Research on the Evaluation of Economic Vitality Based on PCA, GC-TOPSIS and C.V

Taking the Districts of Shanghai for Example

Chaoyu Zhang[1], Han Zhao[2], Siqing You[3]*
[1]244636941@qq.com, [2]1357696193@qq.com, [3]93028603@qq.com

School of Information, Beijing Wuzi University, Beijing, China

Abstract—Based on the data related to the economic vitality of Shanghai in 2019, an appropriate index system from the four aspects of economy, industry, innovation and social life was established. Then three experimental and mathematical statistics analysis methods: PCA, GC-TOPSIS and discrete coefficient method are used to evaluate the economic vitality of Shanghai, and the rank of each district was obtained by a simple ranking aggregation algorithm- Borda Count. After qualitative and quantitative analysis, the conclusions are as follows: First, the Pudong New Area and Minhang District rank at the forefront, mainly because of their superior geographical location and the support of government policies. Secondly, there are two indicators that can better explain Shanghai's economic vitality. They are the social retailgoods and the total export-import volume. According to the conclusions of the study, it is suggested that all regions of Shanghai should adhere to reform and opening up and advocate policy first. Then it emphasizes the opening and the progress of business.

Keywords-Economic Vitality; List aggregation; Variation Coefficient; PCA

1. INTRODUCTION

The economic vitality is used to measure the ability and potential of urban or regional economic development. It is not only a vital aspect of reflecting the level of regional sustainable development, but also the inevitable requirement of realizing high-quality economic development. This paper ranks the economic vitality of various regions in Shanghai, analyzes the reasons why the ranking is like this and the most significant indicators, looks for ways to improve the urban competitiveness, and finally gives suggestions to the districts, especially the less developed areas. It has certain instructive sense for the economic investigate and development of other areas.

At present, domestic scholars have done some research on economic vitality and its evaluation model. In terms of evaluation model creation, by creating an evaluation model of urban economic activity, Siqian Xin aims to attain pertinent solutions to enhance economic dynamism [1]. But he simply believes that the GDP can well evaluate economic vitality and take it as the only measure of evaluation or decision-making. In the application of research methods, Lingfeng Li established an index system, used factor analysis to quantify the urban economic vitality index, and divided 22 indicators into five species. These five parts are individual and

ICEMME 2022, November 18-20, Nanjing, People's Republic of China

DOI 10.4108/eai.18-11-2022.2326741

enterprise economic benefits, civic economic development, civic residential housing and government spending. It is found that the economic benefits of individuals and enterprises are strongly associated with the vitality of the urban economy [2]. In the study of influencing factors of economic vitality, Luyao Jiang took Chongqing as an example to study the impact of population and commercial vitality on regional economic vitality [3]. Haitong Zhao took Tianjin as an example to study the impact of various aspects on areal economic vitality. He analyzed the influence of various elements on economic vitality by factor analysis, and bring up a train of policies to provide incentive to the development of urban vitality [4]. The above studies all use the method of factor analysis. However, after most of the information in multiple indicators is condensed into several relatively few indicators, the significance of the obtained factors can not be completely determined. Therefore, it is difficult to select the most influential indicators and make correct analysis and decision-making.

This paper creatively uses the quantified data, three statistical analysis methods: PCA, GC-TOPSIS and discrete coefficient method, and introduces the idea of Borda Count method to rank the districts of Shanghai. Then, it analyzes the impact of location and policies on economic vitality with qualitative analysis method.

2. DATA HANDLING AND ANALYTICAL METHODS

2.1 Research Object Selection and Data Sorting

As one of China's economic centers, Shanghai has a high level of urban economy, which has certain research significance. The data used in the research are sorted and calculated by the data officially released by the Shanghai Municipal Bureau of statistics and the Shanghai 2020 statistical yearbook.

