Team Number:	90046
Problem Chosen:	B

2019 APMCM summary sheet

Keywords: Keywords1 Keywords2 Keywords3

Contents

I.

Nowadays, leaders and authorities of both countries or regions of a country are seeking ways to improve and boost their regional comprehensive competitiveness. However, the regional comprehensive competitiveness is determined by multiple factors such as economic vitality, living condition and culture depth, which makes it to be a hot topic.

The regional economic vitality is one of the most important part of regional comprehensive competitiveness. Recently, in order to boost the economic vitality, some regions have implemented series of preferential policies to spur the economy vitality. For example, Soochow reduces the investment attraction approval steps and supports the start-ups by providing financial assistance and Wuhan launches policy to lower the settlement threshold and living cost by up to 20% to attract talents. However, due to the existing regional disparity, these policies effect differently. How to seize the key factors and effectively improve the regional economic vitality is really worth studying.

In an effort to research the factors influencing economic vitality and how to improve the regional economic vitality, we have acquired some sources along with the given data to build suitable model and solve the problems.

- To begin our work, we take Eastern region as an example. We build a relational model involving influencing factors of economic vitality such as the total number of registered enterprises, the amount of residents, regional GDP and so on. Then we study the program of action to improve the regional economic vitality and analyse the effects on the change of regional economic vitality from the perspective of changing trend of population and enterprise vitality.
- We continue to use the relational model built in question *1* and the economic vitality measurement index. Then we analyse the short-term and long-term effects on the economic vitality of Eastern region using prediction model when economic policy transformation is implemented. For the population forecast, we adopt the Logistic model. For the quantity of enterprises, we use gray forecast model *GM (1, 1)*. The household consumption, gross domestic production and so on can be classified as economic categories. Therefore, we use time series prediction model based on exponential smoothing to address them.
- Evaluating economic vitality is a complicated issue. Therefore, we select a system of 10 indexes that to a large extent can act as measurement of economic vitality and build a measurement model using PCA. Then we rank the economic vitality of cities in *Attachment 3*.

- According to the conclusions for Problems *1-3*, we provide an economic development proposal for Eastern region discussed in Problem *2* so that the economic vitality in this region presents the benign sustainable development and its competitiveness is stronger.

II.

2.1 STC89C52

- We assume that the given data and surveyed data are valid and authentic.
- We assume that the total number of enterprises in an economic region is the addition of each province in that region.
- We assume that the cancellation rate of enterprises in a region is represented by the average rate of each city in that region.
- We assume that the selection of factors or indexes that may influence the economic vitality is objective.
- We assume that the error or deviation in the data process is within the acceptable range.

2.2 CC1101

Serial Number	Symbol	Description
1	A_{re}	The amount of residents in Eastern region within 1 year.
2	A_{en}	The amount of enterprises in Eastern region within 1 year.
3	I_{CPI}	The consumer price index in Eastern region within 1 year.
4	I_{GDP}	The gross domestic product in Eastern region within 1 year.
5	I_{FAI}	The fixed-asset investment in Eastern region within 1 year.
6	I_{HC}	The household consumption in Eastern region within 1 year.
7	$A_{10\times6}$	The factor matrix.

8	$R_{6\times 6}$	The correlation coefficient matrix.
9	Z	The comprehensive appraise index.
11		4.88
12		9.71
13		13.97
14		46.68
15		3.03

III.

3.1 MCU*1*

We abstract 6 factors that may affect the economic vitality in Eastern region from the given and surveyed data as follows.

- The amount of residents in Eastern region from year 2009 to year 2018.
- The amount of enterprises in Eastern region from year 2009 to year 2018.
- CPI (consumer price index) in Eastern region from year 2009 to year 2018.
- GDP (gross domestic product) in Eastern region from year 2009 to year 2018.
- FAI (fixed-asset investment) in Eastern region from year 2009 to year 2018.
- Household consumption in Eastern region from year 2009 to year 2018.

3.1.1 *The Amount of Residents*

The amount of residents from 2009 to 2018 in each province of Eastern region and the total number are shown in Table 2. The data is surveyed from *National Bureau of Statistics* (unit: 10000)

We also draw the bar chart of the total number changing from 2009 to 2018 in Figure 1.

3.1.2 *The Amount of Enterprises*

Cancellation rate can be defined as the ratio of cancelled enterprises to the whole amount of enterprises. Through the cancellation rate we can easily obtain the data of cancelled enterprises. Therefore, based on the amount and increment of enterprises in *Attachments*, the number of surviving enterprises in each year from 2009 to 2018 can be obtained. The amount of surviving enterprises in each year and changing trend from 2009 to 2018 are shown in Figure 2.

Table 2 The amount of residents from 2009 to 2018 in each province of Eastern region and the total number.

