операций, требуют большей согласованности. Кроме того, расследования и преследования по делам о финансировании терроризма и отмывании денег проходят успешно, но их эффективность может быть усилена, особенно в части конфискации доходов и инструментов преступной деятельности. Для дальнейшего укрепления позиций этим странам следует уделить внимание улучшению координации между государственными органами, устранению барьеров, мешающих эффективной реализации существующих мер, и усилению ресурсной базы национальных ведомств. Кроме того, важно продолжать работу над внедрением практического риск-ориентированного подхода и адаптировать превентивные меры к актуальным угрозам. Усиление международного сотрудничества также останется ключевым элементом, позволяющим обмениваться передовыми практиками и повышать скорость взаимодействия при расследованиях. Таким образом, страны кластера 1 обладают высоким потенциалом для дальнейшего укрепления своих механизмов и достижения лучших результатов в практической реализации мер борьбы с финансовыми преступлениями. Кластер 2 Liechtenstein, Indonesia, Bahrain

**Кластер 2** включает 35 стран, среди которых типичными представителями являются Лихтенштейн, Индонезия и Бахрейн. Эти страны показывают самые высокие результаты среди всех кластеров по Immediate Outcomes (IO), что свидетельствует об их высокой эффективности в реализации мер по борьбе с отмыванием денег (AML) и финансированием терроризма (CFT) на практике. Это выражается в глубоком понимании рисков, успешной координации национальных органов, эффективной работе с финансовыми разведывательными данными и результативности правоприменительных мер. Например, страны демонстрируют значительный успех в таких направлениях, как пресечение деятельности террористических организаций, применение целевых финансовых санкций и расследование финансовых преступлений, сопровождающееся конфискацией активов. Однако, несмотря на

лидирующую позицию по практическим результатам, страны данного кластера занимают второе место по показателям соответствия международным рекомендациям FATF (Recommendations, R). Это означает, что, хотя их нормативно-правовая база и политические инструменты достаточно развиты, в них могут быть выявлены области для совершенствования. В частности, страны могут испытывать сложности в области прозрачности бенефициарной собственности, взаимодействия с нефинансовыми секторами и внедрения риск-ориентированных подходов в некоторых сегментах. Для усиления своих позиций странам из кластера 2 следует сосредоточиться на нескольких ключевых направлениях. Во-первых, стоит продолжать адаптацию законодательства и регулирующих механизмов к актуальным международным стандартам. Это включает развитие более четких норм в отношении прозрачности правовой собственности и усиление надзора за нефинансовыми секторами, такими как DNFBPs (дилеры драгоценных металлов и камней, бухгалтеры и адвокаты). Во-вторых, необходимо инвестировать в укрепление институционального потенциала для дальнейшего развития сотрудничества между национальными ведомствами и международными партнёрами. Кроме того, эти страны могут использовать свои сильные стороны для распространения лучших практик, улучшая механизмы обнаружения подозрительных транзакций и более эффективно внедряя меры по предотвращению преступной деятельности. Совершенствование механизмов обучения и повышения квалификации персонала государственных и частных структур также позволит добиться более высоких результатов. Таким образом, страны кластера 2 демонстрируют отличные результаты в практической реализации мер противодействия AML/CFT и имеют значительный потенциал для дальнейшего совершенствования своей нормативной базы и инструментов предотвращения финансовых преступлений.

**Кластер 3** включает 62 страны, включая такие типичные примеры, как Сербия, Литва и Венгрия, которые демонстрируют оценки ниже среднего по обоим основным показателям: Immediate Outcomes (IO) и Recommendations (R). Это отражает недостатки как в нормативной базе, так и в практическом применении мер по борьбе с отмыванием денег (AML) и финансированием терроризма (CFT). Страны этого кластера занимают третье место среди четырёх групп, что указывает на необходимость значительного повышения эффективности их подходов в этой области. Оценки по Immediate Outcomes свидетельствуют о низкой результативности применения мер AML/CFT. Эти страны сталкиваются с трудностями в обеспечении прозрачности бенефициарной собственности, эффективном расследовании преступлений, связанных с отмыванием денег и финансированием терроризма, а также в конфискации доходов от преступной деятельности. Слабость системы отражается также в

