








# “Evaluating development prospects of smart cities: Cluster analysis of Kazakhstan’s regions”

<b>AUTHORS</b>	Ivan Digel  Dinara Mussabalina  Marat Urdabayev   Nurbakhyt Nurmukhametov  Aigul Akparova 
<b>ARTICLE INFO</b>	Ivan Digel, Dinara Mussabalina, Marat Urdabayev, Nurbakhyt Nurmukhametov and Aigul Akparova (2022). Evaluating development prospects of smart cities: Cluster analysis of Kazakhstan’s regions. <i>Problems and Perspectives in Management</i> , 20(4), 76-87. doi: <a href="https://doi.org/10.21511/ppm.20(4).2022.07">10.21511/ppm.20(4).2022.07</a>
<b>DOI</b>	<a href="http://dx.doi.org/10.21511/ppm.20(4).2022.07">http://dx.doi.org/10.21511/ppm.20(4).2022.07</a>
<b>RELEASED ON</b>	Thursday, 20 October 2022
<b>RECEIVED ON</b>	Wednesday, 15 June 2022
<b>ACCEPTED ON</b>	Tuesday, 11 October 2022
<b>LICENSE</b>	 This work is licensed under a <a href="https://creativecommons.org/licenses/by/4.0/">Creative Commons Attribution 4.0 International License</a>
<b>JOURNAL</b>	"Problems and Perspectives in Management"
<b>ISSN PRINT</b>	1727-7051
<b>ISSN ONLINE</b>	1810-5467
<b>PUBLISHER</b>	LLC “Consulting Publishing Company “Business Perspectives”
<b>FOUNDER</b>	LLC “Consulting Publishing Company “Business Perspectives”



NUMBER OF REFERENCES

**34**



NUMBER OF FIGURES

**3**



NUMBER OF TABLES

**1**

© The author(s) 2023. This publication is an open access article.



## BUSINESS PERSPECTIVES



LLC "CPC "Business Perspectives"  
Hryhorii Skovoroda lane, 10,  
Sumy, 40022, Ukraine  
[www.businessperspectives.org](http://www.businessperspectives.org)

**Received on:** 15<sup>th</sup> of June, 2022  
**Accepted on:** 11<sup>th</sup> of October, 2022  
**Published on:** 20<sup>th</sup> of October, 2022

© Ivan Digel, Dinara Mussabalina,  
Marat Urdabayev, Nurbakhyt  
Nurmukhametov, Aigul Akparova,  
2022

Ivan Digel, Master of Economics,  
Department of Economic Theory and  
Institutional Research, Institute of  
Economics of the Science Committee  
of the Ministry of Education and  
Science of the Republic of Kazakhstan,  
Kazakhstan.

Dinara Mussabalina, Ph.D. Student,  
Department of Economic, L. N.  
Gumilyov Eurasian National University,  
Kazakhstan.

Marat Urdabayev, Ph.D. Student,  
Department of Economic, Al-  
Farabi Kazakh National University,  
Kazakhstan. (Corresponding author)

Nurbakhyt Nurmukhametov, Acting  
Professor, Department of Economic,  
Seifullin Kazakh Agrotechnical  
University, Kazakhstan.

Aigul Akparova, Ph.D. in Economics,  
Department of Economic, Kazakh  
Ablai Khan University of International  
Relations and World Languages,  
Kazakhstan.



This is an Open Access article,  
distributed under the terms of the  
[Creative Commons Attribution 4.0  
International license](https://creativecommons.org/licenses/by/4.0/), which permits  
unrestricted re-use, distribution, and  
reproduction in any medium, provided  
the original work is properly cited.

**Conflict of interest statement:**  
Author(s) reported no conflict of interest

Ivan Digel (Kazakhstan), Dinara Mussabalina (Kazakhstan), Marat Urdabayev (Kazakhstan),  
Nurbakhyt Nurmukhametov (Kazakhstan), Aigul Akparova (Kazakhstan)

# EVALUATING DEVELOPMENT PROSPECTS OF SMART CITIES: CLUSTER ANALYSIS OF KAZAKHSTAN'S REGIONS

## Abstract

This study aims to study Kazakhstan's regions and identify places with the best potential for developing smart cities based on cluster analysis. To analyze the differentiation by the level of development, 17 regions of Kazakhstan are grouped according to 2020 data from the statistical bulletin of the National Bureau of Statistics of the Republic of Kazakhstan. The formation of groups of regions with different values of indicators was carried out based on agglomerative clustering using the single linkage, complete linkage, and Ward's clustering methods. In agglomerative clustering, the algorithm groups regions based on observations into clusters, and indicators determine each area's innovative development level. The instrument to build clustering is the "RStudio" software package. As a result, regions with their essential characteristics were identified, and an assessment of their prospects was obtained with the most significant potential for developing and managing "smart cities" – Atyrau region, Almaty city, and Astana city. The remaining clusters include regions where favorable conditions for the development of innovations have not yet been formed, which require more resources and efforts to build "smart cities." Therefore, they should not be the first to implement this concept. They need a more balanced, integrated approach, ideally supported by experience in implementing the idea in more promising regions. In a sense, clustering also allowed for identifying potential (or even existing) innovation clusters in regions of Kazakhstan. The study results can be used in developing government programs to form smart cities and further study the potential of smart cities.

## Keywords

innovative industries, regional development, digital economy, smart cities, smart management

## JEL Classification

R11, O33, R58

## INTRODUCTION

Nowadays, the critical task is to create conditions for the development of cities of all types, ensuring the growth of their competitiveness and sustainable development of the country's territories. Moreover, the decisive role here is played by interactions and mutual assistance based on the most effective use of limited resources, primarily intellectual. This actualizes the task of creating conditions for the development of modern cities as centers that ensure the development of information parameters of urban development through the formation of smart cities.

Therefore, there is a process of rethinking city management worldwide, and an increasing number of cities are moving to the concept of innovative city development. Developing countries that seek to increase their intellectual potential and ensure sustainable growth of territories are no exception. Smart cities allow to improve the life quality of citizens, reduce socio-economic inequalities, and make city management more effective. The COVID-19 pandemic has reinforced these changes and increased their relevance.

