The Linear Model Using AHP-FCE Method For Economic Vitality

Economic vitality are power of a running economic system, and out-standing vitality will give the local people confidence to building the regional construction. In this article, we build a **Linear AHP-FCE Model**(Analytic Hierarchy Process-Fuzzy Comprehensive Evaluation Model) to evaluate the economic vitality. To begin with, we build a relational model for economic vitality. And then, we choose the Guangdong Province as research object. For different regions, we use **Wilcoxon Signed Ranks Test** to obtain the difference between where policies are implemented and where are not. Then we combine the **Keynesian economic theory** to analyze the short-term and long-term effects of economic policies. Afterwards, we use **Correlation Analysis** to indicators, and select low-correlation indicators for **Analytic Hierarchy Process** to . Then combining with the idea of **Fuzzy Comprehensive Evaluation** establishing the Linear Model for evaluating the economic vitality. Finally, divided regions into different kinds by **K-Means++**, observe the relationship and difference between the regions that owns the similar resources endowment.

Keywords: Analytic Hierarchy Process Correlation Analyze Fuzzy Comprehensive Evaluation K-Means++ Keynesian economic Wilcoxon Signed Ranks Test

I. Introduction

1.1 Background

The regional (or urban or provincial) economic vitality plays an important role in regional comprehensive competitiveness, which is also the main driving force of regional economic development. In recent years, in order to improve economic vitality, some regions have launched a lot of preferential policies to stimulate economic vitality, such as reducing the approval steps for investment, providing capital support to startups and lowering the settlement threshold to attract the talented. However, due to different resource endowments, these policies have different effects in different regions. How to seize the key factors and effectively improve the regional economic vitality is a worth research topic.

1.2 Related Work

At present, the main methods for evaluating economic vitality include factor analysis, analytic hierarchy process, data envelopment analysis, and expert scoring method and so on.

1.3 Contributions

1. Before the Analytic Hierarchy Process, using Correlation Analyze to select the low correlation indicator

2. Firstly combine the Analytic Hierarchy Process, Fuzzy Comprehensive Evaluation, K-Means++ and Correlation Analyze for model building and data analyzing
3. The final model is a Linear Model, which is easy to analyze in economic.
4. Use the improved K-Means Algorithm, K-Means++ for data analyze.
5. Combine with Keynesian economic theory to analyze phenomenon.

II. Models & Basic Problems Solution Process

2.1 Basic Model

2.1.1 *Assumptions*

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 other unlisted factors does not significantly affect economic vitality.
- When the selected indexes are compared and evaluated in pairs, the absolute values can be converted into proportional scale values.
- Cities with similar economic dynamism have similar future developments.

2.1.2 *The Foundation of Model*

- Correlation Analyze
- Analytic hierarchy process
- Fuzzy Comprehensive Evaluation
- K-clustering

2.2 Establishment of index system

Regional economic vitality refers to the capacity and potential of urban economic development. On the one hand, it is the attraction and cohesion for production factors. On the other hand, it is the efficiency of the allocation and use of these production factors.

Guangdong Province ranks first among the provinces in many economic indicators as China's largest economic province. Therefore, we take Guangdong Province as an example to analyze.

The differences in regional economic vitality among provinces are mainly reflected in the differences in resource endowments. We divide resource endowments into seven aspects as the figure shows.

In order to quantify the impact of these seven factors on regional economic vitality, we have selected some indicators, as follows:

Taking fiscal expenditure as an example, fiscal expenditure on science and education can promote technological progress and personnel training in Guangdong Province, to help

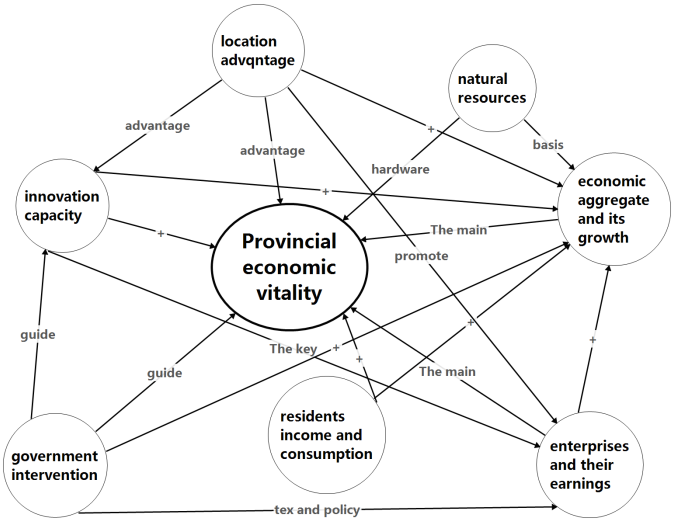

Figure 1 Relational model of influencing factors

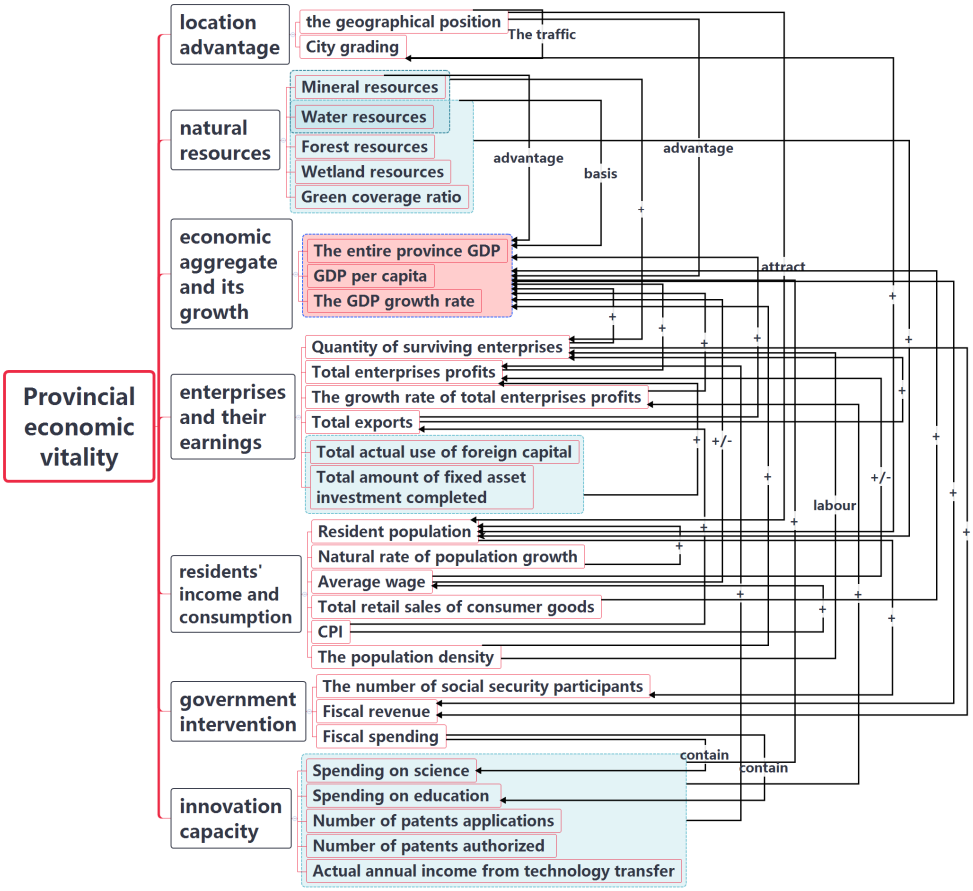

Figure 2 The relationship between these indicators

enterprises improve production capacity and labor quality. With the increase of production capacity, the marginal cost has decreased, and enterprises profits have increased, accumulating capital for expanding production. Increased enterprises vitality has increased Guangdong's GDP. At the same time, it will help increase per capita disposable income, improve residents' living standards, and attract more talents.