ограниченных возможностях международного сотрудничества и недостаточном уровне координации между компетентными органами. Низкий уровень соответствия международным рекомендациям FATF по блоку Recommendations указывает на необходимость реформирования нормативной базы. Страны демонстрируют пробелы в регулировании финансовых институтов и нефинансовых секторов (DNFBPs), в механизмах предотвращения использования юридических лиц и соглашений в преступных целях, а также в системах мониторинга подозрительных транзакций. Для этих стран важно сосредоточиться на усилении правовой и институциональной базы. Это требует обновления законодательства в соответствии с международными стандартами, улучшения координации между надзорными органами и правоприменительными структурами, а также модернизации технической инфраструктуры для выявления и анализа подозрительных транзакций. Кроме того, значительные усилия должны быть направлены на развитие международного сотрудничества, обмен информацией и эффективное использование взаимной правовой помощи. Успех реформ зависит также от инвестиций в обучение кадров, особенно в правоохранительных органах, финансовой разведке и надзорных структурах. Использование современных аналитических технологий может повысить эффективность мониторинга и выявления преступной активности. Кроме того, усиление санкций за несоблюдение требований и расширение механизмов предотвращения преступлений способны укрепить систему AML/CFT. В целом, страны этого кластера находятся на пути реформ, и при реализации вышеуказанных мер они смогут значительно повысить устойчивость к финансовым преступлениям, укрепить доверие к своим финансовым системам и внести больший вклад в глобальные усилия по противодействию отмыванию денег и финансированию терроризма.

После выполнения кластеризации, для упрощения анализа и визуализации результатов, был использован метод главных компонент (PCA). Он позволил сократить размерность признакового пространства до двух компонент, что значительно облегчило интерпретацию и визуализацию кластеров в двумерном пространстве. Результаты кластеризации и снижения размерности были представлены в виде рисунка под номером 4, на котором объекты (страны) отображены в зависимости от их принадлежности к различным кластерам, что позволяет наглядно увидеть структуру данных и связи между объектами в контексте выбранных признаков.

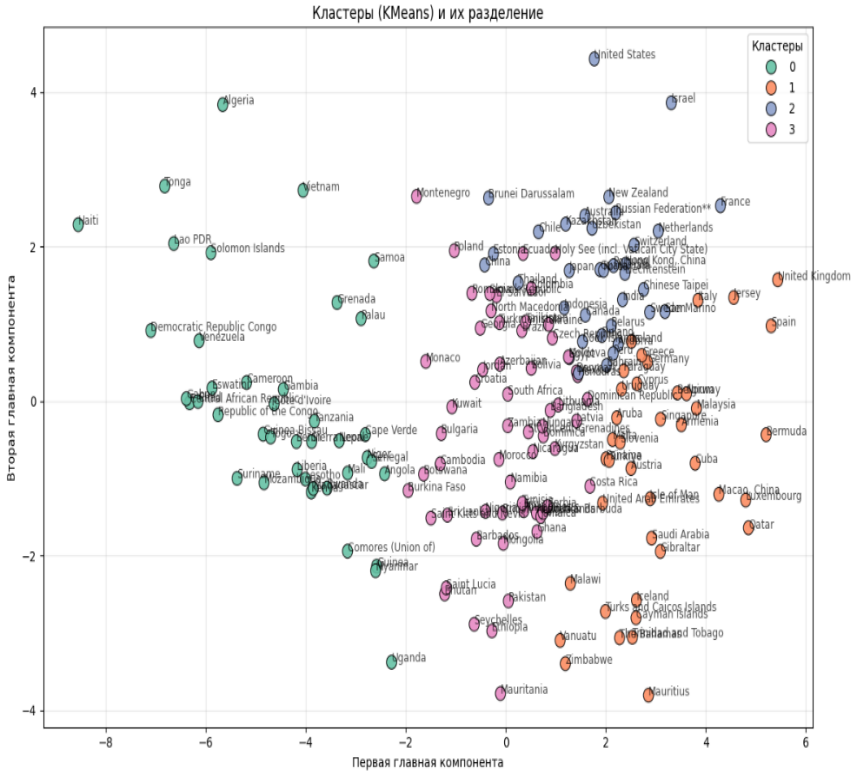


Рисунок 4 – Кластеризация K-means

После выполнения кластеризации мы применили метод опорных векторов (SVM) для построения модели классификации стран. Этот подход позволяет учитывать границы разделения между кластерами и эффективно прогнозировать принадлежность к группе для новых данных. Для визуализации высокоразмерных данных было использовано понижение размерности с помощью метода главных компонент (PCA). Это позволило представить данные в двухмерном пространстве, сохраняя при этом основную информацию о различиях между странами. Кластеры были четко разделены в проекциях на главные компоненты, что подтверждает качество прогноза и кластеризации. результат показан на рисунке 5.