The main goal of a smart city is to improve the resident's quality of life with the help of innovative technologies. Such technologies make urban space more effective in meeting the population's needs and implementing the most modern urban upgrading forms to make citizens' lives comfortable and safe. Developing smart cities requires specific conditions that arise unevenly in different regions of the country. The further development of smart cities is highly dependent on innovation activity because they naturally depend on the introduction of new technological solutions to existing ones. Such "cores" have unique characteristics that can be distinguished using clustering.

Kazakhstan, as a developing country with many regions, is just beginning its transition to the digitalization of the city, introducing local solutions. However, some cities and regions of Kazakhstan already have an excellent infrastructure to implement the concept of a "smart" city. Thus, Kazakhstan needs to keep pace with global changes, such as the worldwide informatization of society, urbanization, and the increasing importance of cities.

Cluster analysis has clear advantages over other types of research. This approach makes it possible to identify promising areas for further sustainable development accurately. It should also be noted that it allows determining the most successful place for the location of a smart city, the type of object required by the name here, and the specific area of the object within the cluster, which is the most economically advantageous. So, in this study, cluster analysis will be used, which should become the basis for determining the level of development of regions and the grouping of cities. It gives an objective picture and predicts the direction of development of smart cities. This is especially important for the zoning of urban areas of Kazakhstan, which will allow rational use of their potential.

Many developing countries, Kazakhstan among them, are experiencing rapid urbanization. The proportion of urban residents and their demands for quality of life are becoming more sophisticated. The development of smart cities is one of the modern ways to meet these demands. Exploring the potential for developing and managing smart cities in Kazakhstan has begun very recently. Therefore, this research topic is relevant and requires a more detailed study.

---

## 1. LITERATURE REVIEW

All studies on the topic agree that smart cities improve the efficiency of placement and use of resources and enhance the population's quality of life by improving security, boosting processes, and providing new services. Existing research on smart cities contains a wide variety of topics. All of them, however, can be divided into several areas. Some studies are solving issues related to the application of specific technologies in smart cities. They are completely out of the scope of this study. Others develop an understanding of the term "smart city" and classify existing cities accordingly. These will be covered briefly. The last group, which is of the most significant interest within the framework of this paper, attempts to identify and use factors for the implementation, development, and management of smart cities.

The history of smart cities began in 1974 when the first Big Data project was launched in Los Angeles.

Further, Gibson et al. (1992) represented information, ideas, and infrastructure that accelerate the creation of smart cities, fast systems, and global networks. Mahizhnan (1999) investigated the case of Singapore being a "smart city." Finally, Hall et al. (2000) proposed a "vision" of a smart city, one of the first to suggest its more or less integrated description. However, there is no consent in academia about the nature and features of smart cities, and one has to stick to one preferred among others or is forced to reside in eclecticism.

Later research provides more profound insights into the formation and development of smart cities. Hollands (2008) explored smart cities by identifying essential issues: assumptions about the smart city as a celebratory tag, that the tag is more of a marketing hype than a practical driver of infrastructure change, and the term itself carrying an uncritical, developmental connotation. Leydesdorff and Deakin (2011) emphasized that smart cities are a process of cultur-

al reconstruction underpinned by policy, academic leadership, and corporate strategy in their leadership. It was also noted that over the past two decades, megacities worldwide had been involved in initiatives to improve urban infrastructure and services aimed at improving the environment, and social and economic conditions, increasing the attractiveness and competitiveness of cities (Jong et al., 2015).

Despite the discussion on various concepts and theories, there is no consensus on a clear definition of the term “smart city” (Hortz, 2016). In general, smart cities extensively use information and communication technologies to help large cities build their competitive advantage (Angelidou, 2014). In the case of Doha, smart city practices are more of an interaction between urban technologies and knowledge economy activities (Conventz et al., 2015). At the same time, Pancholi et al. (2015) noted the practice of integrating Brisbane’s smart technologies into good urban and spatial design practices.

A similar methodological approach to assessing smart cities using cluster analysis has been tested by Cantuarias-Villessuzanne et al. (2021). They analyzed the smart strategies of European cities and developed a clustering of smart cities based on the activities implemented by the cities. Lytras et al. (2019) researched the clustering of smart city services. Finally, Xiang et al. (2019) explored streaming data processing applications in a smart city using a cluster analysis approach.

Muntean (2019) used cluster analysis to predict and solve the parking occupancy problem in Birmingham’s smart city. This approach first groups the dataset to get the appropriate periods throughout the day and then predicts the data in those clusters. Safitri et al. (2020) conducted a cluster analysis of the smart city regions of Banda Aceh, Indonesia. Within the framework of this study, the similarities of the characteristics of each object in the regions of Aceh were determined. In addition, Srinivas and Hosahalli (2021) researched the MapReduce distributed computing environment based on clustering using K-means evolutionary computing for a smart city based on the Internet of Things.

The use of cluster analysis in this paper evokes associations with innovative clusters since the concept of a smart city relies heavily on introducing innovations.

Brakman and van Marrewijk (2013) believe that the effect of clusters in cities lies in their impact on cities’ national and regional economies since their effect favorably affects the existing mechanism in agglomerations. Van Klink and de Langen (2001) believe that a cluster of cities must go through a mandatory development life cycle of creating innovative products. The cluster of cities provides for the emergence in the form of small companies that are developing innovative products in smart cities. In addition, the cluster is growing with the arrival of new companies and highly qualified specialists who carry out innovative “smart” projects. The specialization of a cluster increases until it reaches maturity before a specific period. If companies in a cluster do not have time to adapt to changing economic conditions, then decline sets in.

According to Köcker and Müller (2015), the main objectives of the cluster policy are increasing labor productivity, the speed of development of innovative products, and the competitiveness of small and medium-sized enterprises in the region. Noiva et al. (2016) analyzed a dataset of 142 cities, including annual per capita water use. With these urban water supply and consumption measures, they conducted a hierarchical cluster analysis to identify relative similarities and distances between 142 cases. Kubina et al. (2021) compared standards, implementation, and cluster models for smart cities in North America and Europe using cluster analysis. Finally, Héraud and Muller (2022) studied the interaction between smart cities and innovation clusters, as well as people involved in technology clusters, research centers, factory labs, living labs, etc.

