

## КЛАСТЕРНЫЙ АНАЛИЗ ВИДОВ БЕТОННЫХ СМЕСЕЙ

*Гребёнкин Александр Михайлович*, студент 4 курса бакалавриата института экономики и управления АПК, ФГБОУ ВО РГАУ-МСХА имени К.А. Тимирязева, [ag18032003@mail.ru](mailto:ag18032003@mail.ru)

*Научный руководитель – Дашиева Баярма Шагдаровна*, канд. экон. наук, доцент кафедры статистики и кибернетики, ФГБОУ ВО РГАУ-МСХА имени К.А. Тимирязева, [dashieva.b.sh@rgau-msha.ru](mailto:dashieva.b.sh@rgau-msha.ru)

**Аннотация.** В работе проведен кластерный анализ видов бетонных смесей, изучен их состав и прочностные характеристики. Используются два метода кластеризации: иерархический агломеративный и итеративный (*k-means*). В результате проведенного анализа выделено три кластера видов бетонных смесей, характеризующихся разным содержанием компонент бетона, влияющих на его прочность.

**Ключевые слова:** кластерный анализ, бетонные смеси, классификация, физические свойства, прочность.

## CLUSTER ANALYSIS OF TYPES OF CONCRETE MIXTURES

*Grebenkin Alexander Mikhailovich*, 4th year undergraduate student of the Institute of Economics and Management of the Agro–Industrial Complex, Russian State Agrarian University – Moscow Timiryazev Agricultural Academy, [ag18032003@mail.ru](mailto:ag18032003@mail.ru)

*Scientific supervisor – Dashieva Bayarma Shagdarovna*, Ph.D in Economic Science, Associate Professor of the Department of Statistics and Cybernetics, Russian State Agrarian University – Moscow Timiryazev Agricultural Academy, [dashieva.b.sh@rgau-msha.ru](mailto:dashieva.b.sh@rgau-msha.ru)

**Annotation.** The paper presents a cluster analysis of concrete mixtures, their composition and strength characteristics. Two clustering methods are used: hierarchical agglomerative and iterative (*k-means*). As a result of the analysis, three clusters of concrete mix types are identified, characterized by different contents of concrete components that affect its strength.

**Key words:** Cluster analysis, Concrete mixtures, Classification, Physical properties, Strength.

Кластерный анализ является мощным инструментом в области статистики и машинного обучения, который позволяет выделять однородные группы

объектов одновременно по нескольким признакам и выявлять скрытые структуры в данных [1]. В работе проведен кластерный анализ различных видов бетона, основываясь на характеристиках, таких как: содержание цемента, доменного шлака, летучей золы, воды, суперпластификатора, крупного и мелкого заполнителей, а также возраст и прочность материала. Информация о характеристиках бетонных смесей взята из сайта <https://www.kaggle.com/> (набор данных – **concrete\_data.csv**). Кластерный анализ проводился с использованием языка программирования Python и среды разработки Google Colab.

Набор данных был проверен на наличие пропусков. Далее была проведена стандартизация данных, в связи с тем, что у изучаемых признаков разный масштаб и разные единицы измерения. Оптимальное количество кластеров определено с использованием метода локтя, в результате которого был получен график зависимости суммы квадратов внутрикластерных расстояний (WCSS) от количества кластеров (рисунок 1) [3]. Из рисунка 1 можно предположить, что оптимальное число кластеров равняется трем.