The programs of action are as follows:

	Definition	Policy
Goal 1.	Create opportunities for economic development	1.Lowering the threshold for companies to go public 2.Leverage the agglomeration and synergy effects of industrial clusters 3.Simplify the application process
Goal 2.	Encourage businesses and business opportunities that add value to the community or region	1.Tax Cuts 2.Lower lending rates 3.Lower tariffs
Goal 3.	Balance economic development with quality of life	1.Expand government public investment 2. Complete the social security system 3.Encourage the ability of each business district to provide jobs 4.Make decisions based on local resource endowment

Judging from the factors that affect regional economic vitality, resource endowment is the basis of regional development. Enterprises are users of factors of production and a major source of regional economic vitality. Based on their resource endowment, regional governments can improve the vitality of the regional economy through fiscal and monetary policies.

From the data released by the National Bureau of Statistics, we use the total profit of Guangdong enterprises and its growth rate to measure the vitality of Guangdong enterprises, and use the number of permanent residents and the natural growth rate of Guangdong to measure the population trend in Guangdong.

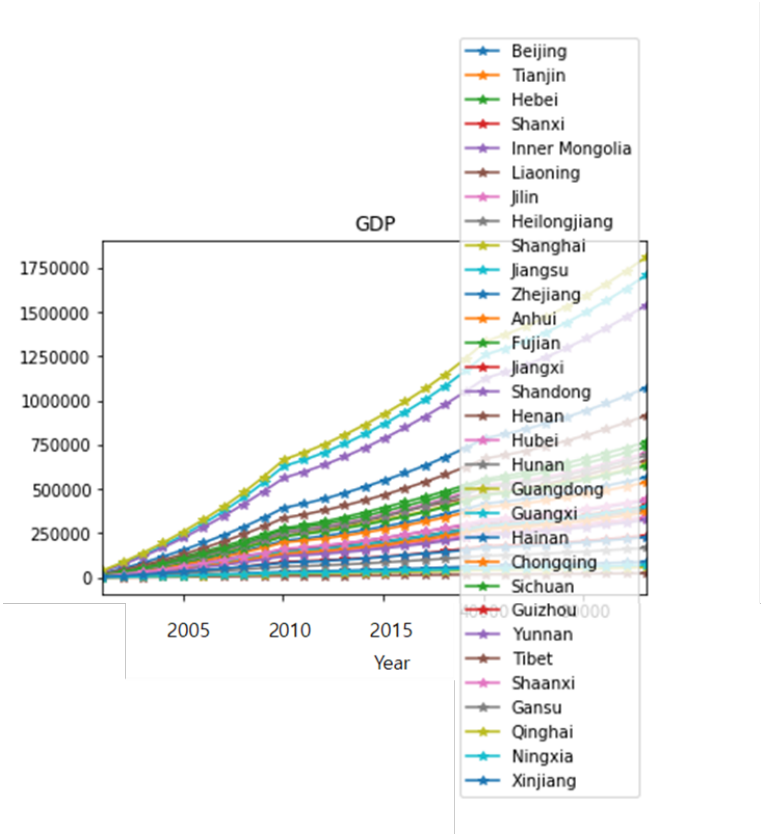

Figure 3 The relationship between these indicators

It can be seen from the figure that the growth rate of profits of Guangdong enterprises is declining, the total profit is increasing, and the vitality of enterprises has begun to decline.It is indicating that Guangdong's economic development has entered a mature period.

The number of permanent residents in Guangdong has continued to increase. Although the natural growth rate of the permanent population has declined, it is still higher than the natural population growth rate of the country (0.5 percent), indicating that Guangdong has a strong appeal to the labor force. Talents are an important source of regional economic vitality, and the continuous attraction of talents is the driving force for the sustainable development of Guangdong Province.

The changes in the resident population did not make up for the impact of the decline in corporate vitality on regional economic vitality. Therefore, the real GDP growth rate of Guangdong Province has continued to decline since 2010, and economic vitality has decreased. Of course, the impact of the 2008 global economic crisis is also an important factor.

2.2.1 *Analysis of the Result*

2.3 Model application for specific region

We choose Guangdong Province as analysis object in this article, focusing on the economic policies around 2009.

In Figure 5, we find that: The GDP growth rates have risen steadily around 2005-

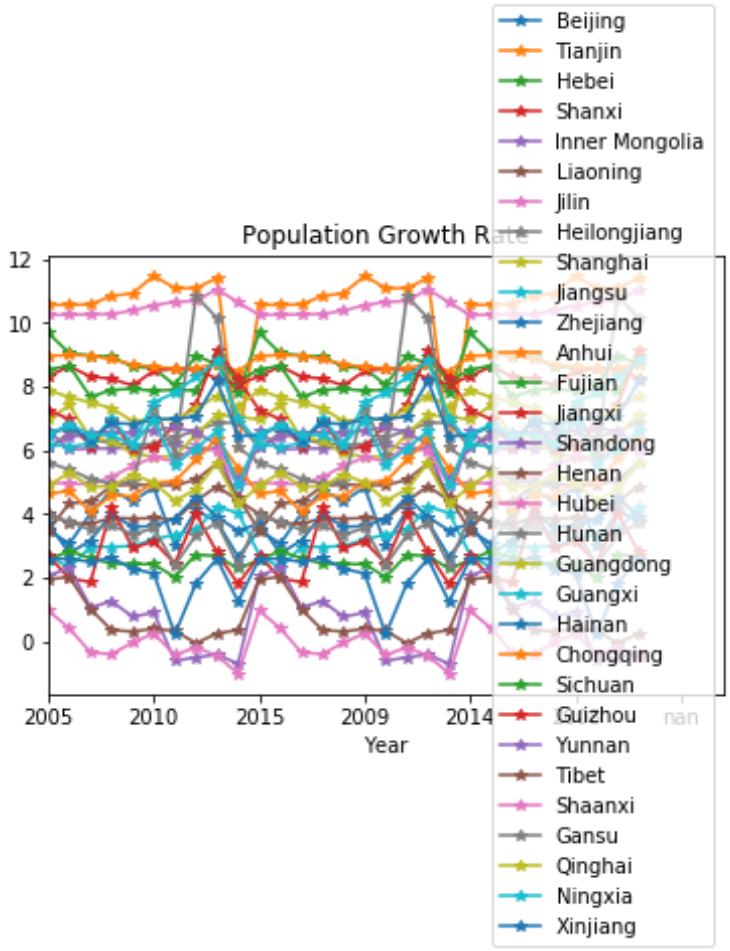

Figure 4 The relationship between these indicators

2015, with no apparent abrupt change, while the population growth rate dropped sharply in 2009. The Enterprise profit growth significantly increased in 2009, and quickly back to the previous level before 2009 in the following peroid.