Nazarova and Demianenko (2018) conducted a cluster analysis of the regions of Ukraine. According to the cluster analysis results, the regions are grouped into six clusters. The dynamics of the quantitative distribution of the regions of Ukraine according to the selected clusters were also analyzed. The study identified cores with a constant composition of regions and presented the characteristics of each cluster.

The analysis of Kazakhstan’s “smart cities” and cluster analysis of the regions and cities were carried out also in Kazakh studies. Urdabayev and Turgel (2021) evaluated the applicability of the “Smart Aqkol” case in the development of “smart cities” in other cities in Kazakhstan. In addition, there is the

“WeAlmaty” project implemented by the British Council, the city government of Almaty, JSC “City Development Center of Almaty,” and Kazakhstan-British Technical University.

Aralbaeva and Berikbolova (2021) considered a cluster analysis of Kazakhstan’s regions regarding the level of innovative development. In their opinion, one of the effective methods to manage the possibilities of sustainable development of cities is the cluster policy of Kazakhstan, with the help of which interconnected forms of suppliers and universities are formed (Kulanov et al., 2020). Cluster policy is an effective form of relations in the internal environment of the regions of the Republic of Kazakhstan, which has recently been a dominant component. The main task of cluster policy is to create favorable conditions for the development of the regional economy, depending on the category of the cluster and the strategy of the regions (Aralbaeva & Berikbolova, 2021).

Satpayeva et al. (2020) proposed methodological tools based on a systematic approach using economic and statistical methods and the 5Ms concept. Furthermore, Mussabalina and Kireyeva (2019) believe that the topic of cluster development in Kazakhstan deserves special attention. Therefore, attempts have been made to maintain and develop a cluster policy aimed at the socio-economic development of Kazakhstan and its regions.

Based on the literature review, it can be concluded that there are studies devoted to the problems of the formation and management of the development of smart cities. Furthermore, different studies are dedicated to the use of cluster analysis to determine the level of development of regions and the grouping of cities. Thus, some studies are related to analyzing “smart cities” and existing clusters in Kazakhstan, as well as articles in which cluster analysis is used as a research method on similar topics. Nevertheless, very few studies use cluster analysis to study the regional environment to form a smart city. Moreover, until now, no research used cluster analysis to zone the potential of the regions of Kazakhstan for developing “smart cities.”

The cluster policy for developing and managing smart city clusters makes it possible to increase innovation activity by strengthening small and medi-

um-sized enterprises, focusing on a common strategic goal and innovation activity.

Therefore, this paper aims to study Kazakhstan’s regions and identify places with the best potential for developing “smart cities” based on cluster analysis.

## 2. METHODOLOGY

This study employs agglomerative clustering with the “RStudio” software using single hierarchical linkage (“closest neighbor”), complete linkage (“farthest neighbor”), and Ward’s clustering methods.

In agglomerative clustering, the algorithm groups up observations in clusters, starting with as many clusters as there are observations and merging them one by one until only one cluster contains all observations. The closer the clusters are to each other, the sooner they merge. Thus, the general algorithm for applying the methods looks as follows. First, suitable quantitative indicators are selected, then the “distances” between all observations are calculated. Next, the observations are combined into clusters according to a specific criterion. It is the researcher’s job to identify a set of clusters that can be meaningfully and usefully interpreted. The method does not imply a clear procedure for choosing the number of clusters.

Ward (1963) first described this method. The function  $D(X, Y)$  calculating the between-clusters distance measures the increase in the “error sum of squares” (ESS) after combining two clusters.

$$D(X, Y) = ESS(XY) - [ESS(X) + ESS(Y)], \quad (1)$$

where  $ESS(.)$  is of the form:

$$ESS(X) = \sum_{i=1}^{N_x} \left| x_i - \frac{1}{N_x} \sum_{j=1}^{N_x} x_j \right|^2, \quad (2)$$

where  $N_x$  is the number of elements in the cluster,  $x_i$  and  $x_j$  are the elements of the cluster. The goal of this method is to choose such a sequence of clustering steps that minimizes  $D(X, Y)$  (the increase in ESS at each step).

The single and the complete linkage methods work differently. They still start from as many clusters

as there are observations, ending with only one, but the merging criteria differ from Ward’s method. Both methods use Euclidean distance:

$$D_{ij} = \sqrt{\sum_i (a_i - b_i)^2}, \quad (3)$$

$$D_{ij} = \sqrt{(P_{1x} - P_{2x})^2 + (P_{1y} - P_{2y})^2}, \quad (4)$$

where  $a_i$  and  $b_i$  are the chosen variables of the respective observations A and B.

At each step, two clusters with the smallest distance are merged. The methods differ in how the new distance is calculated. For the single linkage, the distance is calculated between the two closest points of two clusters (i.e., the closest neighbors), for the complete linkage, the distance is calculated between the two farthest points of the two given clusters (i.e., the farthest neighbors) (Sharma & Neha, 2019).

Before clustering, the data were standardized using z-score to allow comparison of the indicators with different units of measurement. The data source is the National Statistics Bureau of the Agency for Strategic Planning and Reforms of the Republic of Kazakhstan. The objects to the cluster are 17 administrative units of the highest order of Kazakhstan, of which 14 are “regions” and 3 are “cities of republican significance.”

The selected set of indicators determines the level of innovative development of regions. Since the development of “smart cities” is closely related to digital technologies, it is possible to use the same

indicators as an approximation. They are of two types: indicators of the region’s innovation responsiveness (gross regional product, labor productivity, employment, fixed assets, returns on assets and ecological properties of production) and measures of innovation activity in the region (spending on R&D, the share of firms using innovations, production of innovative goods).

