Analysis of China's Regional Economic Vitality and Its Influential Factors Based on System Cluster Model and Computer Multiple Regression Model

Jinfeng Pang\textsuperscript{1}, Jintao He\textsuperscript{1}, Lan Luo\textsuperscript{2}, Yao Li\textsuperscript{1}, Yiren Mao\textsuperscript{2}, and Hanrui Zhang\textsuperscript{2}

\textsuperscript{1} College of Internet of Things Engineering, Hohai University, Changzhou, 213022, China  
\textsuperscript{2} School of Business Administration, Hohai University, Changzhou, 213022, China

*Correspondence should be addressed to Jinfeng Pang; 1065055828@qq.com

Abstract. The economic vitality of China’s regions (provinces or cities) is affected by many factors. In this paper, the influencing indicators are listed by referring to the relevant factors that affect economic vitality. Based on the data obtained from the Internet, a systematic (hierarchical) clustering model of relevant indicators in Chinese provinces and cities is established, introducing the Elbow Rule and Factor Analysis to analyze and summarize the correlation of various indicators. A computer-based multiple regression model is used to establish a relationship model that affects economic vitality. Based on this relationship model, effective and reasonable suggestions for improving economic vitality are proposed.

1. Introduction
The economic vitality of a region (or city or province) is an important part of the overall regional competitiveness. In recent years, in order to improve economic vitality, some regions have introduced a number of preferential policies to stimulate economic vitality, such as reducing the steps of investment approval, providing funding support for start-ups, and lowering the threshold for attracting talents. However, due to different resource endowments, these policies have different roles in different regions. How to grasp the key factors and effectively improve the vitality of the regional economy is a subject worthy of study.

2. Model assumption
To simplify the problems and make it convenient for us to simulate real-life conditions, we make the following basic assumptions, each of which is properly justified. 

- Assuming that the data source is reliable and accurate.
- Assuming that abnormal data is ignored.
- Assuming that factors with minimal economic impact are ignored.
- Assuming that the employment demand of a city is within the normal variable range.

3. Model establishment and solution: systematic (hierarchical) clustering model
After consulting relevant data and obtained national data, this paper proposes the relevant factors that affect economic vitality: the number of companies living, foreign investment, consumption level index,
permanent population, foreign trade investment capital, productivity level (technical level), ecological environment, social welfare enterprises, energy production, economic aggregates and their growth (regional GDP, per capita GDP, GDP growth rate).

![Figure 1. Allocation of registered capital data for second-tier cities and business units in Beijing, Shanghai, Guangzhou and Shenzhen.](image)

From the perspective of the six pie charts, whether it is a first-tier city, a second-tier city, or even the whole country, the proportion of each capital segment of the registered capital of an enterprise unit is not much different, indicating that the differences in registered capital between cities cannot reflect clearly Because of the difference in economic vitality between cities, this factor is not included in the main influencing factor indicators.

Next, the reduction factor and incremental processing of the impact factor data of the provinces nationwide are obtained here to make the data unified. Then, according to the Pauta Criterion (the $3 \sigma$ criterion), outliers are eliminated from the obtained data.

Then, according to the method proposed in the problem analysis, the system (hierarchical) clustering model based on SPSS is established to analyze and solve the problem. The steps of the system clustering model are as follows:

- Step 1: treat each object as a class and calculate the minimum distance between the two;
- Step 2: merge the two classes with the smallest distance into a new class;
- Step 3: recalculate the distance between the new class and all classes;
- Step 4: repeat steps 2, 3, and 3 until all classes are merged into one;
- Step 5: finish.