2.2 Evaluation Index Selection

Urban economic vitality is affected by many factors, so evaluating it is a complex and comprehensive evaluation process. After consulting relevant literature, this paper finally determines the evaluation system of urban economic vitality which includes 4 influencing factors and 12 specific evaluation indexes. The specific indexical hierarchy is demonstrated in the table 1. The vitality of economic growth mainly reflects the existing economic development and prospective state of affairs in the region. The vitality of education and innovation mainly reflects the ability of regional talent training and technological innovation. The vitality of industry and capital mainly shows the affinity of the region to industry and capital. Social vitality mainly mirrors the living standards of local residents These four parts are different but correlated, which together constitute the evaluation index model.

2.3 Analytical Methods

Each objective weighting method has its own advantages and disadvantages, so using a single analysis method will usually get different analysis results, resulting in poor results. To avoid this situation, this paper uses a great variety of methods to dissect. Then leads into the idea of list aggregation

TABLE 1. INDICATOR SYSTEM

Influential Factors	Indicators	Dimensional Units
vitality of economic growth	GDP	hundred million RMB
	per capita GDP	ten thousand RMB
	general public budget revenue	hundred million RMB
vitality of education and innovation	expenditure for education	hundred million RMB
	expenditure for science and technology	hundred million RMB
vitality of industry and capital	total output value of industries above designated size	hundred million RMB
	investment in the fixed assets	hundred million RMB
	total export-import volume	hundred million RMB
	added value of the tertiary industry	hundred million RMB
social vitality	Permanent resident population	ten thousand person
	non-resident population	ten thousand person
	total retail sales of consumer goods	hundred million RMB

to evaluate the economic vitality of Shanghai.

1) Initial Evaluation Methods

After consulting relevant materials and comparing, it is decided to use three methods for initial evaluation. The three methods are PCA, GC-TOPSIS and discrete coefficient method.

Principal Component Analysis is a sort of dimension decrease analytically approach. Its goal is to integrate many original indicators with certain relation into a new set of irrelevant inclusive parameters to replace the original indicators. This method can remove the dependence between the original evaluation indexes but the meaning of principal component interpretation is generally a little fuzzy, which is not as distinct and precision as of the signification of original variables. The main procedures are as follows:

Step 1: standardize the original data:

Suppose that there are m index variables for principal component analysis: $x_1, \cdots, x_m$, a total of n evaluation objects, and the value of the j index of the i evaluation object is x.

$$\tilde{x}_{ij} = \frac{\tilde{x}_{ij} - \overline{x}_j}{s_j}, (i = 1, 2, \cdots, n; j = 1, 2, \cdots, m) \tag{1}$$

Step2: calculate the correlation coefficient matrix R:

Correlation coefficient matrix $R_{ij} = (R_{ij})_{m \times m}$.

$$r_{ij} = \frac{\sum_{k=1}^{n} \tilde{x}_{ki} \times \tilde{x}_{kj}}{n-1}, (i, j = 1, 2, \cdots, m) \tag{2}$$

Step 3: calculate eigenvalues and eigenvectors:

First, calculate the proper value of the correlation coefficient matrix R: $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \geq 0,$ and the corresponding eigenvector: $u_1, u_2, \cdots, u_m$. New m index variables are composed of eigenvectors. In the formula, y_m is the m-th principal component. The calculated eigenvalues should be organized from large to small.

$$\begin{cases} y_1 = u_{11}\tilde{x}_1 + u_{21}\tilde{x}_2 + \cdots + u_{n1}\tilde{x}_n \\ y_2 = u_{12}\tilde{x}_1 + u_{22}\tilde{x}_2 + \cdots + u_{n2}\tilde{x}_n \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_m = u_{1m}\tilde{x}_1 + u_{2m}\tilde{x}_2 + \cdots + u_{nm}\tilde{x}_n \end{cases} \tag{3}$$

Step 4: select principal components and calculate the final appraisal value:

Calculate the information contribution rate of each eigenvalue $\lambda_j (j = 1, 2, \cdots m)$ .Then calculate the cumulate contribution rate of each eigenvalue arranged in order.

$$b_j = \frac{\lambda_j}{\sum_{k=1}^{m} \lambda_k} (j = 1, 2, \cdots m), \alpha_p = \frac{\sum_{k=1}^{p} \lambda_k}{\sum_{k=1}^{m} \lambda_k} \tag{4}$$

In the formula, b_j indicates the information contribution rate of the j-th principal component. And α_p is the cumulative contribution rate of principal components: $y_1, y_2, \cdots y_p$ .When α_p is close to 1, the first p index variables are selected as p principal components to replace the original m index variables, so as to analyze the p index variables. Finally, calculate the comprehensive score.

$$z = \sum_{j=1}^{p} b_j y_j \tag{5}$$

In this formula, b_j represents the contribution rate weight of this principal component in the selected several compositions.