After analyzing the background, we draw the major factor:Financial Crisis. In 2008, a world-wide financial crisis attacked the economic environment, and disequilibrated the monetary system. In 2009, facing with the serve international financial crisis, People's Government of Guangdong Province introduced several policies to cope with the crisis:

- 16 livelihood engineering policies.
  In terms of social security, training of rural labor, three integrated medical insurance, rural social security and minimum living security, and New Rural Cooperative Medical Care. In terms of living, newly built low-rent housing and affordable housing. In terms of sanitation and environmental protection, transforming health centers and community hospitals, treating rivers and building sewage plants. In terms of education, solving the problem of substitute teachers, rural education school building renovation, and vocational education engineering problem, comprehensive implementation of free compulsory education in urban and rural areas.
- 20 price policies.

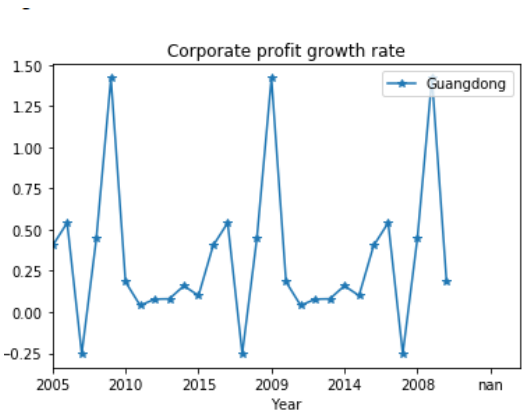

Figure 5 The Enterprise profit growth rate of Guangdong Province from 2005 to 2015

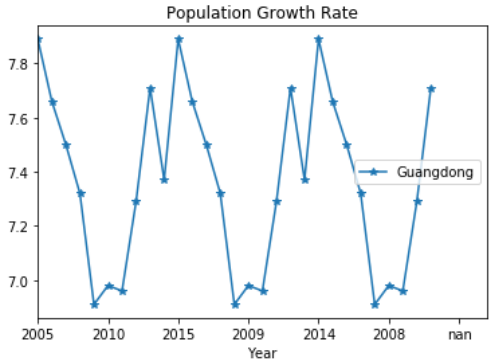

Figure 6 The Guangdong population growth rate of Guangdong Province from 2005 to 2015

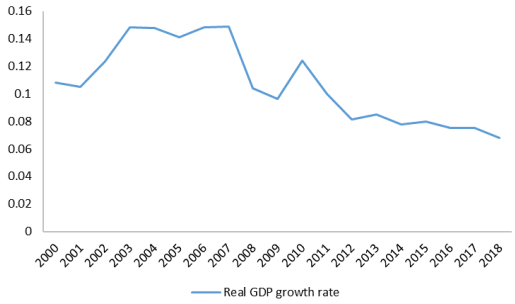

Figure 7 The GDP growth of Guangdong Province from 2005 to 2015

Strictly control the introduction of charging items involving enterprises. Use price leverage to promote energy conservation and emission reduction, improve the price mechanism, and promote industrial upgrading. Continue to rectify the fees and charges related to agriculture, strengthen farmers and benefit farmers, push "Guangdong prices", and strengthen pricing power.

- Eight preferential tax measures support the development of small and medium-sized enterprises.
- Adjust the economic structure and use fiscal expenditure to promote the development of high-tech zones.

  Newly-built industrial transfer industrial parks will provide preferential construction

	short term	long term
Total enterprise profits	no statistical significance	no statistical significance
Enterprise profit after tax	statistical significance	no statistical significance
GDP	statistical significance	statistical significance(strong)
GDP per capita	statistical significance	statistical significance(strong)
Resident population	statistical significance	statistical significance(strong)
Average wage	statistical significance(weak)	no statistical significance
green coverage ratio	no statistical significance	no statistical significance

land indicators to promote research and development in high-tech zones. Introduce high-end industrial projects and incubate independent intellectual property rights projects. Accelerate the transformation of major innovations.

- Nine measures fully support the transformation and upgrading of enterprises.
- Customs facilitation measures promoted exports to domestic sales, achieving substantial growth.
- New ten projects.
- Expand effective domestic demand. Home appliances go to the countryside and cars go to the countryside.

2.3.1 *Analysis Method*

- Empirical Analysis
- Keynesian macroeconomics

2.3.2 *Empirical Analyze*

Based on clustering of existing indicators of all cities, provinces similar to Guangdong include Fujian, Jiangsu, and Zhejiang. According to these policies, short-term or long-term indicators for these three regions and Guangdong The comparison of the change rate (the change rate of the indicators that have an impact on changes in economic vitality in the first question) can determine whether Guangdong's policy changes have significantly affected these indicators and thus affected economic vitality.

The indicators we selected are: total enterprise profits, GDP, GDP per capita, resident population, average wage, green coverage ratio. The short-term definition is 3 years, and the long-term definition is so far (2018). Guangdong Province is the experimental group, and others three regions served as the control group, and Wilcoxon Signed Ranks Test was performed on the two groups of data. ($\alpha = 0.05$) The results are as follows:

In the short term, these policy changes have an impact on corporate after-tax profits, GDP, resident population, and Average wage, and thus have an important impact on economic vitality. However, in the long run, the impact of policy changes has been transformed into GDP (and GDP per capita). Taking the short-term residential population as

an example, the test results are as follows:

Ranks

		N	Mean Rank	Sum of Ranks
Other - GuangDong	Negative Ranks	7[a]	5.71	40.00
	Positive Ranks	2[b]	2.50	5.00
	Ties	0[c]		
	Total	9		

a. Other < GuangDong

b. Other > GuangDong

c. Other = GuangDong

Figure 8

Test Statistics[a]

	Other - GuangDong
Z	-2.073[b]
Asymp. Sig. (2-tailed)	.038
Exact Sig. (2-tailed)	.039
Exact Sig. (1-tailed)	.020
Point Probability	.006

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

Figure 9

2.3.3 *Keynesian Macroeconomics Analyze*

According to the AD-AS model and the classical model of macroeconomics, short-term aggregate supply will not change much, and prices are sticky, that is, commodity prices and labor prices are almost unchanged.

In the short term, tax cuts and subsidies have boosted corporate sales by stimulating consumption and government spending, increasing their operating income and thereby increasing their profits. The increase in corporate profits has stimulated production and increased GDP and GDP per capita. With the increase in per capita GDP, the disposable income of residents has increased, and their lives have improved, thereby attracting more talents to Guangdong Province.

In the long run, social security, education, price mechanisms and economic restructuring play a role. In the long run, prices are elastic, and costs increase as business income increases, so corporate profits remain unchanged or even decline. With the improvement of technological level, the production capacity of enterprises has increased, which has increased the potential total output and total supply. With the adjustment of the economic structure, the consumption of residents, enterprises and governments has increased, and aggregate demand has increased. Finally, at equilibrium, total output increases, manifested in increased GDP and per capita GDP, and increases regional economic vitality.

Obviously, policy makers have fully considered the short-term and long-term effects of policies on economic vitality, and aim to stabilize economic growth to achieve high-quality development.

2.4 Mathematical model analyzes and measures the regional economic vitality

2.4.1 *Basic Method*

We choose 10 indicators for Problem 3. which are available, as shown in Table 1.