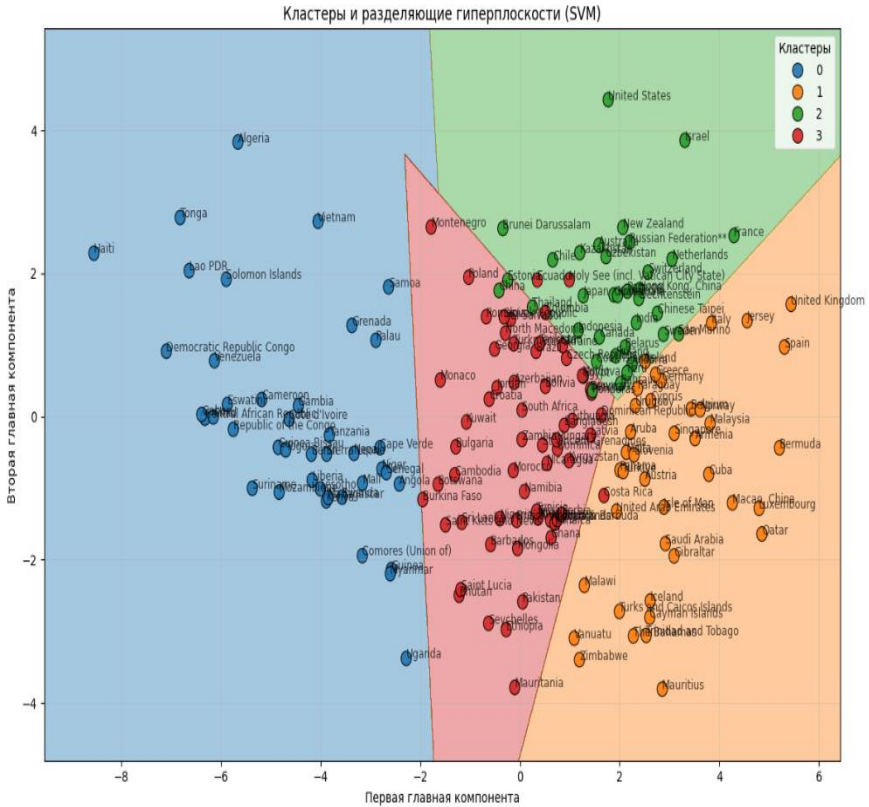


Рисунок 5 – Применение метода опорных векторов

Для оценки прогностической способности модели была проведена разбивка данных на обучающую и тестовую выборки в соотношении 70% на 30%. Обученная модель показала точность классификации на уровне **94.4%** при проверке на тестовой выборке, что свидетельствует о высокой предсказательной способности модели, графически это можно увидеть на рисунке 6 и 7.

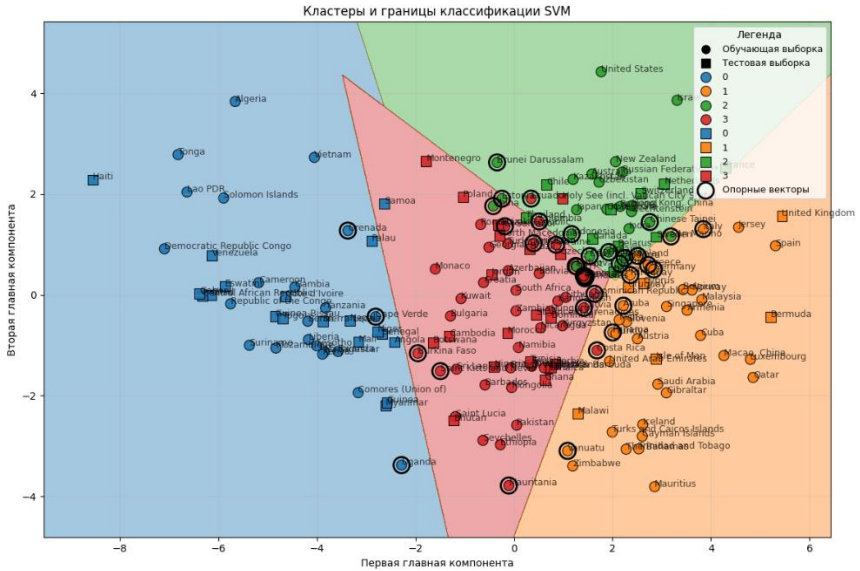


Рисунок 6 – Деление на обучающую и тестовую выборку

Точность модели: 94.44%

Отчет по классификации (precision, recall, F1-score):

	precision	recall	f1-score	support
0	1.00	0.95	0.97	19
1	1.00	1.00	1.00	6
2	0.90	0.90	0.90	10
3	0.90	0.95	0.92	19
accuracy			0.94	54
macro avg	0.95	0.95	0.95	54
weighted avg	0.95	0.94	0.94	54