The grey correlation TOPSIS approach coalesces the idea of fuzzy positive and negative desirable points. The distance between the evaluation object and the plus-minus ideal points is taken as the evaluation standard. This method effectively settles the dispute that it is tough to correctly quantitate and reckon the evaluation indexes. It reduces the influence of human prejudiced features to a certain extent. It has low requirements for sample size and does not need typical distribution rules.

The coefficient of variation method directly uses the original data value of each indicator to decide the magnitude affecting the final evaluation value. When the evaluation index is vague for the evaluation objective, it is more appropriate to use the C.V method for utilization. It is applicable to the determination of the internal index significance of each constituent element.

The disadvantage is that it does not pay enough attention to the specific economic significance of the index. The main steps are as follows:

Step 1: calculate the coefficient of variation

Suppose there are n samples to be judged and p evaluating indicator to form the original data matrix, x_{ij} signifies the value of the j-th appraisal indicator of the i-th sample.

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{np} \end{pmatrix}, v_j = \frac{s_j}{\overline{x}_j}, j = 1, 2, \cdots, p \tag{6}$$

Step 2: Calculate the weight of each index:

$$w_j = \frac{v_j}{\sum_{j=1}^{p} v_j} \tag{7}$$

The calculated final index weight: W= $\{w_1, w_2, \cdots, w_p\}$.

2) Ranking Aggregation Method

Borda Count is a sort of voting method. Every voter has his own set of evaluation indicators, so everyone has a ranking for all candidates. Each candidate obtains corresponding points according to different ranking, and the candidate with the highest points wins the election. This idea can be used to simply rank and aggregate: the number of objects defeated by the object in a list (for example, in an arrangement with length m, the score of the first place is m-1, and the score of the m-th place is zero). Finally, the final ranking is obtained by simple summation.

3. RESULTS AND DISCUSSION

3.1 Ranking of Economic Vitality in Shanghai

Firstly, use three methods to calculate the significance extent value of each indicator. Then multiply the standardized value of the original data by the weight. Finally, use the idea of the Borda Count method to aggregate the ranking to obtain the final ranking.

1) Calculation of Initial Weight and Ranking

SPSS software is used to execute factorial analysis on the economic vitality of Shanghai. First, the kmo and Bartlett tests were carried out, and the results are shown in Table 2. After testing, the kmo value is 0.723, Bartlett test value is 402.132, SIG. = 0.000, indicating that it meets the conditions of factor analysis and can be analyzed by PCA.

TABLE 2. KMO AND BARTLETT TESTS

Kmo and Bartlett tests		
Kmo sampling suitability quantity		.723
Bartlett sphericity test	Approximate chi square	402.132
	degree of freedom	66
	conspicuousness	.000

TABLE 3. TOTAL VARIANCE INTERPRETATION

Total variance interpretation			
component	Initial eigenvalue	variance percentage	Cumulative %
1	9.792	81.602	81.602
2	1.39	11.583	93.185
3	0.409	3.41	96.595
……	……	……	……
11	0.001	0.008	99.997
12	0	0.003	100
Extraction method: principal component analysis method.			

