

## ОЦЕНКА РЕКРЕАЦИОННОГО ПОТЕНЦИАЛА ГОРОДОВ БЕЛАРУСИ МЕТОДАМИ КЛАСТЕРНОГО АНАЛИЗА

**Ю. В. Зиновик, А. А. Сазонов**

*Белорусский государственный университет, Ленинградская 16, 220030,  
г. Минск, Беларусь, jualialuix@gmail.com*

В работе формировалась первичное оценивание рекреационного потенциала для районных городов Беларуси на основе методики кластерного анализа. Сформированные кластеры позволили разделить все районные города по потенциалу рекреационной инфраструктуры и провести последующий анализ с выявлением городов, требующих рекреационного развития.

**Ключевые слова:** кластерный анализ; рекреационный потенциал; оценивание.

## ASSESSMENT OF THE RECREATION POTENTIAL OF CITIES OF BELARUS BY METHODS OF CLUSTER ANALYSIS

**J. Zinovik, A. A. Sazonau**

*Belarusian State University,  
Minsk, Nezavisimosti Ave., 4, 220030, Belarus*

The work formed a primary assessment of the recreational potential for regional cities of Belarus based on the cluster analysis technique. The formed clusters made it possible to divide all regional cities according to the potential of recreational infrastructure and carry out subsequent analysis to identify cities requiring recreational development.

**Keywords:** cluster analysis; recreational potential; assessment.

Методы кластерного анализа в географических исследованиях находят все новое применение. Так, кластерный анализ используется для группировки регионов, где кластерный анализ может помочь выделить группы регионов с похожими характеристиками, что облегчит сравнение и анализ; выявление паттернов: этот метод помогает выявить пространственные паттерны в данных, что важно для понимания распределения явлений, таких как население, климат, или экономические показатели. Анализ кластеров может быть полезен при принятии решений в области геопланирования, например, при определении оптимальных мест для строительства инфраструктуры или размещения предприятий.

В свою очередь, метод К-средних (K-means) является одним из наиболее популярных методов кластерного анализа, и у него есть несколько преимуществ. Этот алгоритм эффективен на больших объемах данных, прост

и понятен при интерпретации результатов. Однако, стоит отметить, что K-средних также имеет ограничения, такие как чувствительность к выбросам и неспособность обрабатывать кластеры произвольной формы. Тем не менее, во многих случаях он остается эффективным инструментом для кластерного анализа.

На подготовительном этапе были собраны данные по объектам рекреационной инфраструктуры городов-районных центров Беларуси следующих категорий: заведения азартных игр, кинотеатры и театры, музеи и выставки, заведения общественного питания, парки и зоопарки, центры спорта. Источником данных стали базы данных OpenStreetMap, Яндекс Карты.

Далее, для проведения кластерного анализа, необходимым представлялось выделить количественный параметр у данных. С этой целью были подсчитаны все уникальные категории выделенной инфраструктуры в пределах каждого районного города. Затем, данные значения были выражены в процентном отношении к общему количеству всех категорий в каждом городе.

Прежде, чем провести дифференциацию рекреационного потенциала районных городов Беларуси на отдельные группы, проводилось предварительное выделение оптимального количества таких групп (кластеров).

Оптимальным количеством рекреационной классификации городов экспертным методом определено в выделении пяти кластеров. Именно данное количество кластеров позволило учесть и разномасштабность исследуемых территорий и выделить закономерность кластеризации.

Все необходимые расчёты, а также выполнение самого кластерного анализа осуществлялось в программной среде ArcGIS Pro с использованием метода кластеризации «K-means» и метода инициализации «Случайные местоположения начальных объектов». В результате была получена следующая классификация районных городов Беларуси (рис. 1).

Таким образом, районные города Беларуси были разделены на пять кластеров оценки рекреационного потенциала.

Чтобы выявить закономерность данной кластеризации необходимо взглянуть на построенные диаграммы размаха (рис. 2).

На (рис. 2) предоставлены основные параметры столбцов диаграмм, которые позволяют судить о разбросе данных: минимальное и максимальное значения, значения первого и третьего квартилей, медианы, а также межквартильный размах (мера вариативности одномерного распределения случайной величины. Определяется через квантили, а именно принимается равной разнице между 75-м и процентилями, то есть между третьим и первым квартилями) [1].