Cause the correlation between different factors will result that, the same grey factor accounting for more for the consquence, we use the **Pearson Correlation Coefficient** to select the resonable factors for models.

What's more we use **Analytic Hierarchy Process** to estimate the weight of corresponding index.

Then, combing the weight we caculated above, using **Fuzzy Comprehensive Evaluation** to estimate the linear euqation for economic vitality.

In the end, according the model we built, rank the economic vitality of cities in Attachment 3.

2.4.2 *Correlation analysis*

Pearson Correlation Coefficient is used to measure whether two data sets are on the same line, and it is used to measure the linear relationship between distance variables.

- Pearson Equation

Table 1 Symbols refer to indicators

Symbol of indicator	Indicators
A1	Tier 1 cities
A2	Fixed capital
A3	Quantity of Labours
A4	Water Source
A5	The quantity of Enterprise
A6	Proportion of scientific research funds in government fiscal expenditure
A7	Proportion of education research funds in government fiscal expenditure
A8	Primary industry
A9	Secondary industry
A10	Tertiary Industry

$$r = \frac{N\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{N\sum x_i^2 - (\sum x_i)^2}\sqrt{N\sum y_i^2 - (\sum y_i)^2}} \tag{1}$$

Table 2 The Pearson Correlation Class Definition

Person Correlation Value Range	Correlation Degree
0.8-1.0	very strong correlated
0.6-0.8	strong correlated
0.4-0.6	moderately correlated
0.2-0.4	weak correlated
0.0-0.2	very weakly correlated or uncorrelated

Table 2. is the comparison table for the Correlation level.

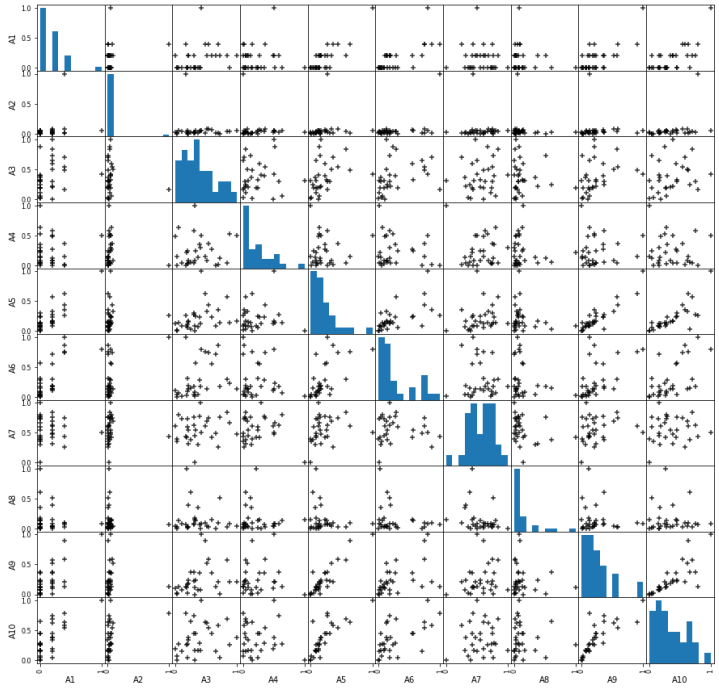

Figure 10 The correlation image of A1-A10 factors

Figure 10. is the scatter plot for indicators A1-A10, we can see that there are significant positive correlation between A9 and A10, A10 and A5, A5 and A9, A5 and A10, A6 and A9. The indicator A5 is the quantity of Enterprises, the measurement data for it is Attachcment 1 in problem paper. The scatter plot reflects that the quantity of Enterprises have a great effect on economic evaluation.

Following is the Pearson Correlation Coefficient result for indicator A1-A10 We can see that there are strong correlation between Tertiary Industry and Tier 1 cities, The quantity of Enterprise, Secondary Industry. The same to A1, A2, A7, A9, A10. Finally the indicators we choose for Analytic Hierarchy Process Model is **A3, A4, A5, A6, A8**

Table 3 Correlation Table For A1-A10

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10
A1	1	0.37891	0.26557	0.05251	0.83905	0.70896	-0.05816	-0.16172	0.71132	0.70035
A2	0.37891	1	0.43711	-0.04916	0.724	0.28786	0.35184	-0.18352	0.83534	0.68193
A3	0.26557	0.43711	1	0.06033	0.34021	0.32298	0.1274	-0.08743	0.31773	0.23733
A4	0.05251	-0.04916	0.06033	1	0.04125	-0.12967	-0.09464	-0.09507	0.04601	-0.06777
A5	0.83905	0.724	0.34021	0.04125	1	0.60563	0.22813	-0.20981	0.90213	0.78977
A6	0.70896	0.28786	0.32298	-0.12967	0.60563	1	-0.05942	-0.22727	0.50686	0.58177
A7	-0.05816	0.35184	0.1274	-0.09464	0.22813	-0.05942	1	-0.2724	0.19682	0.08291
A8	-0.16172	-0.18352	-0.08743	-0.09507	-0.20981	-0.22727	-0.2724	1	-0.12155	-0.0535
A9	0.71132	0.83534	0.31773	0.04601	0.90213	0.50686	0.19682	-0.12155	1	0.79683
A10	0.70035	0.68193	0.23733	-0.06777	0.78977	0.58177	0.08291	-0.0535	0.79683	1

2.4.3 *Analytic Hierarchy Process*

Analytic hierarchy process(AHP), refers to the decision method that separate the element always related to the decision into the levels of goals, criteria, plans, etc.., and carries out qualitative and quantitative analysis on this basis.

- Hierarchical structure model The goals,first-level-factors and second-level-factors are divided into the top, middle and bottom layers according to their mutual relations. The top level is the target variable to measure. And after the Correlation analysis, we get the model as follows:

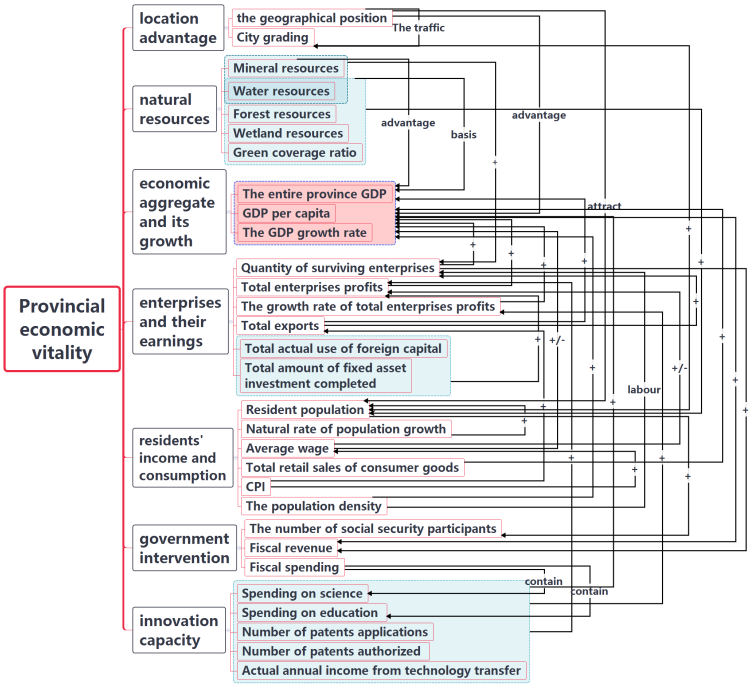

Figure 11 The correlation image of A1-A10 factors

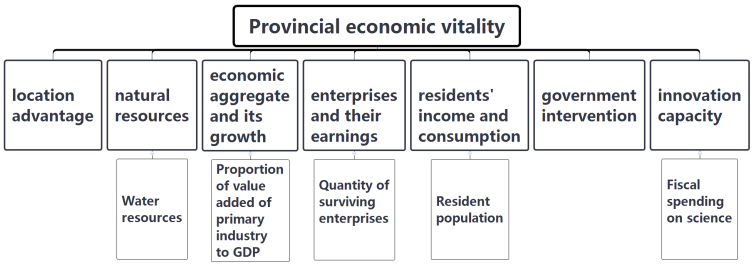

Figure 12 The correlation image of A1-A10 factors

- Construct judgment matrix For a criterion, compare the factors in pairs, and rank them according to their importance.