TABLE 4. PRINCIPAL COMPONENTS EXTRACTED

	component	
	1	2
GDP	0.984	0.151
per capita GDP	0.225	0.919
general public budget revenue	0.993	0.066
total output value of industries above designated size	0.873	-0.309
investment in the fixed assets	0.96	-0.23
total export-import volume	0.985	-0.007
added value of the tertiary industry	0.942	0.29
expenditure for education	0.778	0.321
expenditure for science and technology	0.987	0.037
permanent resident population	0.894	-0.362
non-resident population	0.961	-0.22
total retail sales of consumer goods	0.972	0.053

The total variance interpretation is shown in Table 3. Only the eigenvalues of the first two factors are greater than one, and the adding of their contribution rates is close to one, so determine the first two factors to be the major parts. The variance interpretation rates of the two principal components are 81.602% and 11.583% respectively, and the cumulative variance interpretation rate is 93.185%. It shows that most of the indicators selected in this paper have clear meaning and meet the requirements of PCA. The two principal components extracted are shown in Table 4. Then calculate the importance of the two major parts in the new index system, that is, the significance of the principal components. Finally, multiply the standardized data by the weight to obtain the comprehensive score and rank.

Taking Python language as the main body, the standardized data are calculated by using package files such as numpy to obtain the dimensionless distance and grey related extent. Through them, the relative proximity between each city and the ideal object is obtained, so as to measure the closeness between each city and the ideal object. The results are shown in Table 5. Finally, the economic vitality of the city is evaluated according to the relative closeness.

TABLE 5. GC-TOPSIS RESULTS

	R_i^+	R_i^-	S_i^+	S_i^-	C_i^+
Pudong New Area	1.00	0.60	1.00	0.31	0.76
Minhang District	0.66	0.87	0.43	0.84	0.34
Jiading District	0.64	0.90	0.40	0.88	0.32
Jing'an District	0.65	0.92	0.40	0.89	0.31
Songjiang District	0.64	0.91	0.39	0.89	0.30
Baoshan District	0.63	0.93	0.37	0.92	0.29
Huangpu District	0.63	0.94	0.37	0.93	0.28
Xuhui District	0.62	0.94	0.36	0.93	0.28
Yangpu District	0.62	0.95	0.35	0.94	0.27
Qingpu District	0.62	0.95	0.35	0.94	0.27
Fengxian District	0.61	0.96	0.34	0.95	0.26
Jinshan District	0.61	0.96	0.34	0.96	0.26
Changning District	0.61	0.97	0.33	0.96	0.26
Putuo District	0.61	0.98	0.32	0.97	0.25
Hongkou District	0.61	0.98	0.32	0.98	0.25
Chongming District	0.60	1.00	0.30	1.00	0.23

TABLE 6. C.V RESULTS

	mean value	Standard deviation	CV	weight
GDP	2353.3881	2844.82622	1.21	0.08
per capita GDP	15.3856	8.63136	0.56	0.04
general public budget revenue	235.505	230.68139	0.98	0.06
total output value of industries above designated size	2089.7356	2665.53728	1.28	0.08
investment in the fixed assets	452.6094	478.44699	1.06	0.07
total export-import volume	2108.5663	4957.18857	2.35	0.15
added value of the tertiary industry	1738.2587	2246.91972	1.29	0.08
expenditure for education	12.9	13.52339	1.05	0.07
expenditure for science and technology	43.4556	30.27013	0.7	0.05
permanent resident population	61.1069	58.27272	0.95	0.06
non-resident population	151.7588	120.19218	0.79	0.05
total retail sales of consumer goods	3446.6831	11199.35564	3.25	0.21

Calculate the mean value and SD of the original data of each indicator. Then use both of them to calculate the cov of each indicator. Finally, the significance of each index is the normalized outcome of the variation coefficient. The computed result are shown in Table 6.

2) Calculation of Final Ranking

Due to the differences in the ranking obtained by various methods, each method has different preferences for indicators. Therefore, the ranking obtained by the first three methods should be aggregated to get the total ranking. The number of votes (score) of each object under each method is m-1 (m is the number of objects). The scores of the three means are added to get the

final ranking. The final ranking are shown in Table 7.