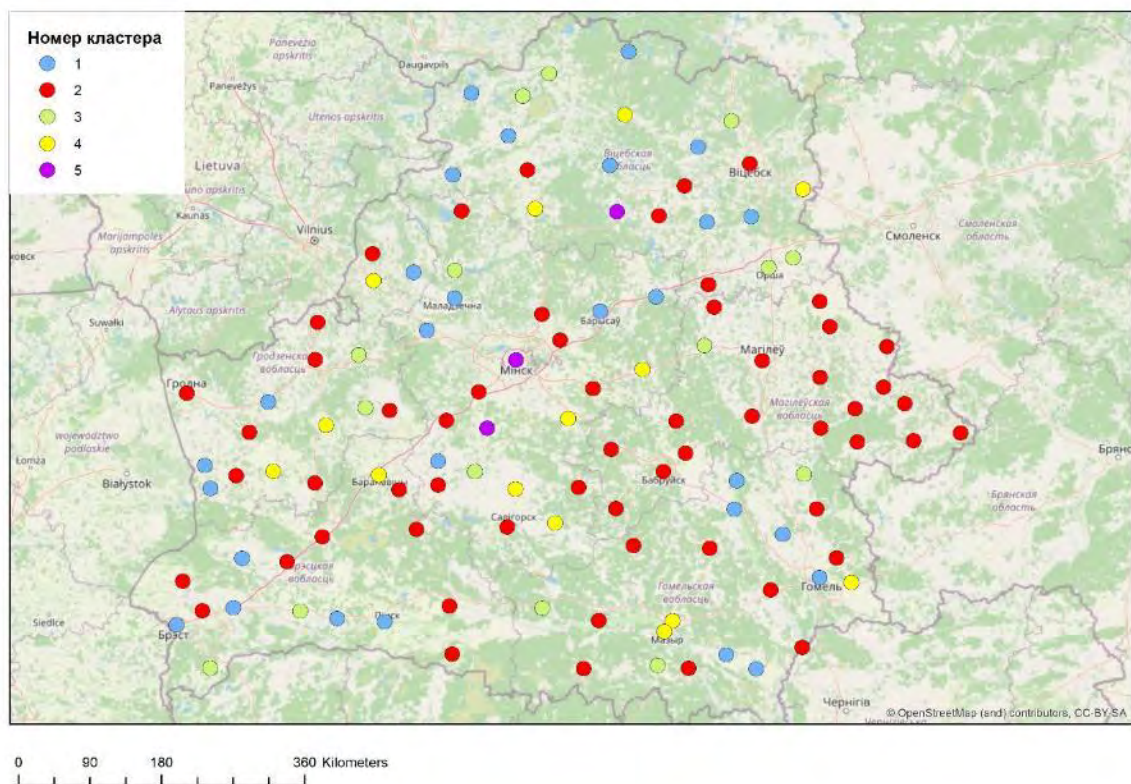


Рис. 1. Распределение районных городов

Стоит обратить внимание на оси построенной диаграммы.

Ось ординат представлена стандартизированным значением атрибутов (включает z-преобразование, при котором среднее значение всех значений вычитается из каждого значения и делится на стандартное отклонение для всех значений). Таким образом, стандартизация z-показателя помещает все атрибуты в один и тот же масштаб, позволяя визуализировать несколько распределений на одной диаграмме [2].

Ось абсцисс представлена выделенными категориями оценивания рекреационного потенциала (заведения азартных игр, кинотеатры и театры, музеи и выставки, заведения общественного питания, парки и зоопарки, центры спорта).

Анализируя полученную диаграмму, можно отметить относительно равномерное распределение для второго и четвертого кластеров. Иная ситуация прослеживается для первого, третьего и пятого кластеров, при этом значения пятого кластера носят аномальный характер и значительно выбиваются из общего ряда анализируемых категорий (в частности категории музеи и выставки, заведения общественного питания).

Для объяснения полученных результатов и выявления закономерности стоит проанализировать более подробную диаграмму размаха, которая

позволяет продемонстрировать составляющие значения при выделении того или иного кластера (рис. 3).