### 3. RESULTS

As a result of the clustering, there are three dendrograms (Figures 1, 2, and 3), one for each method are used. The visual analysis was the primary decision-making tool to cut dendrograms in this study, supplemented with considerations about the possible interpretation of the clusters. The red frames represent the selected clusters. As a unit of measurement in the pictures, the scale uses the Euclidean distance represented by standard deviations.

#### 3.1. Single linkage

Figure 1 shows a dendrogram characteristic of the single linkage method. Most regions are connected into one large cluster, and the remaining clusters account for a small fraction of the remaining observations. Here the dendrogram is easily cut into three clear clusters:

1. Astana city and Almaty city.
2. Atyrau region.
3. All other regions.

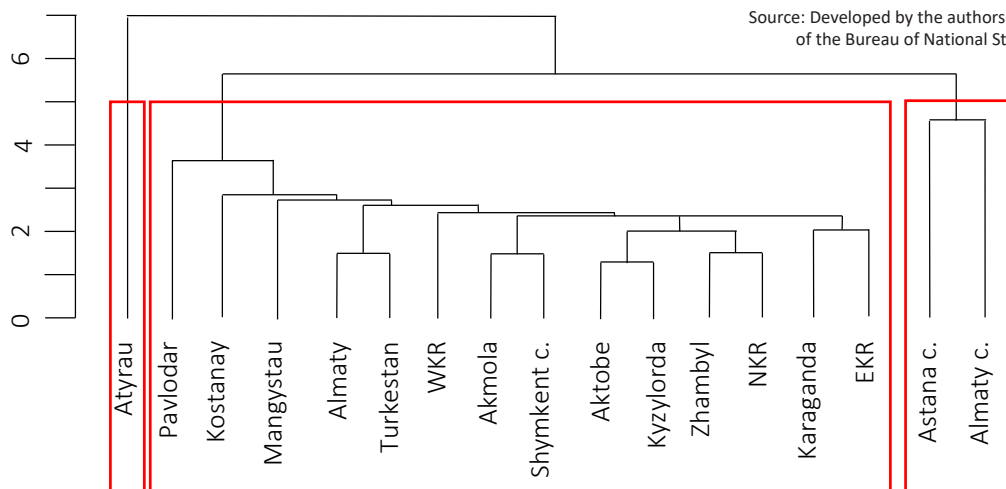
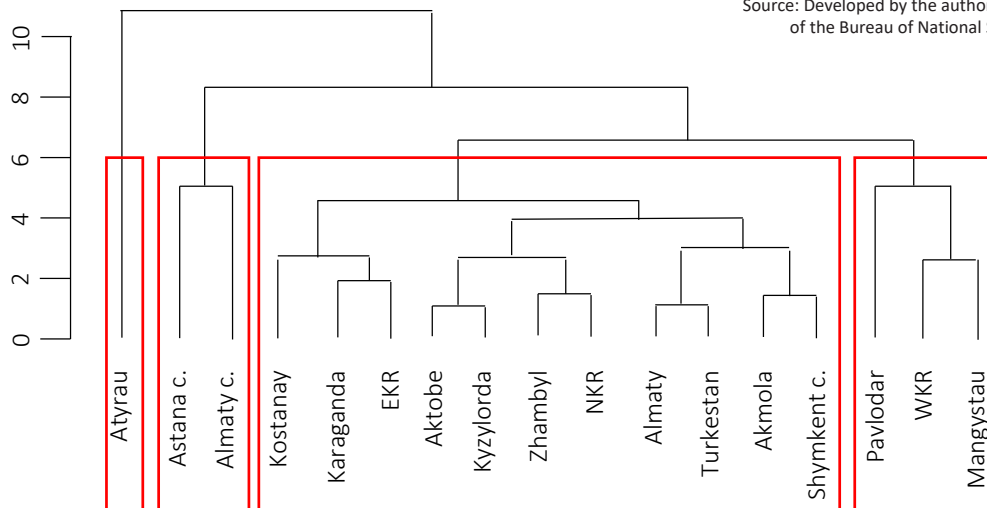


Figure 1. Dendrogram of clusters of regions of Kazakhstan, single linkage

Source: Developed by the authors using the data of the Bureau of National Statistics (2021).



**Figure 2.** Dendrogram of clusters of regions of Kazakhstan, complete linkage

### 3.2. Complete linkage

The dendrogram could be separated into the same clusters as for the single linkage. However, there is an opportunity to create the fourth cluster. Here are all four as follows:

1. Astana city and Almaty city.
2. Atyrau region.
3. Pavlodar region, Mangystau region, West Kazakhstan region.
4. All other regions.

ly used approaches and give reasons to create four clusters as it was for complete linkage. However, the fourth cluster here is different from the one of the complete linkage method:

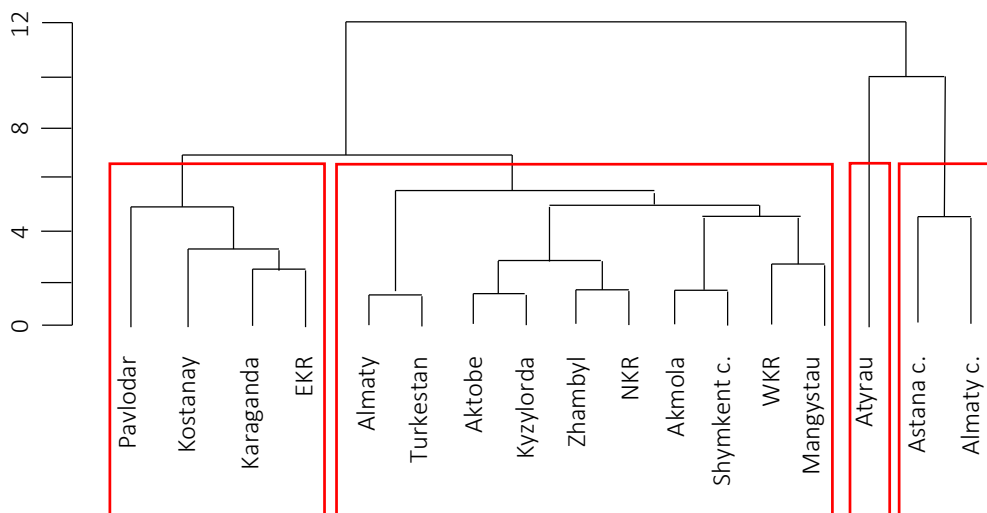
1. Astana city and Almaty city.
2. Atyrau region.
3. Pavlodar region, Kostanay region, Karaganda region, East Kazakhstan region.
4. All other regions.