TABLE 7. FINAL RANKING

District	PCA Ranking	GC-T Ranking	C.V Ranking	Aggregated ranking
Pudong New Area	1	1	1	1
Minhang District	2	2	2	2
Jiading District	3	3	3	3
Jing'an District	4	4	5	4
Songjiang District	5	5	4	5
Huangpu District	6	7	6	6
Baoshan District	7	6	7	7
Xuhui District	8	8	8	8
Yangpu District	9	9	9	9
Qingpu District	10	10	10	10
Fengxian District	11	11	11	11
Jinshan District	12	12	12	12
Changning District	13	13	13	13
Putuo District	14	14	14	14
Hongkou District	15	15	15	15
Chongming District	16	16	16	16

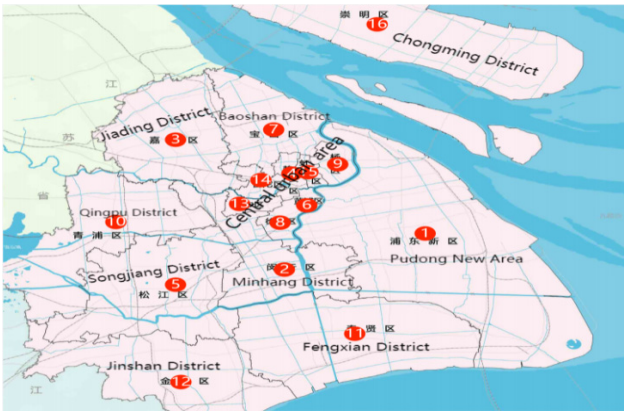

Figure 1. Ranking distribution

The ranking distribution is shown in Figure 1. It can be aware of that the ranking of Shanghai districts is related to their geographical location. Most areas with high economic vitality are distributed near the central urban area and along the Yangtze River. Most of these places have developed water, air and railway transportation. For example, Minhang District has one of the

most important nodes of the national transportation lifeline- Hongqiao Station. The Pudong New Area is an important transportation hub of Shanghai and an advanced international logistics port. Pujiang bridge, undersea tunnel, maglev train and subway line form a dense transportation network.

The Pudong New Area ranked first in the ranking. In April 1990, the CPC Central Committee officially approved the opening up of the Pudong. Since the 18th CPC National Congress, the CPC Central Committee has attached great importance to the development and opening up of Pudong and placed a series of national strategic tasks in Pudong. Over the past thirty years, Pudong is fully implementing the national strategic task and vigorously promoting innovation and reform. It strives to be at the forefront of the country in the way of high-quality development and domestic life. In the course of the 14th Five-Year Plan period, it will undertake the task of high-standard reform and opening up and building a leading area for socialist modernization: building a world-class industrial cluster, promoting the high-quality development of manufacturing and modern service industries.

Minhang District ranked second, is not only the geographical center of Shanghai but also the central area of industrial layout. Caohejing emerging technology development zone, established in 1988 and located at the junction of Minhang District and Xuhui District, is the first batch of national economic and technological development zones focusing on the development of emerging technology, the introduction of foreign capital and advanced technology. Minhang Economic Development Zone, founded in 1983, is one of the first two national development zones established in Shanghai. With complete policy guarantee, strong financial support and open positioning, they have brought a high-quality industrial foundation to Minhang District, and more and more enterprises have settled in. Hongkou District ranks low and covers a small area. It is in the stage of transformation of the old city, but due to policy mistakes, the upgrading of business and the promotion of purchasing power are unbalanced, resulting in the stagnation of transformation progress. Perhaps after a long period of hard work, it can usher in a take-off development based on the location and economic strength.

It can be seen from the above analysis that the improvement of economic vitality is inseparable from the guidance of correct political decision-making.

3.2 Analysis of Indicators

The development of foreign trade can exchange the needs of different countries to ensure the best distribution of resources. It can save social manpower to obtain better profitable productivity. It can also absorb the progressive scientific and technological attainments of the outside world. It will make the country confront the rival force and contests in the international community, but the pressure is the driving motivation, which can advance domestic enterprises to constantly update technology, improve labor productivity and accelerate internationalization. The social retailgoods describe the business and reflect the size of the business. It can be used to study the changes in the regional retail market and is a considerable indicator of economic affluence.