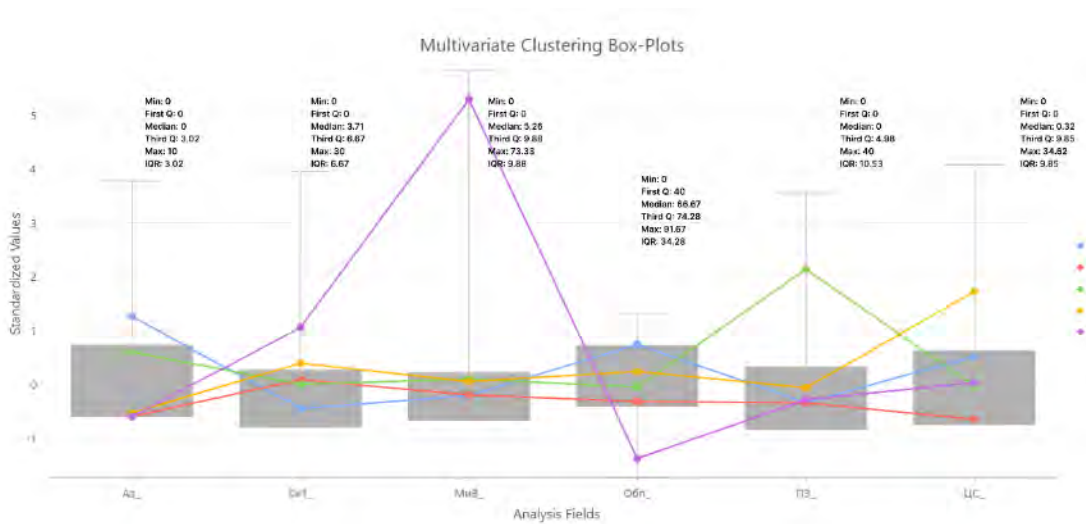


Рис. 2. Диаграмма размаха построенных кластеров



Рис. 3. Диаграмма размаха построенных кластеров (со слагающими)

Главный принцип выделения кластеров строится на максимальном их различии между собой, то есть, чем больше выражена разница значений, тем больше вероятность выделения их в отдельные составляющие. Так, из представленной диаграммы (рис. 3) можно сказать, что решающую роль в выделении второго кластера сыграли наибольшее максимальное значение кинотеатров и театров, наибольшее максимальное и минимальное значения заведений общественного питания. При выделении третьего кластера основными параметрами выступили наибольшее максимальное значение заведений азартных игр, парков и зоопарков, оптимальные (близкие к медиане) значения кинотеатров и театров и центров спорта.

Пики пятого кластера, указывают на контрасты внутри группировки городов по рекреационному потенциалу. Такое распределение свидетельствует о необходимости пересмотреть данные на выявление ошибок при подготовке данных.

Таким образом, анализируя слагающие пятого кластера видно, что его выделение произошло на наибольшем максимальном значении музеев и выставок, наибольшем минимальном значении заведений общественного питания.

Об аномальности выделения пятого кластера также свидетельствует количество вошедших в данный кластер городов (рис. 4).

Четвёртый кластер имеет наиболее оптимальное (близкое к медианным значениям) распределение.

Первый кластер, в отличие от других четырёх, имел главный фактор распределения на основе минимальной частоты встречаемых значений.

Из проведённого кластерного анализа по формированию рекреационной оценки районных городов Беларуси, можно сделать ряд выводов.



Рис. 4. Диаграмма количественного распределения городов по кластерам

Если исходить из проведённого анализа, то наибольший потенциал рекреационного развития имеют города с оптимальными значениями распределения, относительно выделенных категорий рекреационной инфраструктуры. Это города, вошедшие в четвёртый кластер, затем немного меньший балл могут получить города второго кластера, далее первого, третьего и пятого.

Следует отметить, что проведённый анализ позволил продемонстрировать методику многофакторного кластерного распределения пространственных данных. Сами же полученные результаты имеют экспериментальный характер.

Таким образом, использование методов кластерного анализа в географических исследованиях может значительно облегчить анализ пространственных данных и выделение закономерностей, что полезно для принятия решений в различных областях.

### **Библиографические ссылки**

1. Understanding Statistics. Oxford University Press, 1996 [Electronic resource]. URL: [https://en.wikipedia.org/wiki/Interquartile\\_range](https://en.wikipedia.org/wiki/Interquartile_range) (date of access: 12.11.2023).

2. How Multivariate Clustering works : ArcGIS Pro [Electronic resource]. URL: <https://pro.arcgis.com/en/pro-app/3.0/tool-reference/spatial-statistics/how-multivariate-clustering-works.htm> (date of access: 12.11.2023).