### 3.3. Ward's method

The results of Ward's method allow this study to create three clusters similar to the two previous-

For ease of interpretation, all the results were gathered in Table 1. In each cluster, indicator averages were calculated and distributed by minimum and maximum criteria among clusters.

Source: Developed by the authors using the data of the Bureau of National Statistics (2021).



**Figure 3.** Dendrogram of clusters of regions of Kazakhstan, Ward's method

**Table 1.** Description and interpretation of the regional clusters of Kazakhstan

Source: Authors' elaboration.

No.	Regions	Maximum	Minimum
<b>Three clusters (the same for all three used methods)</b>			
1	<b>"Long-Term Growth Zone"</b> Pavlodar, Kostanay, Karaganda, East Kazakhstan region, Almaty, Aktobe, Kyzylorda, Zhambyl, NKR, Akmola, Shymkent city, Mangystau, North Kazakhstan Region, Turkestan	<ul style="list-style-type: none"> <li>Return on fixed assets</li> </ul>	<ul style="list-style-type: none"> <li>Labor productivity in the sectors "Professional and scientific or technology activity", "Manufacturing", "Construction", "Transportation and storage", "Information and communications"</li> <li>Fixed assets</li> <li>Spending on R&amp;D</li> <li>GRP</li> </ul>
2	<b>"Innovative Hub of the West"</b> Atyrau	<ul style="list-style-type: none"> <li>Labor productivity in the sectors "Professional and scientific or technology activity", "Manufacturing", "Construction", "Transportation and storage"</li> <li>Fixed assets</li> <li>Production of innovative goods</li> <li>Emissions to air from stationary sources</li> </ul>	<ul style="list-style-type: none"> <li>Employed population</li> <li>Return on fixed assets</li> <li>Innovation activity</li> <li>Ratio of GRP to emissions</li> </ul>
3	<b>"Cores of Smart Cities"</b> Astana c., Almaty c.	<ul style="list-style-type: none"> <li>Labor productivity in the sector "Information and communications"</li> <li>Employed population</li> <li>Ratio of GRP to emissions</li> <li>Spending on R&amp;D</li> <li>Innovation activity</li> <li>GRP</li> </ul>	<ul style="list-style-type: none"> <li>Production of innovative goods</li> <li>Emissions to air from stationary sources</li> </ul>
<b>Four clusters (Complete linkage)</b>			
1	<b>"Obsolete Industries Zone"</b> Pavlodar, West Kazakhstan Region, Mangystau	<ul style="list-style-type: none"> <li>Emissions to air from stationary sources</li> </ul>	<ul style="list-style-type: none"> <li>Labor productivity in the sector "Manufacturing"</li> <li>Ratio of GRP to emissions</li> <li>Innovation activity</li> <li>Production of innovative goods</li> </ul>
2	<b>"Long-Term Growth Zone"</b> Kostanay, Karaganda, East Kazakhstan region, Almaty, Aktobe, Kyzylorda, Zhambyl, North Kazakhstan region, Akmola, Shymkent c., Turkestan	<ul style="list-style-type: none"> <li>Return on fixed assets</li> </ul>	<ul style="list-style-type: none"> <li>Labor productivity in the sectors "Information and communications", "Professional and scientific or technology activity", "Construction", "Transport and storage"</li> <li>Fixed assets</li> <li>Spending on R&amp;D</li> <li>GRP</li> </ul>
3	<b>"Innovative Hub of the West"</b> Atyrau	<ul style="list-style-type: none"> <li>Labor productivity in the sectors "Professional and scientific or technology activity", "Processed industry", "Transport and storage", "Construction"</li> <li>Fixed assets</li> <li>Production of innovative goods</li> </ul>	<ul style="list-style-type: none"> <li>Employed population</li> <li>Return on fixed assets</li> </ul>
4	<b>"Cores of Smart Cities"</b> Astana c., Almaty c.	<ul style="list-style-type: none"> <li>Labor productivity in the "Information and Communication" industry</li> <li>Employed population</li> <li>Ratio of GRP to emissions</li> <li>Spending on R&amp;D</li> <li>Innovation activity</li> <li>GRP</li> </ul>	<ul style="list-style-type: none"> <li>Emissions to air from stationary sources</li> </ul>
<b>Four clusters (Ward's method)</b>			
1	<b>"Traditional Industrial Zone"</b> Pavlodar, Kostanay, Karaganda, East Kazakhstan region	<ul style="list-style-type: none"> <li>Return on fixed assets</li> <li>Emissions to air from stationary sources</li> </ul>	<ul style="list-style-type: none"> <li>Ratio of GRP to emissions</li> </ul>
2	<b>"Long-Term Growth Zone"</b> Almaty, Aktobe, Kyzylorda, Zhambyl, NKR, Akmola, Shymkent c., Mangystau, North Kazakhstan Region, Turkestan	<ul style="list-style-type: none"> <li>No indicators</li> </ul>	<ul style="list-style-type: none"> <li>Labor productivity in the "Information and Communication", "Professional and scientific or technology activity", "Manufacturing", "Construction", "Transport and storage"</li> <li>Fixed assets</li> <li>Innovation activity</li> <li>Production of innovative goods</li> <li>GRP</li> <li>Spending on R&amp;D</li> </ul>



**Table 1 (cont.).** Description and interpretation of the regional clusters of Kazakhstan

No.	Regions	Maximum	Minimum
3	<b>"Innovative Hub of the West"</b> Atyrau	<ul style="list-style-type: none"> <li>Labor productivity in the sectors "Professional and scientific or technology activity", "Processed industry", "Transport and storage", "Construction"</li> <li>Fixed assets</li> <li>Production of innovative goods</li> </ul>	<ul style="list-style-type: none"> <li>Employed population</li> <li>Return on fixed assets</li> </ul>
4	<b>"Cores of Smart Cities"</b> Astana c., Almaty c.	<ul style="list-style-type: none"> <li>Productivity in the "Information and Communication" industry</li> <li>Employed population</li> <li>Ratio of GRP to emissions</li> <li>Spending on R&amp;D</li> <li>Innovation activity</li> <li>GRP</li> </ul>	<ul style="list-style-type: none"> <li>Emissions to air from stationary sources</li> </ul>

The main result of the analysis is three clusters. The choice of method does not affect the composition of these clusters in any way, but it does affect the aggregation sequence.