Converted to flowchart:
After that, based on the data, the provinces and cities in the country were plotted on the number of enterprises surviving, foreign investment, consumption level index, permanent population (number of population), foreign investment capital, productivity level (technology level), ecological environment, social welfare enterprises, energy production. A systematic clustering pedigree of the ten factors of economic aggregate and its growth (regional GDP, per capita GDP, GDP growth rate), as shown below:

![Flowchart](image1)

**Figure 2. Clustering flowchart**

From the pedigree chart, it can be clearly seen that it has clustered the data of each region according to a certain rule. The value of k can be 2, 3, 4 ..., then determining an appropriate value of k is the solution to this model. The problem also affects analysis and evaluation to a certain extent.

![Pedigree Diagram](image2)

**Figure 3. Pedigree diagram**
Therefore, this article uses SPSS software to draw elbow rule curves for ten impact factors. Elbow Role, that is, when a certain set of data reaches a certain critical point, the degree of distortion will be greatly improved, and then the discount slowly decreases. Then this critical point can be considered as a point with better clustering performance. This The value of the points can be used as the k value of the cluster. The elbow rule curve is as follows:

![Figure 4. Curve of elbow rule](image)

By observing the elbow rule diagram, it is found that when the k value is 4, the clustering coefficient gradually changes from steep to gentle. Therefore, this paper considers that the clustering category is 4 which is more reasonable. That is, the four major influencing factors that affect economic vitality are: demographic factors (including consumption level index, number of resident population), corporate factors (including number of surviving enterprises, productivity level (technical level), foreign investment, foreign trade investment capital, social welfare enterprises ), environmental factors (including ecological environment, energy production) and economic benefits (including economic aggregates and their growth (regional GDP, per capita GDP, GDP growth rate)). The framework can be summarized as a mind map:

![Figure 5. Clustering graph](image)

Then, through the data perspective chart processing of Excel software, this article draws the area stacking chart of these four factors influencing economic vitality. The picture is as follows:
Figure 6. Area map of four major influencing factors

It is obvious from the figure that enterprise factors play a dominant role.

Factor Analysis

This paper introduces a method of scoring the economic vitality of each region based on four factors. Factor scores are calculated by using the component score coefficient matrix times the standardized index, that is:

\[ S_i = \alpha_{i1} x_1 + \alpha_{i2} x_2 + \alpha_{ip} x_p \ (p = 1, 2 \ldots m) \]

Where: \( S_i \) represents the score of the ith factor; \( x_1, x_2 \ldots x_p \) represents the standardized value of the index; \( \alpha_{i1}, \alpha_{i2} \ldots \alpha_{ip} \) represents the component score coefficient. The total factor score is equal to the weighted arithmetic mean of each factor score, namely:

\[ S = \sum b_i S_i \left( \sum b_i = 1 \right) \]

Where: \( S \) represents the total factor score; \( S_i \) represents the score of the ith factor; \( b_i \) represents the contribution degree of the ith factor (contribution degree of fractional factor = contribution rate of variance after rotation of fractional factor/interpretation rate of total variance).

4. Regression model

Based on this model, based on the four major influencing factors derived from the system clustering model, this paper conducts regression analysis on the four types of factors, analyzes the relationship between the influencing factors and economic vitality, and evaluates the population and enterprise factors. In this paper, the statistical analysis software Stata is used to perform regression processing on the four major factors to obtain the regression coefficient table of the four major factors, as follows:
Figure 7. Table of regression coefficients

| Regression Table | Ecov  |
|------------------|-------|
| Businessgrowth   | 0.915*** |
| Population       | -0.087 |
| GDP              | 0.499** |
| Urbangreenspacearea | -0.330* |
| N                | 10    |

*** p<0.01  ** p<0.05  * p<0.1

From the regression coefficient table of the four influencing factors, it can be seen that Businessgrowth has the best significance, while Population has the worst. GDP and Urbangreenspacearea have the best significance at the confidence interval of 0.05 and 0.1 respectively. May safely draw the conclusion from this, the greatest influence on the regional economic vitality of the factors for the enterprise, match with accumulation area on the map, also suggests that we are in a position to accept a null hypothesis. While demographic factors and runs counter to the results of the present picture, and significant effect is poorer, so we reject the null hypothesis, and can be thought of the relatively small population impact on regional economic vitality.