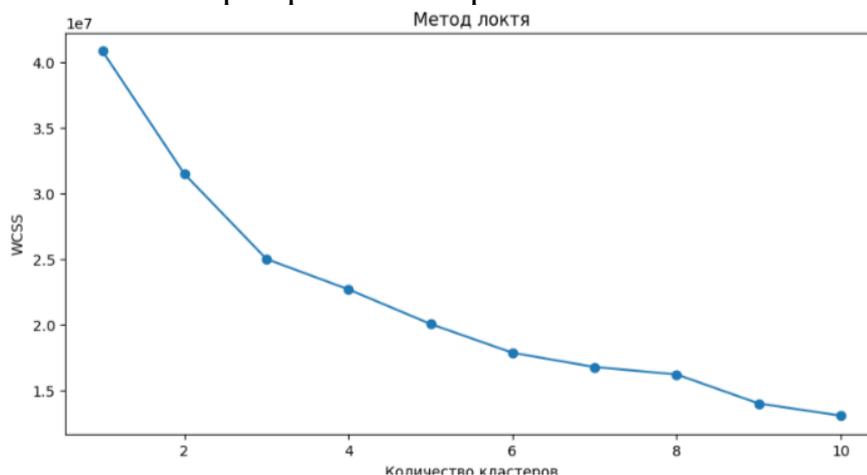


Рисунок 1 – График метода локтя

При проведении иерархического агломеративного кластерного анализа была выбрана функция близости – Евклидово расстояние и метод объединения – метод Варда, при использовании которого минимизируется минимальная дисперсия внутрикластерных расстояний. На рисунке 2 приведен фрагмент матрицы расстояний. В результате построенная дендрограмма позволяет сделать вывод о том, что выделяются 3 кластера видов бетонных смесей (Рисунок 3).

```
# Сформируем матрицу расстояний
Z = linkage(data, method='ward')
print(Z)

[[1.75000000e+02 1.77000000e+02 0.00000000e+00 2.00000000e+00]
 [1.79000000e+02 1.03000000e+03 0.00000000e+00 3.00000000e+00]
 [1.82000000e+02 1.03100000e+03 0.00000000e+00 4.00000000e+00]
 ...
 [2.05200000e+03 2.05500000e+03 2.47517727e+03 4.38000000e+02]
 [2.05100000e+03 2.05400000e+03 3.38419641e+03 5.92000000e+02]
 [2.05600000e+03 2.05700000e+03 4.09045362e+03 1.03000000e+03]]
```

Рисунок 2 – Матрица расстояний

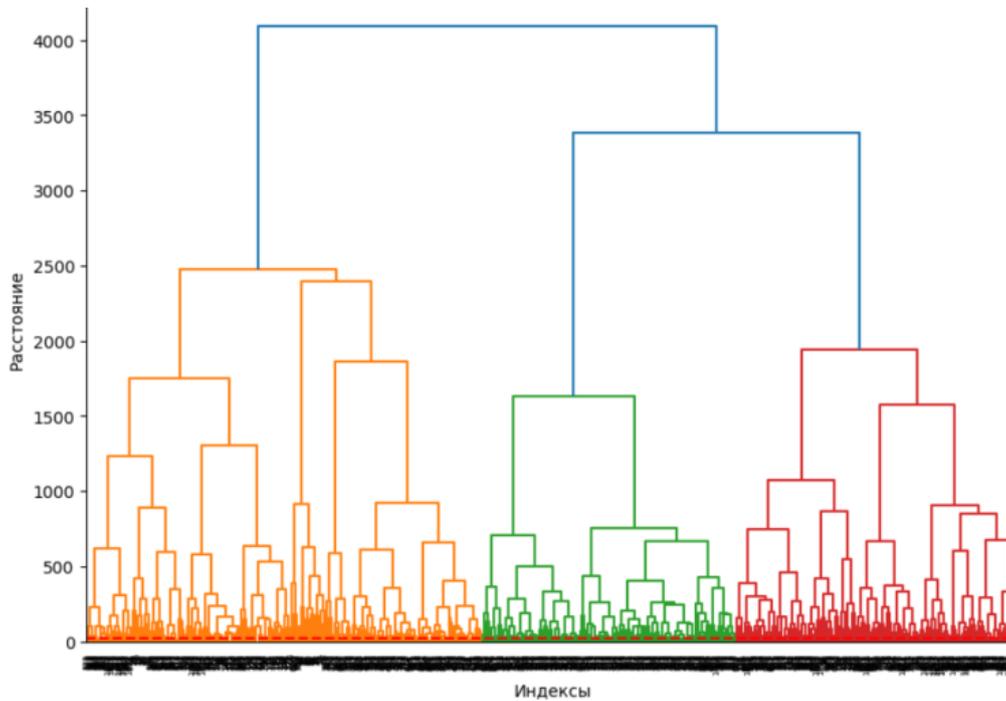


Рисунок 3 – Дендрограмма

Далее выполнили итеративную кластеризацию методом K-Means. При использовании данного метода задано 3 кластера, исходя из предыдущего анализа. Начальные центры кластеров были заданы по самым наиболее отдаленным точкам кластеров. В результате получена следующая диаграмма скопления точек, распределенных на три кластера (Рисунок 4).