"Long-term growth zone" is the largest of the clusters. It has the highest average return on fixed assets but the lowest average labor productivity, fixed assets, GRP, and R&D spending. It represents regions with almost no advantages for smart cities and needs to develop some before implementing them. It is easy to see that both strengths and weaknesses of the cluster are of economic nature, whereas its ecological and innovative indicators are somewhere in-between. The highest average return on fixed assets means this cluster is somewhat promising for future investors. Still, it will be necessary to carefully manage policy to stimulate investments in R&D, IT, and infrastructure instead of, for example, mining. However, with only one strength and many weaknesses, this cluster is not the best choice to start the development of smart cities in Kazakhstan. The following clusters seem to be more suitable.

"Cores of smart cities" are two cities of republican importance: Almaty and Astana. They have developed industrial IT clusters with the highest average labor productivity. The most significant average spending on R&D and innovation activity would allow for easier development and introduction of "smart" technologies. The highest average GRP and employment indicate significant market capacities and purchasing power, which could support the demand for services in a smart city. On top of that, Astana and Almaty have the most emissions-efficient economies, but the volume of

emissions is still the biggest among clusters. The combination of these factors makes this cluster the most suitable for the pilot projects of smart cities in Kazakhstan.

"Innovative hub of the West" is, in fact, only one region, Atyrau. However, it is so different from all other clusters that its "closest neighbors" are only cities from the cluster "cores of smart cities." However, Atyrau region has its own advantages – the highest labor productivity in the sectors "Professional and scientific or technology activity," "Processed industry," "Transport and storage," and "Construction;" the highest average value of fixed assets and the highest amount of produced innovative goods. Those indicate that there exist firms capable of producing the staff necessary for the development of smart cities. These strengths make the cluster suitable for smart cities as well, but the weaknesses here are more influential than in the previous sector. In particular, this cluster has the lowest average employment and returns of fixed assets.

Applying the complete linkage method or Ward's method leaves "Cores of smart cities" and "Innovative hub of the West" unchanged but picks out certain regions from the "Long-term growth zone":

- The complete linkage method gives an "Obsolete industries zone" cluster. This is a cluster of regions with non-innovative industry, which is also inefficient in terms of environmental pollutant emissions and has low labor productivity. The development of smart cities in these regions will require a complete

overhaul. Significant attention will need to be paid to reducing pollution, stimulating entrepreneurial activity, and increasing labor productivity.

- With Ward's method, it is possible to distinguish a "Traditional industrial zone" cluster. Unlike the "Obsolete industries zone," this one has the advantage of the highest average return on fixed assets, and only a low return on GRP in relation to emissions remains among the disadvantages. This cluster includes regions with efficient enterprises in terms of capital investments, which, nevertheless, still pollute the environment heavily. The efficiency of capital investments can be used to manage a financial basis for developing smart cities.

Summing up, there are definitely three clusters. Changing the approach makes it possible to determine an additional cluster from the "Long-term growth zone" and interpret it meaningfully. The choice of specific clustering depends on management preferences. Regardless of the approach, the most promising for developing smart cities are clusters "Cores of smart cities" and "Innovative hub of the West," which already have conditions for implementing pilot projects. Their cases should further serve as an example to develop smart cities in the remaining cluster(s) regions. It is important to note that the three most promising regions best suited to developing "smart cities" are well-distributed geographically. Those are Astana city for the northern part of the country, Almaty, for the southern one, and Atyrau, for the western region.

## 4. DISCUSSION

Cluster analysis to assess the potential of smart cities in Kazakhstan has not been used before. However, there are similar studies for other countries. For example, Cantuarias-Villesuzanne et al. (2021) used seven PCA ascending hierarchical classifications. They identified three clusters of cities in Europe: cities with new smart strategies, cities focused on technology, and smart cities focused on quality of life. The study focuses on cities with good development potential as smart cities and pays little attention to the rest. This makes

sense on a European scale with many relatively developed cities. At the same time, there have been no similar studies regarding the identification and development of smart cities in Kazakhstan.

The study's results allowed dividing the regions of Kazakhstan into three groups: long-term growth zone, cores of smart cities, and innovative hub of the West. The results are consistent with previous studies that used cluster analysis (Kubina et al., 2021; Héraud & Muller, 2022). Kubina et al. (2021) used the clustering method to analyze North America and Europe regions. Those cities were divided into three clusters: balanced innovators cities, post-industrial opportunists cities, and evolutionary cities. In addition, Héraud and Muller (2022) analyzed European cities and highlighted five municipalities representing the best practices in implementing the smart city concept. This review showed that the selected regions for developing smart cities could receive significant positive effects in the form of increased competitiveness, increased labor productivity, and improved infrastructure.

The fact that the cities of Astana and Almaty would be the most favorable for smart cities is not surprising since these are the two most developed cities in the country, being attractors for capital and migration. Suddenly, the city of Shymkent, which became the third million-plus city in the country, and several years ago received a "promotion" in the form of obtaining the status of a "city of republican significance," does not have such great potential. Despite the new position, its characteristics remain at the level of the least developed regions in the study context. At the same time, the Atyrau region turned out to be so unexpectedly original that it needed to be singled out in a separate cluster, which could not be expected after a superficial study. The presence of a large group of regions with low potential is also not surprising. Still, the opportunity to single out a cluster in a "transitional state" is also unexpected and made it possible to understand the capabilities of these regions better. Obviously, changing the clustering method affects the resulting clusters. However, two clusters are resistant to method changes. "Cores of smart cities" and "Innovative hub of the West" remain the same regardless of the method used. This may indicate the "robustness" of their uniqueness.