$$a_{ij} = \frac{1}{a_{ji}} \tag{2}$$

According to Table 4. , the judgment matrix for selected factors is shown as Table 4.

Table 4 Quantized value for importance

Factor i compare to factor j	Quantized value
equally important	1
a little important	3
more important	5
highly important	7
extremely important	9
The intermediate value of two adjacent judgments	2,4,6,8

Table 5 The judgment matrix

	A3	A4	A5	A6	A8
A3	1	3	1	5	6
A4	1/3	1	3	5	6
A5	1	1/3	1	2	2
A6	1/5	1/5	1/2	1	1/2
A8	1/6	1/6	1/2	2	1

- Consistency check

$$CI = \frac{\boldsymbol{\lambda} - n}{n-1}$$

- Random consistency index

$$RI = \frac{CI_1 + CI_2 + .. + CIn}{n-1}$$

The random consistency index RI is only related to the order of the judgment matrix

Table 6 Random consistency index

Order of matrix	1	2	3	4	5	6	7	8	9
RI	0	0	0.58	0.90	1.12	1.24	1.32	1.41	1.45

- Inspection coefficient

$$CR = \frac{CI}{RI}$$

In general, if CR<0.1, the judgment matrix is considered passing the consistency test, otherwise it does not satisfy the consistency. And the CI, CR result for model in our article is:

- Weight The result of weight is:

Table 7 CI, CR Result

CI	CR
0.12097950266067814	0.09756411504893399

Table 8 Indicator weight result

Indicator	A3	A4	A5	A6	A8
Weight	0.39958842	0.3023969	0.16909787	0.05806765	0.07084915

- Fuzzy Comprehensive Evaluation Fuzzy comprehensive evaluation is based on fuzzy mathematics. According to the membership theory of fuzzy mathematics, this comprehensive evaluation ethod transforms qualitative evaluation into quantitative evaluation. In this article, we use the idea of Fuzzy comprehensive evaluation, combing with weight vector get in Hierarchical structure model, establishing the linear model:

$$Y = a_0X_0 + a_2X_1 + a_2X_2 + .. + a_nX_n$$

2.4.4 *Result*

In this article, all the data used will be normalized with min-max method. And the rank result of Attachment 3 is shown as Table 9 and Table 10 for Economic Vatility Result.

	Region	Economic vitality
1	Guangdong	0.772408265429782
2	Sichuan	0.535046236832942
3	Jiangsu	0.493016758932572
4	Shandong	0.488955294814793
5	Henan	0.436299380557526
6	Hunan	0.43461026175968
7	Zhejiang	0.426187806194284
8	Guangxi	0.398440787909183
9	Anhui	0.377930881025595
10	Hubei	0.370743805691688
11	Jiangxi	0.360109879102562
12	Yunnan	0.34417384482531
13	Hebei	0.337658667512227
14	Xizang	0.314014237969484
15	Fujian	0.297976991556249
16	Heilongjiang	0.282114524812096

Table 9: Economic Vatility Result

	Region	Economic vitality
17	Guizhou	0.244697359654887
18	Liaoning	0.23346453779024
19	Chongqing	0.208725196941852
20	Shanghai	0.201226624086537
21	Shaanxi	0.193675619757365
22	Beijing	0.178633579424476
23	Xinjiang	0.177690693556435
24	InnerMonger	0.164571342154985
25	Shanxi	0.159329569009769
26	Jilin	0.148750434136653
27	Gansu	0.118752774275182
28	Tianjin	0.102353631579446
29	Hainan	0.093534310113533
30	Qinghai	0.055677610851501
31	Ningxia	0.019793860258241

Table 10: Economic Vatility Result

2.4.5 *Analysis of the Result*

From the Table 10., we can easily see that Guangdong Province's economic vitality is the highest, which fits our general expression. However , the ranking of Beijing and Shanghai is not very high, which obeys our common sense. The reason that cause the low evaluation is that, the final five factors in our model are Quantity of Labours, Water Source, The quantity of Enterprise,Proportion of scientific research funds and Primary industry, which do not measure the political status, but more about their endowment of resources. And from the angle of resource endowment, the result is reasonable. It illustrates a fact that, some regions with excellent resource endowment but underdeveloped economy can be increased by adjusting through policies, in other words, economic policies is the major determinant of regional economic development.

2.5 Model-based Development recommendations

2.5.1 *Basic Method*

- Using the K-Means++ Algorithm to classify the region, observing the relationship between regions with similar resource endowments.
- According the economic vitality score, and the k-means result observing, raising policy based on economic theory

2.5.2 *Solution Process and Result*

- K-Means++ K-Means++ is the improved method of K-Means.K-means++ selects K cluster centers according to the following idea: Assuming that n initial clustering centers (0<n<K) have been selected out. When selecting the n+1 clustering center, the further away from the current n clustering center is higherly possible to be selected.
- K-Means++ Result

Table 11 The K-Means++ Result

I class	Guangdong, Fujian
II class	Bejing, Tianjin, Shanghai
III class	Jiangsu, Zhejiang
IV class	Hebei, Heilongjiang, Anhui, Jiangxi, Henan, Hunan, Guangxi, Hainan, Sichuan, Qinghai, Ningxia
V class	Inner Mongolia, Liaoning, Jilin, Hubei, Chongqing, Shaanxi
VI class	Shandong
VII class	Shanxi, Guizhou, Yunnan, Tibet, Gansu

- Analyze and Policy recommendations

  1. Optimize resource allocation and reduce economic differences

  The government should continue to increase guidance on industrial transfer, improve the efficiency of resource utilization, and give play to the advantages of agglomeration and synergy of industrial clusters. The eastern and western regions of Guangdong have undertaken the transfer of labor-intensive industries in the Pearl River Delta region, and the industrial structure of the Pearl River Delta region has been upgraded to high-tech industries and emerging industries. Transmit the development dividend of the Pearl River Delta region to the underdeveloped regions in northern Guangdong, share the fruits of economic development, reduce the income gap, and pay more attention to equity while improving the efficiency of economic development. Improve the quality of economic development and economic vitality.

  2. Exert positive externalities and radiate neighboring provinces.

  Looking at the economic vitality index model of problem three, Guangdong's economic vitality index is 0.7724, far exceeding the second-ranked Sichuan province (0.5350). From the results of cluster analysis, Guangdong and Fujian are in the same category. In the process of industrial transfer, along with the flow of capital and talents, a

technology spillover effect occurred, and technical support was provided to neighboring provinces. Guangdong Province should strengthen its economic ties with neighboring provinces, form an industrial network of linkage and cooperation, improve production efficiency and economic vitality.