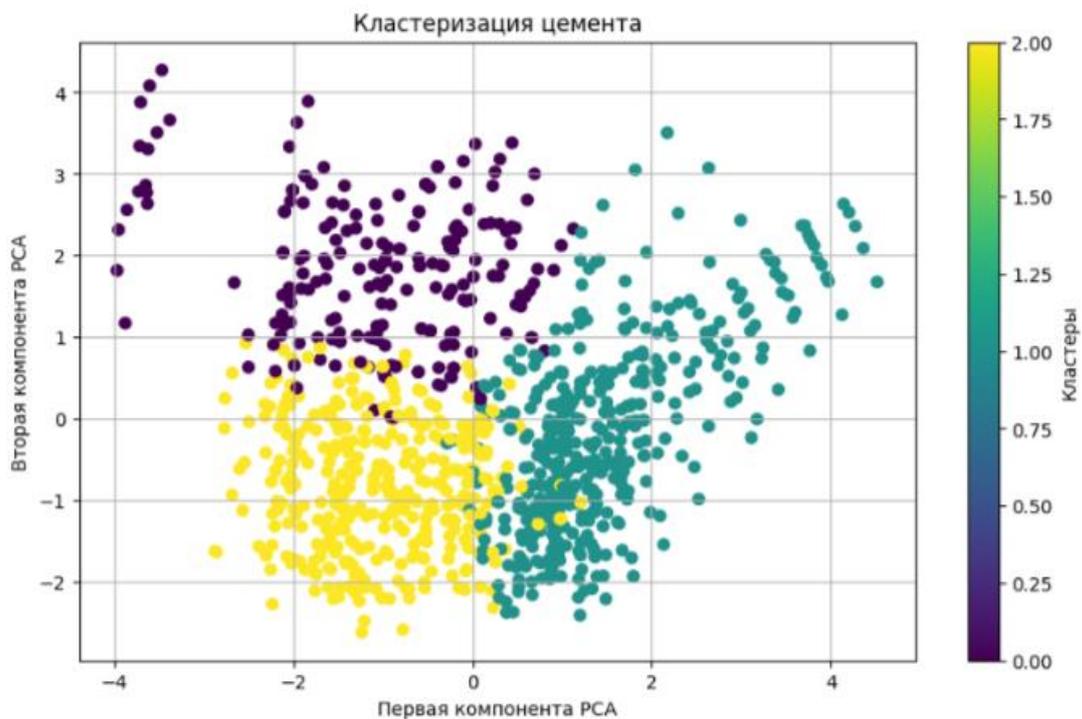


Рисунок 4 – Точечная диаграмма с распределением объектов по классам

Теперь можно интерпретировать кластеры и увидеть, чем они отличаются друг от друга. Для интерпретации в датасет добавим новый столбец «Cluster», а после средние значения характеристик по каждому кластеру (рисунок 5).

Проведем сравнение кластеров: выделим крепкие и хрупкие характеристики бетонных смесей. Для этого рассчитаем среднее значение характеристик для каждого кластера и сравним их с помощью графика, чтобы понять, какой кластер лучше. Среднее в `positive_ntr` должно быть выше, в `negative_ntr` – ниже. Словарь средних значений характеристик для каждого кластера: {0: {'Крепкие': {'Cement': 374.57191780821915, 'Strength': 41.704360730593606, 'Coarse Aggregate': 959.8159817351599}, 'Хрупкие': {'Water': 183.50844748858447, 'Fly Ash': 21.04041095890411, 'Blast Furnace Slag': 48.36940639269406}}, 1: {'Крепкие': {'Cement': 204.67124600638977, 'Strength': 31.965654952076676, 'Coarse Aggregate': 955.2258785942493}, 'Хрупкие': {'Water': 187.95303514376997, 'Fly Ash': 43.778594249201284, 'Blast Furnace Slag': 167.6153354632588}}, 2: {'Крепкие': {'Cement': 220.35232974910394, 'Strength': 30.898709677419355, 'Coarse Aggregate': 1013.3383512544804}, 'Хрупкие': {'Water': 171.35591397849464, 'Fly Ash': 117.90537634408602, 'Blast Furnace Slag': 8.829032258064517}}}. Для сравнения кластеров построены круговые диаграммы содержания компонентов бетона в % (Рисунок 6).