Future research may focus on the following areas:

- The most obvious is expanding the list of indicators for clustering. The list of indicators used in this paper is not exhaustive. For example, indicators of the development of political institutions were not included. However, under a unitary autocracy, the assumption of a slight difference in political institutions in the regions is quite reasonable. However, one may try to identify and include any differences in the analysis. Expanding the list of indicators of technological development can also improve the investigation, but the research will face difficulties in collecting data.
- The next option is to change the administrative level. The present analysis is made for the highest level of the administrative system of Kazakhstan, which contains only regions and cities of republican significance. The change to a lower level could give more data points to single out individual cities or parts of regions in which smart city projects can be launched and efficiently managed.
- Using a different approach to clustering is an option, as well. In this paper, hierarchical agglomerative methods are used. At the same time, the cluster analysis toolkit is much more comprehensive. It includes hierarchical divisive methods and a large group of non-hierarchical methods that can give other valuable results.

---

## CONCLUSION

This study aimed to evaluate the features of the regions of Kazakhstan to highlight those having the best potential to develop and manage “smart cities.” The analysis gave either expected and unsuspected results. The study highlighted deep regional inequality affecting the potential to successfully develop and manage smart cities. There are three regions with the highest potential – Atyrau region, Almaty city, and Astana city. They have a lot of capital, high R&D and innovational activity, employment rate, and labor productivity to support “smart cities” development. The remaining regions will demand much more resources and effort to build “smart cities” and thus should not be the first to implement the concept.

This inequality must be considered for the development of smart cities in Kazakhstan. The best solution in forming a policy for the development of smart cities will be the launch of pilot projects in the “cores of smart cities.” If the projects are successful, they can include the “innovative hub of the West” cluster. This paper can be helpful for the development of smart cities, which is essential when developing projects for clustering and sustainable growth.

## AUTHOR CONTRIBUTIONS

Conceptualization: Ivan Digel, Marat Urdabayev, Nurbakhyt Nurmukhametov, Aigul Akparova.

Data curation: Ivan Digel, Dinara Mussabalina, Marat Urdabayev.

Funding acquisition: Dinara Mussabalina.

Investigation: Ivan Digel, Marat Urdabayev, Nurbakhyt Nurmukhametov.

Methodology: Ivan Digel, Marat Urdabayev, Aigul Akparova.

Project administration: Ivan Digel, Dinara Mussabalina.

Resources: Nurbakhyt Nurmukhametov.

Software: Marat Urdabayev, Aigul Akparova.

Supervision: Dinara Mussabalina.

Validation: Aigul Akparova.

Visualization: Ivan Digel, Marat Urdabayev, Nurbakhyt Nurmukhametov, Aigul Akparova.

Writing – original draft: Ivan Digel, Marat Urdabayev.

Writing – review & editing: Ivan Digel, Marat Urdabayev.