3. Pay attention to scientific and educational investment and cultivate and attract talents.

In the process of compiling the regional economic vitality index, we found that Guangdong ¡¯s expenditure on science and education accounted for a high proportion of fiscal expenditure, but also only accounted for only 2.4

4. Optimize tax structure and release economic vitality

Taxation is the main source of local finance. Perfecting the taxation system and building a modern fiscal system will help reduce tax losses. At the same time, the reform of the tax structure helps to structurally increase personal disposable income and reduce corporate tax burdens. This will help increase consumer demand and investment demand, promote production, and increase economic vitality.

2.6 Model Evaluation

2.6.1 *Correlation Analyze Test*

In the process of data analysis, we found that delete or change several data records would not change the overall correlation result, and the selected indicators for Analytic Hierarchy Process.

2.6.2 *Analytic Hierarchy Process*

The main factor that affects the result of AHP is the contrast matrix, while adjusting the element in contrast matrix, the result will change a lot. As a result, we have to determine the contrast cautiously.

2.6.3 *K-means++*

While dealing with the K-means++ algorithm, quantities of normalized data takes part in the process, so that several data error will lead a small effect to result

2.7 Strength and Weakness

- Strength

  1. Easy to understand.

  2. In the influencing factor model of economic vitality, qualitative indicators are basically perfect.

  3. Correlation analysis was done when applying AHP to ensure that indicators are independent of each other.

- Weakness

  1. The model of influencing factors of economic vitality does not consider the inter-

action between one primary indicator and the secondary indicator of another primary indicator.

2. Due to the difficulty in obtaining data, the selection of quantitative indicators is not as comprehensive as possible.

3. When analyzing the impact of policies on economic vitality, we conducted empirical tests without controlling other factors. Only by comparing the development status of similar regions, we can infer the impact of policies.

4. The subjective factor is introduced when applying the analytic hierarchy process: the contrast matrix relies on the prior knowledge of experts.

5. The model is highly dependent or sensitive to data.

III. Conclusions

The factors affecting economic vitality cover a wide range, and there is more or less correlation between the factors. In addition, both population and business vitality have a positive correlation with economic vitality. For Guangdong Province, the economic policy changes in 2009 stimulated the economic vitality in the short term, but its impact was not significant in the long run. The total number of companies and multiple economic indicators are relevant. The final result obtained in the linear model for calculating economic vitality shows that the weight of natural resources is large. And the model as a whole is more inclined to resource endowment. In addition, it affected the results of subsequent K-mean++ clustering. Finally, a series of policies proposed for Guangdong Province based on previous models help it sustainably increase economic vitality and enhance regional competitiveness.

IV. Future Work

- Use the Rubin Causal Model to perform a causal analysis of the influencing factors of economic vitality in problem one (using the value of economic vitality obtained in problem III), and remove indicators that do not have a causal relationship.

- You can use the difference-in-difference model to quantitatively evaluate the effect of policies.

- Use the PLS-SEM model to analyze the policy change or the impact mechanism of a policy on economic vitality, that is, to find the path of the policy to influence economic vitality (through the problem one index), and calculate the effect coefficient. In addition, it can predict the policy for the future Impact of economic vitality.

V. References

[1] Dong S B , Yang L P . Linear programming models for estimating the weights in the analytic hierarchy process[J]. Systems Engineering Electronics, 2006, 28(11):1682-1552.

[2] Kwong C K , Bai H . A fuzzy AHP approach to the determination of importance weights of customer requirements in quality function deployment[J]. Journal of Intelligent Manufacturing, 2002, 13(5):367-377.

[3] Mountlake Terrace Government, Economic Vitality Element, visit, 2015,7.

[4] Gong L , Jin C . Fuzzy Comprehensive Evaluation for Carrying Capacity of Regional Water Resources[J]. Water Resources Management, 2009, 23(12):2505-2513.

[5] 何汝群. 珠江-西江经济带城市经济活力评价研究 [D]. 广西师范大学,2019.

[6] Guangzhi B U , Yuwen Z . Grey Fuzzy Comprehensive Evaluation Based on the Theory of Grey Fuzzy Relation[J]. Systems Engineering-theory Practice, 2002.

[7] Bartal Y , Charikar M , Raz D . Approximating min-sum k -clustering in metric spaces[C] Acm Symposium on Theory of Computing. ACM, 2001.