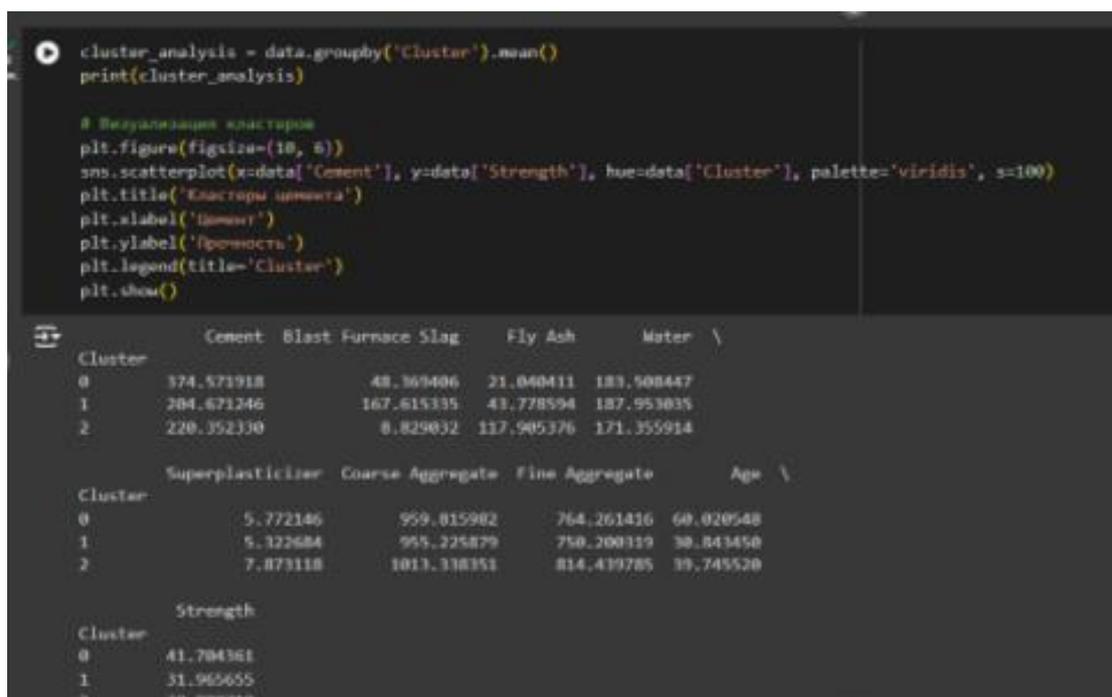


Рисунок 5 – Средние значения признаков в каждом кластере

Сравнение трех кластеров по значениям характеристик:

Среднее содержание крепких материалов в расчете на 1 м<sup>3</sup> смеси:

Кластер 0: цемент – 374,57 кг, что указывает на высокую прочность (41.70 Мпа) и устойчивость к разрушению; крупный заполнитель – 959,82 кг, что говорит о стабильной структуре.

Кластер 1: цемент – 204,67 кг, что может указывать на более низкую прочность (31.97 МПа); крупный заполнитель – 955.23 кг.

Кластер 2: цемент – 220,35 кг; прочность – 30.90 МПа; крупный заполнитель – 1013,34 кг, что может улучшить прочностные характеристики.

Среднее содержание хрупких материалов в расчете на 1 м<sup>3</sup> смеси:

Кластер 0: летучая зола – 21,04 кг, вода – 183,51 кг, доменный шлак – 48,37 кг.

Кластер 1: летучая зола – 43,78 кг, что может улучшить прочностные характеристики; вода – 187,95 кг, что улучшает растворимость; доменный шлак – 167.62 кг, что повышает устойчивость к агрессивной среде.

Кластер 2: летучая зола – 117,91 кг, что указывает на высокую степень полимеризации; вода – 171,36 кг; доменный шлак – 8,83 кг, что может негативно влиять на прочность [4].

Таким образом, кластерный анализ, проведенный двумя способами: иерархическим агломеративным и итеративным методом k-means, показал, что выделяется 3 кластера видов бетонных смесей: кластер 0 с крепким бетоном, кластер 1 с относительно крепким составом и кластер 2 с хрупким составом [2].