## REFERENCES

1. Angelidou, M. (2014). Smart city policies: A spatial approach. *Cities*, 41(1), S3-S11. <https://doi.org/10.1016/J.CITIES.2014.06.007>
2. Aralbaeva, G. G., & Berikbolova, U. D. (2021). Cluster Analysis of the regions of Kazakhstan by the level of innovative development. *International Research Journal*, 9(111), 133-137. <https://doi.org/10.23670/IRJ.2021.9.111.059>
3. Brakman, S., & van Marrewijk, C. (2013). Reflections on cluster policies. *Cambridge Journal of Regions, Economy and Society*, 6, 217-231.
4. Bureau of National Statistics. (2021). *Annual statistical collections of the Republic of Kazakhstan*. (In Kazakh). Retrieved from <https://stat.gov.kz/official/industry>
5. Cantuarias-Villessuzanne, C., Weigel, R., & Blain, J. (2021). Clustering of European Smart Cities to Understand the Cities Sustainability Strategies. *Sustainability*, 13(2), 513. <https://doi.org/10.3390/su13020513>
6. Ceylan, R. F., Ozkan, B., & Mu-lazimogullari, E. (2020). Historical evidence for economic effects of COVID-19. *The European Journal of Health Economics*, 21(6), 817-823. <https://doi.org/10.1007/s10198-020-01206-8>
7. Conventz, S., Thierstein, A., Wiedmann, F., & Salama, A. M. (2015). When the Oryx takes off: Doha a new rising knowledge hub in the Gulf region? *International Journal of Knowledge-Based Development*, 6(1), 65-82. <https://doi.org/10.1504/IJKBD.2015.069443>
8. El Mendili, S., El Bouzekri, Y., El Idrissi, & Hmina, N. (2016). Benchmarking study on smart city data analytics. *4th IEEE International Colloquium on Information Science and Technology (CiSt)* (pp. 841-846). <https://doi.org/10.1109/CIST.2016.7805004>
9. Gibson, D.V., Kozmetsky, G., & Smilor, R.W. (1992). *The Technopolis Phenomenon: Smart Cities, Fast Systems, Global Networks*. Lanham: Rowman & Littlefield Publishers.
10. Hall, R. E., Bowerman, B., Braverman, J., Taylor, J., Todosow, H., & Von Wimmersperg, U. (2000). The vision of a smart city. *2nd International Life Extension Technology Workshop*. Paris, France. Retrieved from <https://digital.library.unt.edu/ark:/67531/metadc717101/>
11. Héraud, J., & Muller, E. (2022). Smart Cities and Innovation Clusters. *Open Journal of Business and Management*, 10(1), 387-401. <https://doi.org/10.4236/ojbm.2022.101023>
12. Hollands, R. G. (2008). Will the real smart city please stand up? *City*, 12(3), 303-320. <https://doi.org/10.1080/13604810802479126>
13. Hortz, T. (2016). The Smart State test: a critical review of the Smart State Strategy 2005-2015's Knowledge-Based Urban Development. *International Journal of Knowledge-based Development*, 7(1), 75-101. <https://doi.org/10.1504/IJKBD.2016.075434>
14. Jong, M., Joss, S., Schraven, D., Zhan, C., & Weijnen, M. (2015). Sustainable-smart-resilient-low carbon-eco-knowledge cities; making sense of a multitude of concepts promoting sustainable urbanization. *Journal of Cleaner Production*, 109, 25-38. <https://doi.org/10.1016/J.JCLEPRO.2015.02.004>
15. Köcker, G. M., & Müller, L. (2015). *Cluster Programmes in Europe* (Report). European Cluster Observatory. Retrieved from <https://ec.europa.eu/docsroom/documents/12925/attachments/1/translations/en/renditions/pdf>
16. Kulanov, A., Issakhova, A., Koshkina, O., Issakhova, P., & Karshalova, A. (2020). Venture financing and the fuel and energy complex: Investing in alternative energy. *International Journal of Energy Economics and Policy*, 10(5), 531-538. <http://dx.doi.org/10.32479/ijeep.9963>
17. Kubina, M., Šulyová, D., & Vodák, J. (2021). Comparison of Smart City Standards, Implementation and Cluster Models of Cities in North America and Europe. *Sustainability*, 13, 3120. <https://doi.org/10.3390/SU13063120>
18. Leydesdorff, L., & Deakin, M. (2011). The Triple-Helix Model of Smart Cities: A Neo-Evolutionary Perspective. *Journal of Urban Technology*, 18(2), 53-63. <https://doi.org/10.1080/10630732.2011.601111>
19. Lytras, D. M., Visvizi, A., & Sarirete, A. (2019). Clustering Smart City Services: Perceptions, Expectations, Responses. *Sustainability*, 11(6), 1669. <https://doi.org/10.3390/su11061669>
20. Mahizhnan, A. (1999). Smart cities: the Singapore case. *Cities*, 16(1), 13-18. [https://doi.org/10.1016/S0264-2751\(98\)00050-X](https://doi.org/10.1016/S0264-2751(98)00050-X)
21. Muntean, M. V. (2019). Car Park Occupancy Rates Forecasting based on Cluster Analysis and kNN in Smart Cities. *11th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)* (pp. 1-4). <https://doi.org/10.1109/ECAI46879.2019.9042098>
22. Mussabalina, D. S., & Kireyeva, A. A. (2019). Assessment of the level of innovative development of the regions of Kazakhstan and the possibility of their further clustering. *Economics: the strategy and practice*, 1(14), 149-161. <https://esp.ieconom.kz/jour/article/view/126/123>
23. Nazarova, G., & Demianenko, A. (2018). Analysis of the human security in Ukraine in a regional perspective. *Social and labour relations: theory and practice*, 8(2), 1-7. [https://doi.org/10.21511/slrrp.8\(2\).2018.01](https://doi.org/10.21511/slrrp.8(2).2018.01)
24. Noiva, K., Fernández, J. E., & Wescoat, J. L. (2016). Cluster analysis of urban water supply and demand: Toward large-scale comparative sustainability planning. *Sustainable Cities and Society*, 27, 484-496. <https://doi.org/10.1016/j.scs.2016.06.003>
25. Pancholi, S., Yigitcanlar, T., & Guaralda, M. (2015). Public space

- design of knowledge and innovation spaces: Learnings from Kelvin grove Urban Village, Brisbane. *Journal of Open Innovation*, 1(1), 1-17. <http://dx.doi.org/10.1186/s40852-015-0015-7>
26. Safitri, W. D., Ikhwan, M., Firman-syah, D., Rusdiana, S., Rahayu L., & Akhdansyah, T. (2020). Partial distance strategy analysis on city characteristics to improve reliable smart cities services. *3rd Smart Cities Symposium (SCS)* (pp. 354-358). <https://doi.org/10.1049/icp.2021.0933>
  27. Satpayeva, Z. T., Kireyeva, A. A., Kenzhegulova, G., & Yermekbayeva, D. (2020). Gender equality and women business of framework 5Ms in Kazakhstan: Analysis and basic directions. *Journal of Asian Finance, Economics and Business*, 7(3), 253-263. <https://doi.org/10.13106/jafeb.2020.vol7.no3.253>
  28. Sharma, Sh., & Neha, B. (2019). Comparative study of single linkage, complete linkage, and ward method of agglomerative clustering. *International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*.
  29. Srinivas, K. G., & Hosahalli, D. (2021). Evolutionary Computing Assisted K-Means Clustering based MapReduce Distributed Computing Environment for IoT-Driven Smart City. *International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)* (pp. 192-200). <https://doi.org/10.1109/ICCCIS-51004.2021.9397217>
  30. Thierstein, A., Conventz, S., Wiedmann, F., & Salama, A. M. (2013). When the Oryx takes off: Doha a new rising knowledge hub in the Gulf-region? *International Journal of Knowledge-Based Development*, 6(1), 65-82. <https://doi.org/10.1504/IJKBD.2015.069443>
  31. Urdabayev, M. T., & Turgel, I. D. (2021). Development of the smart city on the example of Aqkol project: concepts and main trends. *Economics: the strategy and practice*, 16(2), 188-196. <https://doi.org/10.51176/1997-9967-2021-2-188-196>
  32. Van Klink, A., & de Langen, P. (2001). Cycles in industrial clusters: the case of the shipbuilding industry in the Northern Netherlands. *Tijdschrift Voor Economische En Sociale Geografie*, 92(4), 449-463. <http://dx.doi.org/10.1111/1467-9663.00171>
  33. Ward, J. H. Jr. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*, 58(301) 236-244. <https://doi.org/10.1080/01621459.1963.10500845>
  34. Xiang, Z., Jinghua, C., Wei, S., Quan, G., & Tao, W. (2019). Flow data processing paradigm and its application in smart city using a cluster analysis approach. *Cluster Computing*, 22(2), 435-444. <https://doi.org/10.1007/s10586-018-2839-y>