VI. Appendix

Listing 1: The python Source code of Algorithm

```
K-means.py
import csv
import math
import random
from functools import reduce
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
import operator
#   KMeans++算法，优化后的KMeans的算法
class KMeansPP():
def __init__(self,pBasePoints,index,index_data,pN = 7,pPointsCSVName =
    "kmeans_points2.csv",pSetsCSVName =
    "kmeas_sets2.csv",index_diverse='index_diverse2.csv'):
"""
初始化KMeans++算法的构造函数
:param pBasePoints: 所要计算的数据，为点的二维数组
:param pN: 要分成的簇的个数
:param pPointsCSVName: 要写入的点集的CSV文件
:param pSetsCSVName: 要写入的簇的CSV文件
"""
self.arr=[]
```

```
self.item=[]
self.diverse=[]
self.index = index
self.index_data = index_data
self.__N = pN
self.__PCSVName = pPointsCSVName
self.__SCSVName = pSetsCSVName
self.index_diverse=index_diverse
self.__M = len(pBasePoints)#数据的个数
self.__basePoints = pBasePoints
self.__initBaseCenterPoint() #kmeans++算法初始化中心点
#self.__centerPoints = random.sample(self.__basePoints,self.__N) #kmeans算法初始化中心点
self.__initSetsAndNewCenter()#初始化簇集合
self.__kmeans()
self.writeToCSV()
pass

#   初始化N个点
#   这里改进为Kmeans++算法
def __initBaseCenterPoint(self):
self.__centerPoints = []
self.__centerPoints.append(self.__basePoints[random.randint(0, self.__M - 1)])#
    首先初始化一个中心点
while len(self.__centerPoints) < self.__N:#添加中心点直到N个
tempDX = [min([KMeansPP.f_dAB(a,b) for b in self.__centerPoints])**2 for a in
    self.__basePoints]
#D(x)的平方的列表。这一步中的a是遍历了所有的点，然后将a再分别与中心点集合进行遍历求出两点距离求出最短距离
DXSum = sum(tempDX)#kmeans++公式中的分母
DXP = []#轮盘法的值域范围计算，从开始的0到最后的1
for i in range(len(tempDX)):
if i == 0:
DXP.append(tempDX[0]/DXSum)
else:
DXP.append(DXP[i-1]+tempDX[i]/DXSum)
#   因为中心点到其他中心点的最短距离必定是0，+所以必定不会选中中心点
indice =KMeansPP.f_Roulette(DXP)
self.__centerPoints.append(self.__basePoints[indice])
pass

#   初始化新中心点和中心点集合
def __initSetsAndNewCenter(self):
self.__sets={}
self.__sets = {str(k):[] for k in self.__centerPoints}
#print('The __sets 111111111 is:{}'.format(self.__sets))
self.__newCenterPoints = []

#   计算新的中心点
def __countNewCenterPoints(self):
self.__newCenterPoints = []
```

```
#print('The __sets __countNewCenterPoints is:{}'.format(self.__sets))
pDim = len(self.__basePoints[0])
for i in range(self.__N):#重新计算每个簇的中心点
tp = self.__sets[str(self.__centerPoints[i])]#获取簇集合
point = tuple([sum([eval(p)[i] for p in tp])/len(tp) for i in range(pDim)])
#计算新的点。先i遍历维度，然后遍历每个点，对每个点的维度i取出来作为集合再求平均值。实际上就是矩阵的转置
self.__newCenterPoints.append(point)
#print('The self.__newCenterPoint33333333333333333333s
    is:{}'.format(self.__newCenterPoints))
#print('The __sets __countNewCenterPoints2 is:{}'.format(self.__sets))
pass

#   求AB距离
@staticmethod
def f_dAB(A,B):
dim = min(len(A),len(B))
return sum([(int(A[i]) - int(B[i])) ** 2 for i in range(dim)]) ** 0.5

#   轮盘法，返回下标
@staticmethod
def f_Roulette(_list):
tr = random.random()
for i in range(len(_list)):
if i == 0 and _list[i] > tr:
return 0
else:
if _list[i] > tr and _list[i - 1] <= tr:
return i

#   划分集合,kmeans算法
def __kmeans(self):

#   {其他点:[这个点到N个中心点的距离]，……}
t_dList = {str(b):[KMeansPP.f_dAB(a, b) for a in self.__centerPoints] for b in
    self.__basePoints}
#先遍历b为其他点，a为中心点。计算点b到其他所有的中心点的距离
#print('The t_dList is:{}'.format(t_dList))
for k,v in t_dList.items():
#print('The k is:{}'.format(k))
#print('The v is:{}'.format(v))
#print('The self.__centerPoints[v.index(min(v))]
    is:{}'.format(self.__centerPoints[v.index(min(v))]))
#print('The self.__centerPoints is:{}'.format(self.__centerPoints))
self.__sets[str(self.__centerPoints[v.index(min(v))])].append(k)#将距离最小的添加到对应的簇里
#print('The self.__sets is:{}'.format(self.__sets))

self.__countNewCenterPoints()#计算新中心点
#print('%%%%%%%%%%%%%%5The __sets __kmeans1 is:{}'.format(self.__sets))
```

```
#   当各个簇之间有点变动时，就继续
#print('$$$$$$')
j=0
for k,v in self.__sets.items():
self.diverse.append([])
for i in range(len(v)):
#
#print(v[i])
self.arr=self.index_data
self.item=v[i]
index=self.find_all_index()
self.diverse[j].append(self.index[index])
j+=1
print(self.diverse)
self.diverse=[]
if sum([KMeansPP.f_dAB(self.__centerPoints[i],self.__newCenterPoints[i]) for i in
    range(self.__N)]) > 0:
self.__centerPoints = self.__newCenterPoints[:]#把新中心点作为中心点
self.__initSetsAndNewCenter()#重置集合和新中心点清空self.__sets
#print('##############3The __sets is:{}'.format(self.__sets))
#print('The __centerPoints is:{}'.format(self.__centerPoints))
self.k_means()#递归调用


pass

#   k_means算法的对外接口
def k_means(self):
self.__kmeans()
print('############################')
return self.__sets,self.__centerPoints

def find_all_index(self):
#print(item)
for i,a in enumerate(self.arr):
if(operator.eq(a,eval(self.item))):
print(operator.eq(a,eval(self.item)))
return i


def writeToCSV(self):
with open(self.__SCSVName,"w",newline="") as fpc:
fpcWriter = csv.writer(fpc)
fpcWriter.writerow(self.__centerPoints)
maxIndex = max([len(v) for k, v in self.__sets.items()])
j=0
for k,v in self.__sets.items():
self.diverse.append([])
for i in range(len(v)):
```

```
#
#print(v[i])
self.arr=self.index_data
self.item=v[i]
index=self.find_all_index()
self.diverse[j].append(self.index[index])
j+=1



print(self.diverse)

fpcWriter.writerows([[v[i] if len(v) > i else "" for (k, v) in self.__sets.items()]
    for i in range(maxIndex)])
pass

with open(self.__PCSVName,"w",newline="") as fpp:
fppWriter = csv.writer(fpp)
fppWriter.writerows([[self.__basePoints[i*10 + j] if i*10+j < self.__M else "" for j
    in range(10)] for i in range(self.__M//10)])
pass
pass
with open(self.index_diverse,"w",newline="") as fpp:
fppWriter = csv.writer(fpp)
fppWriter.writerows([i] for i in self.diverse)
pass
pass
#读取x为list

'''
#print(df.values)
data = np.array(df)
df=df.to_dict(orient='index')
data_ = np.delete(data, 0, axis=1)
#print(data)
data_ = data_.tolist()
#print(data_)
#print(data)
index = data[:,0].tolist()
#print(index)
'''

file_path = r'./数据.xlsx'
#A='地方科研经费支出占财政支出比例'
#B='教育支出占财政支出比例'
#df = pd.read_excel(file_path,sheetname='Sheet10',hearder=True,usecols=[A,B])
df = pd.read_excel(file_path,sheetname='K-clustering',hearder=True)
#df=df.drop([31])#delete the last row
```

```
data = np.array(df)
df=df.drop('地区',axis=1)
#df=df.apply(lambda x: (x - np.min(x)) / (np.max(x) - np.min(x)))
data_ = np.array(df).tolist()
#print(data_)
#print(data)

index = data[:,0].tolist()
print(df)
print('$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$')
max_min_scaler = lambda x : (x-np.min(x))/(np.max(x)-np.min(x))

#df['A2'].apply(max_min_scaler)