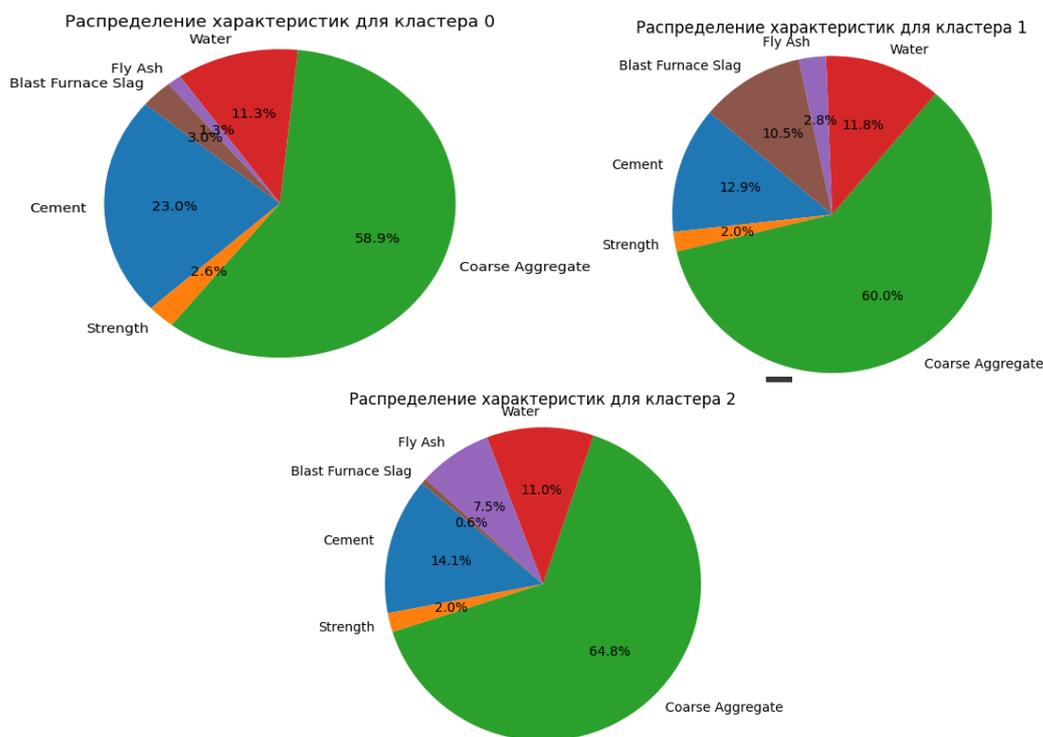


Рисунок 6 – Удельный вес содержания компонентов бетона по каждому кластеру

### Библиографический список

1. Балакин, А. А. Кластерный анализ: теория и практика. – Москва: Наука, 2019. – 240 с.
2. Белов, И. П. Исследование свойств цементов. – Санкт-Петербург: Стройиздат, 2018. – 300 с.

3. Лозинский, А. В. Научные основы кластерного анализа. – Уфа: БашГУ, 2016. – 180 с.

4. Масленников, Ю. Г. Математические методы в исследовании цемента. – Ростов-на-Дону: Ростовский университет, 2020. – 175 с.

5. Математическая статистика : учебник / А. П. Зинченко, М. В. Кагирова, Ю. Н. Романцева [и др.]. – Москва : Российский государственный аграрный университет - МСХА им. К.А. Тимирязева, 2018. – 199 с. – ISBN 978-5-9675-1681-8. – EDN JAWIE.

6. Быков, Д. В. Кластерный анализ на основе многомерных средних по результатам ВСХП-2016 с использованием Python / Д. В. Быков, А. В. Уколова // Бухучет в сельском хозяйстве. – 2022. – № 12. – С. 834. – DOI 10.33920/sel-11-2212-05. – EDN CMRECI.

7. Уколова, А. В. Типизация личных подсобных хозяйств методом нейросетевого кластерного анализа / А. В. Уколова, Д. В. Быков // Экономика сельского хозяйства России. – 2023. – № 6. – С. 97-107. – DOI 10.32651/236-97. – EDN WIJZBS.