Рисунок 7 – Точность модели классификации

Модель продемонстрировала высокую точность классификации, с общей точностью в 94%. При этом наблюдаются различия в показателях между кластерами. Для кластера 0 точность составила 1, что означает отсутствие ложных срабатываний, а полнота – 0,95, что свидетельствует о том, что 95% объектов были правильно классифицированы, но 5% (1 объект) оказались ошибочно отнесены к другим кластерам. F1-score, как сбалансированная метрика точности и полноты, составил 0,97, что

подтверждает высокую надежность модели в отношении этого кластера. С учетом 19 объектов в кластере (support), эти показатели можно считать значимыми. Кластер 1 показал идеальные результаты, с точность, полнота и F1-score равными 1, что свидетельствует о полной точности и полноте классификации. Однако малый объем данных в кластере (support = 6) делает метрики чувствительными к изменениям выборки. Для кластера 2 точность составил 0,9, что говорит о том, что 90% предсказанных объектов действительно принадлежали этому кластеру, а 10% оказались ложными срабатываниями. Полнота также равна 0,9, что показывает, что 10% объектов кластера не были распознаны моделью. F1-score, объединяющий эти показатели, составил 0,9, что свидетельствует о хорошем, но не идеальном уровне работы модели для данного кластера. Support в 10 объектов подтверждает умеренную значимость этих метрик. В кластере 3 точность равен 0,9, что указывает на 10% ложных срабатываний, а полнота составляет 0,95, отражая, что 95% объектов кластера были корректно классифицированы. F1-score достиг 0,92, демонстрируя сбалансированную точность и полноту. Как и в случае с кластером 0, этот кластер имеет 19 объектов (support), что делает метрики устойчивыми.

Проведенный кластерный анализ позволил выявить структурные различия между странами на основе их эффективности и соответствия международным стандартам в сфере борьбы с отмыванием денег и финансированием терроризма (AML/CFT). Использование иерархического метода Уорда на первом этапе дало возможность определить предварительное количество кластеров и визуализировать структуру данных. В дальнейшем метод K-means, подтвержденный методом локтя, позволил получить более четкое и интерпретируемое разбиение стран на четыре кластера. Сравнение функционалов качества разбиения показало преимущество метода K-means, обеспечившего минимальную суммарную дисперсию внутри кластеров. Это подтвердило высокую обоснованность выбранного способа кластеризации. Анализ содержательного наполнения каждого из кластеров позволил установить как сильные стороны, так и ключевые зоны для улучшения в разных группах стран — от государств с низким уровнем реализации и нормативной базы до стран, демонстрирующих высокие результаты в практическом применении мер AML/CFT. Визуализация с применением метода главных компонент (PCA) способствовала более наглядному представлению структуры данных, а построенная классификационная модель на основе метода опорных векторов (SVM) показала высокую точность прогнозирования (94,4%), что подтверждает состоятельность подхода и применяемых методов анализа. Таким образом, полученные результаты могут быть использованы как в целях дальнейшего научного анализа, так и как прикладной инструмент для разработки целенаправленных рекомендаций по усилению мер AML/CFT в различных странах в зависимости от их кластерной принадлежности.

Список использованных источников:

1. Татарников Вадим Владимирович, Пестунов Игорь Алексеевич, Бериков Владимир Борисович Алгоритм усреднения центроидов для построения кластерного ансамбля // КО. 2017. №5. URL: <https://cyberleninka.ru/article/n/algorithm-usredneniya-tsentroidov-dlya-postroeniya-klasterного-ansamblya> (дата обращения: 23.03.2025).
2. Бизнес-аналитика на основе больших данных: обучение без учителя на языках Python и R : учеб.-метод. пособие / С. В. Рындина. – Пенза : Изд-во ПГУ, 2020. – 76 с.
3. Ахмедов К. М., Дзидзава Э. Т. ОБЗОР МЕТОДОВ ЕДИНОЙ КЛАСТЕРИЗАЦИИ // Вестник магистратуры. 2021. №1-1 (112). URL: <https://cyberleninka.ru/article/n/obzor-metodov-edinoy-klasterizatsii> (дата обращения: 23.03.2025).
4. Харисова А.Ф., Бакуменко Л.П. Применение метода главных компонент для анализа производственных показателей на предприятиях // Экономика и менеджмент инновационных технологий. 2017. № 2 [Электронный ресурс]. URL: <https://ekonomika.snauka.ru/2017/02/13907>.