df['A3'].apply(max_min_scaler)
df['A4'].apply(max_min_scaler)
df['A5'].apply(max_min_scaler)
df['A6'].apply(max_min_scaler)
df['A8'].apply(max_min_scaler)
df['B1'].apply(max_min_scaler)
#print(data)
#print(index) #归一化
KMeansPP(data_,index,data_)

\newpage
AHP.py
# -*- coding: utf-8 -*-
import numpy as np
import pandas as pd
import xlrd
import openpyxl

'''
def write_excel_xlsx(path, sheet_name, value):
index = len(value)
workbook = openpyxl.Workbook()
sheet = workbook.active
sheet.title = sheet_name
for i in range(0, index):
for j in range(0, len(value[i])):
sheet.cell(row=i+1, column=j+1, value=str(value[i][j]))
workbook.save(path)
print("xlsx格式表格写入数据成功！")


def read_excel_xlsx(path, sheet_name):
workbook = openpyxl.load_workbook(path)
# sheet = wb.get_sheet_by_name(sheet_name)这种方式已经弃用，不建议使用
sheet = workbook[sheet_name]
for row in sheet.rows:
```

```
for cell in row:
print(cell.value, "\t", end="")
print()


book_name_xlsx = 'xlsx格式测试工作簿.xlsx'

sheet_name_xlsx = 'xlsx格式测试表'

value3 = [["姓名", "性别", "年龄", "城市", "职业"],
["111", "女", "66", "石家庄", "运维工程师"],
["222", "男", "55", "南京", "饭店老板"],
["333", "女", "27", "苏州", "保安"],]


write_excel_xlsx(book_name_xlsx, sheet_name_xlsx, value3)
read_excel_xlsx(book_name_xlsx, sheet_name_xlsx)

'''




A = np.array([[1,3,1, 5,6],
[1/3,1,3,5,6],
[1,1/3,1,2,2],
[1/5,1/5,1/2,1,1/2],
[1/6,1/6,1/2,2,1]]) #对比矩阵

# 文件路径
file_path = r'./test.xlsx'

#A = df.values
def AHP_weight(A):
m,n = len(A),len(A[0])
RI = [0, 0, 0.58, 0.90, 1.12, 1.24, 1.32, 1.41, 1.45, 1.49, 1.51]
R = np.linalg.matrix_rank(A)
V, D = np.linalg.eig(A) #V特征值， D特征向量
lambda_max = np.max(list(V)) #最大特征值
index = list(V).index(lambda_max)
C = D[:,index] #对应特征向量
CI = (lambda_max-n)/(n-1)
CR = CI/RI[n]
weight = C/np.sum(C)#权重向量
if CR<0.10:
print("CI=",CI)
print("CR=",CR)
print('对比矩阵A通过一致性检验，各向量权重向量Q为：{}'.format(weight))
else:
```

```
print("对比矩阵A未通过一致性检验，需对对比矩阵A重新构造")
#A =get_comparison_matrix(A)

AHP_weight(A)

def get_comparison_matrix(A):
m,n = len(A),len(A[0])
for w in range(m):
for p in range(w):
A[w,p]=1/A[p,w]
return A
def AHP_1(A1):#一级指标对比矩阵
AHP_weight1 = AHP_wehight(A1)
return AHP_weight_1

#c = np.insert(a, 0, values=b, axis=0)
#d = np.insert(a, 0, values=b, axis=1)

def AHP_2(A1,A2):#二级指标对比矩阵
m,n = len(A1),len(A1[0])
A2_list = []
for i in range(m):
A2_list.append(AHP_weight(A2))

\newpage
plot.py
# -*- coding: utf-8 -*-'
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
#

filepath= r'./数据(1).xlsx'
df = pd.read_excel(filepath,sheetname='企业利润总额增长率',hearder=True)
print(df.index)
#df = df.cumsum()
list_year=df['Year']
print(list_year)
df=df.drop('Year',axis=1)

print(df)
df.index = list_year
plt.figure(figsize=(50,50));
df.plot(marker='*',title='Corporate profit growth rate');
plt.legend(loc='best')



A='Guangdong'
```

```
filepath= r'./亚太赛数据(1).xlsx'
df = pd.read_excel(filepath,sheetname='企业利润总额增长率',hearder=True,usecols=[A])

print(df)
df.index = list_year
plt.figure(figsize=(50,50));
df.plot(marker='*',title='Corporate profit growth rate');
plt.legend(loc='best')


filepath= r'./数据(1).xlsx'
df = pd.read_excel(filepath,sheetname='GDP',hearder=True)
print(df.index)
list_year=df['Year']
df = df.cumsum()
df=df.drop('Year',axis=1)
df=df.drop('nationwide',axis=1)
df.index = list_year
print('The df index is :{}'.format(df.index))
print(list_year)
plt.figure(figsize=(50,50));
df.plot(marker='*',title='GDP');
#plt.legend(loc='right')



filepath= r'./数据(1).xlsx'
df = pd.read_excel(filepath,sheetname='GDP_growthrate',hearder=True)
print(df.index)
df = df.cumsum()
list_year=df['Year']
print(list_year)
df=df.drop('Year',axis=1)
df=df.drop('nationwide',axis=1)
print(df)
df.index = list_year
plt.figure(figsize=(50,50));
df.plot(marker='*',title='GDP Growth Rate');
plt.legend(loc='right')


filepath= r'./数据(1).xlsx'
df = pd.read_excel(filepath,sheetname='GDP_growthrate',hearder=True,usecols=[A])
print(df.index)
df = df.cumsum()
print(df)
df.index = list_year
plt.figure(figsize=(50,50));
df.plot(marker='*',title='GDP Growth Rate');
```

```
plt.legend(loc='right')




filepath= r'./数据(1).xlsx'
df = pd.read_excel(filepath,sheetname='PopulationTrend',hearder=True)
print(df.index)
#df = df.cumsum()
list_year=df['Year']
print(list_year)
df=df.drop('Year',axis=1)
df=df.drop('nationwide',axis=1)

print(df)
df.index = list_year
plt.figure(figsize=(50,50));
df.plot(marker='*',title='Population Growth Rate');
plt.legend(loc='right')



filepath= r'./数据(1).xlsx'
df = pd.read_excel(filepath,sheetname='PopulationTrend',hearder=True,usecols=[A])
print(df.index)
#df = df.cumsum()

print(df)
df.index = list_year
plt.figure(figsize=(50,50));
df.plot(marker='*',title='Population Growth Rate');
plt.legend(loc='right')

\newpage
Rank.py
# -*- coding: utf-8 -*-
import csv
import math
import random
from functools import reduce
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
import operator
file_path = r'./数据.xlsx'
#A='地方科研经费支出占财政支出比例'
#B='教育支出占财政支出比例'
#df = pd.read_excel(file_path,sheetname='Sheet10',hearder=True,usecols=[A,B])
```

```
df = pd.read_excel(file_path,sheetname='K-clustering',hearder=True)
#df=df.drop([31])#delete the last row

data = np.array(df)
df=df.drop('地区',axis=1)
#df=df.apply(lambda x: (x - np.min(x)) / (np.max(x) - np.min(x)))
data_ = np.array(df).tolist()
#print(data_)
#print(data)

index = data[:,0].tolist()
print(df)
print('$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$')
max_min_scaler = lambda x : (x-np.min(x))/(np.max(x)-np.min(x))
print(df['A3'])
#df['A2'].apply(max_min_scaler)
df['A3'].apply(max_min_scaler)
df['A4'].apply(max_min_scaler)
df['A5'].apply(max_min_scaler)
df['A6'].apply(max_min_scaler)
df['A8'].apply(max_min_scaler)
A3 = np.array(df['A3'])
A4 = np.array(df['A4'])
A5 = np.array(df['A5'])
A6 = np.array(df['A6'])
A8 = np.array(df['A8'])
print(A3.size)

def max_min(A):
max = np.max(A)
min = np.min(A)
for i in range(A.size):
A[i] = (A[i]-min)/(max-min)
return A

A3 = max_min(A3)
A4 = max_min(A4)
A5 = max_min(A5)
A6 = max_min(A6)
A8 = max_min(A8)




#A = 0.39958842* A3+0.3023969*A4+0.509787*A5+0.05806765*A6+0.07084915*A8
A = 0.39958842* A3+0.3023969*A4+0.16909787*A5+0.05806765*A6+0.07084915*A8
print(A)
B =np.argsort(A)
B+=2
```

```
print(B)
with open('./Enterprise.xlsx',"w",newline="") as fpp:
fppWriter = csv.writer(fpp)
fppWriter.writerows([i] for i in A)
pass

with open('./Enterprise_rank.xlsx',"w",newline="") as fpp:
fppWriter = csv.writer(fpp)
fppWriter.writerows([i] for i in B)
pass
```