5. Promotion strategies and development recommendations

5.1. Vigorously develop the tertiary industry
Under the environment that China advocates for green development, the government should vigorously develop high-tech industries and energy-saving and environmental protection industries. On the one hand, it protects the environment to achieve the effect of energy conservation and emission reduction, on the other hand, it can quickly expand employment fields and jobs and avoid labor surplus, Increase the income of residents, thereby stimulating the increase of regional economic vitality. The tertiary industry is an important sign of a country's economic development. It has the characteristics of less investment, short cycle, quick results, and higher wages for employees. Actively developing the tertiary industry can not only strengthen the overall quality culture in the region, but also promote the transformation of agricultural society to industrial society.

5.2. Strengthen investment
Investment is a component of economic vitality. The increase in investment intensity is of great significance to the promotion of economic vitality. Investment is divided into government investment, private investment and foreign investment. Government investment is mainly to promote the development of the public domain, strengthen the construction of regional public infrastructure, and allow residents to enjoy more social benefits, which can attract more people and indirectly promote the improvement of economic vitality. In the context of China's market economy, private investment and foreign investment can stimulate the emergence of high-tech industries and emerging industries, thereby enhancing economic growth. Therefore, while rationally planning government investment, private investment and foreign investment should be actively encouraged.

5.3. Improve the level of social security
Optimize and secure regional fiscal revenues through reasonable economic policies to increase the level of social security to enhance regional economic vitality. To a certain extent, a higher level of social security will stimulate residents' consumption, thereby expanding demand, increasing output, and promoting the income of employed people, thereby improving the living standards of residents, enhancing people's happiness, and promoting sustainable regional economic development.
5.4. Improve business economic benefits
The number of enterprises is an important indicator of the economic vitality of a region. If companies want to survive in a highly competitive society and avoid being eliminated, they can only continuously produce products and services that meet social needs. Therefore, enterprises must introduce advanced technology, vigorously develop high-tech industries, and use technological innovation to promote production, so as to achieve the effect of advancing with the times and maintain the continuous improvement of corporate economic vitality.

6. Conclusion
Based on the analysis of China's regional economic vitality and its influencing factors, this paper, through the establishment and solution of the model, the application of system (hierarchical) clustering model and computer multiple regression model, combined with Elbow Rule and Factor Analysis, concludes that enterprise factors are the most influential factors on the regional economy. At the same time, this article puts forward the promotion strategy and the development suggestion, the key aspect includes the industry economy, the social security level and the enterprise.

References
[1] Yi Wei, Construction of Sichuan Province Regional Economic Vitality Evaluation Index System [J], Journal of Sichuan Vocational and Technical College, 2015, 25 (1): 23-25.
[2] Yi Wei, Construction of Sichuan Regional Economic Vitality Measurement System [J], Cooperative Economy and Science and Technology, 2016 (9): 16-17
[3] Zhang Mengqi, Analysis and Evaluation of Urban Vitality [D], Wuhan University, 2018.
[4] Li Changjun, Persist in Problem Orientation and Improve Jilin's Economic Vitality [N], Jilin Party School Newspaper, 2018-10-15 (002)
[5] Lei Shuzheng, Xu Dengyao, Li Yan, Comprehensive Evaluation and Analysis of Urban Vitality in Sichuan Provinces [J], Economic Forum, 2017 (09): 26-29
[6] Jin Yanjie, Evaluation of China's Urban Economic Vitality, Geographical Science [J], 2007, 27 (1): 9-16.
[7] Lou Haimiao, Sun Qiubi. Research on the Evaluation of the Economic Vitality of Different Provinces in China Based on Factor Analysis. Journal of Fuzhou University (Philosophy and Social Sciences Edition) [J], 2005, (3): 17-18.
[8] Lu Mingyang, Construction and Evaluation of Urban Vitality Index System [D], Dongbei University of Finance and Economics, 2011.