According to the weight analysis calculated in the research process, the social retailgoods and the total export-import volume are the two most important indicators affecting urban economic vitality. Their weights are 0.15 and 0.21 respectively, which are far greater than other indicators. In other words, efforts to develop these two indicators will be more in favour of the

improvement of economic vitality.

4. CONCLUSION

Three methods: PCA, GC-TOPSIS and discrete coefficient method are used for quantitative analysis, then make a qualitative analysis on the policy, the conclusions are as follows: firstly, most of the top economic vitality are distributed near the central city and along the Yangtze River, except for the Chongming district. Secondly, the development level of foreign trade and commerce can better affect the regional economic vitality. Thirdly, for the sake of improving economic vitality, it not only needs the concern of the government, but also needs to formulate appropriate policies in combination with the actual situation.

More important than the evaluation of regional economic vitality is how to improve it. Here are some suggestions: firstly, scientific and effective policy formulation is essential. The local government should continue to focus on the development of the Pudong New Area, establish a pioneering area of socialist modernization. And they need to enlarge monetary investment and policy assistance for slower developing regions. Secondly, the district government must take the development of foreign trade and commerce seriously, to add vitality for local industries and improve people's living standards. Thirdly, Pudong should play a leading role in radiation, enhance its core functions and build it into a functional highland for China's deep integration into the global economic pattern.

Acknowledgement. Beijing Social Science Fund 19GLC051

REFERENCES

[1]Siqian Xin."Analysis and Decision of Regional Economic Vitality and its Influencing Factors." Frontiers in Economics and Management 1.11(2020)

[2]Lingfeng Li and Wenxi Kong. Quantitative Assessment of Urban Economic Vitality[J]. Scientific Journal of Economics and Management Research,2020,2(11)

[3]Luyao Jiang, Shi Liu ,and Weijia Xiao."The Impact of Population and Business Vitality on Regional Economic Vitality--Take Chongqing as an Example." International Core Journal of Engineering 6.9(2020)

[4]Haitong Zhao, Weina Wu and Yihan Chu."Research on Regional Economic Vitality based on Factor Analysis." Frontiers in Economics and Management 1.2(2020)

[5]Daojie Zhou,Na Zhao and Yafeng Gao. Model Based on Regional Economic Vitality and Its Influencing Factors[J]. International Journal of Social Sciences in Universities,2021,4

[6]Sihao Fu."Research on the Influence of Regional Economic Development on Shenzhen's Economic Vitality." Finance and Market 5.3(2020)

[7]Yifei Jin,Yang Li and Qian Feng. Evaluation of Urban Economic Vitality Based on Principal Component Analysis[J]. The Frontiers of Society, Science and Technology,2020,2(5)

[8]Chengxiang Xu,Xiaoyang Yu,Jiayi Zhang and Zhile Xia. Research on Regional Economic Vitality Based on Factor Analysis[J]. International Journal of Mathematics & Computation,2020,31(2)

[9]Yidan Liu , Yujia Cheng and Fang Sun. Analysis of Urban Economic Vitality Index based on Principal Component Analysis[J]. International Journal of Intelligent Information and Management Science,2020,9(2)

[10]Qing Lv and Zhiyang Wu. Evaluation of Economic Vitality of Beijing and Analysis of Influencing Factors - Based on Entropy Method and Grey Correlation Analysis Model[J]. Open Journal of Social Sciences,2020,08(04)
[11]Jian Luo."The Prediction of Economic Vitality in Hunan Province and Its Influencing Factors." Statistics and Application 09.01(2020)
[12]Zhiyang Wu and Zejiong Zhou. Forecast and Analysis on Economic Vitality of Economic Zone Based on the New Trend of Enterprises[J]. Scientific Journal of Intelligent Systems Research,2020,2(8)
[13]Ying Long."Does block size matter? The impact of urban design on economic vitality for Chinese cities." Environment and Planning B: Urban Analytics and City Science 46.3(2019)
[14]Lan Feng."How do population inflow and social infrastructure affect urban vitality? Evidence from 35 large- and medium- sized cities in China." Cities 100.prepublish(2019)