Year	2018	2017	2016	2015	2014	2013	2012	2011	2010	2009
Beijing	2154	2171	2173	2171	2152	2115	2069	2019	1962	1860
Tianjin	1560	1557	1562	1547	1517	1472	1413	1355	1299	1228
Hebei	7556	7520	7470	7425	7384	7333	7288	7241	7194	7034
Shandong	10047	10006	9947	9847	9789	9733	9685	9637	9588	9470
Jiangsu	8051	8029	7999	7976	7960	7939	7920	7899	7869	7810
Shanghai	2424	2418	2420	2415	2426	2415	2380	2347	2303	2210
Zhejiang	5737	5657	5590	5539	5508	5498	5477	5463	5447	5276
Fujian	3941	3911	3874	3839	3806	3774	3748	3720	3693	3666
Guangdong	11346	11169	10999	10849	10724	10644	10594	10505	10441	10130
Hainan	934	926	917	911	903	895	887	877	869	864
Total	53750	53364	52951	52519	52169	51818	51461	51063	50665	49548

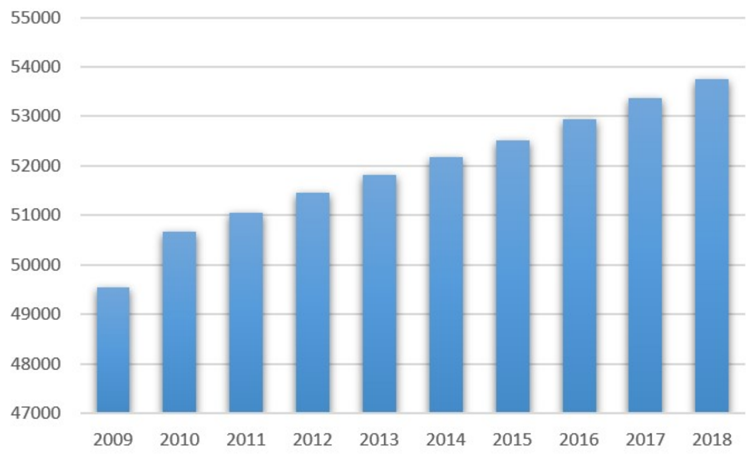

Figure 1 The bar chart of the total number changing from 2009 to 2018

3.1.3 *Consumer Price Index*

Consumer price index from 2009 to 2018 in each province of Eastern region and the average number are shown in Table 3. The data is surveyed from *National Bureau of Statistics* (last year: 100)

We also draw the line chart of the average number changing from 2009 to 2018 in Figure 3.

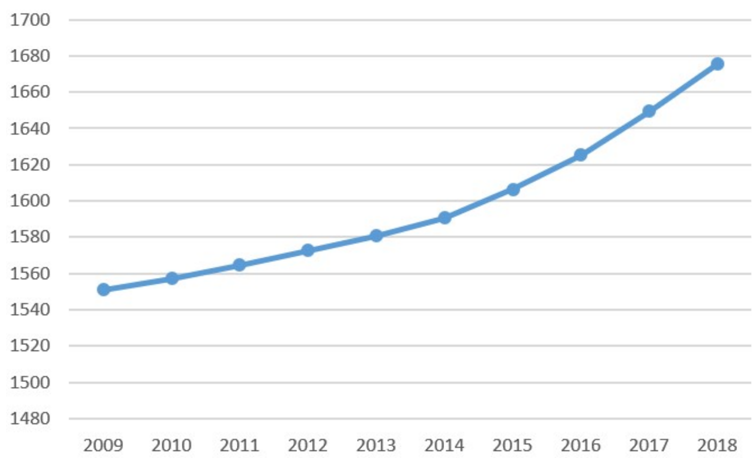

Figure 2 The line chart of the number of enterprise changing from 2009 to 2018

Table 3 Consumer price index from 2009 to 2018 in each province of Eastern region and the average number.

Year	2018	2017	2016	2015	2014	2013	2012	2011	2010	2009
Beijing	102.50	101.90	101.40	101.80	101.60	103.30	103.30	105.60	102.40	98.50
Tianjin	102.00	102.10	102.10	101.70	101.90	103.10	102.70	104.90	103.50	99.00
Hebei	102.40	101.70	101.50	100.90	101.70	103.00	102.60	105.70	103.10	99.30
Shandong	102.50	101.50	102.10	101.20	101.90	102.20	102.10	105.00	102.90	100.00
Jiangsu	102.30	101.70	102.30	101.70	102.20	102.30	102.60	105.30	103.80	99.60
Shanghai	101.60	101.70	103.20	102.40	102.70	102.30	102.80	105.20	103.10	99.60
Zhejiang	102.30	102.10	101.90	101.40	102.10	102.30	102.20	105.40	103.80	98.50
Fujian	101.50	101.20	101.70	101.70	102.00	102.50	102.40	105.30	103.20	98.20
Guangdong	102.20	101.50	102.30	101.50	102.30	102.50	102.80	105.30	103.10	97.70
Hainan	102.50	102.80	102.80	101.00	102.40	102.80	103.20	106.10	104.80	99.30
Average	102.18	101.82	102.13	101.53	102.08	102.63	102.67	105.38	103.37	98.97

3.1.4 *Gross Domestic Product*

Gross Domestic Product from 2009 to 2018 in each province of Eastern region and the total number are shown in Table 4. The data is surveyed from *National Bureau of Statistics* (unit: 100 million yuan)

We also draw the bar chart of the total number changing from 2009 to 2018 in Figure 3.

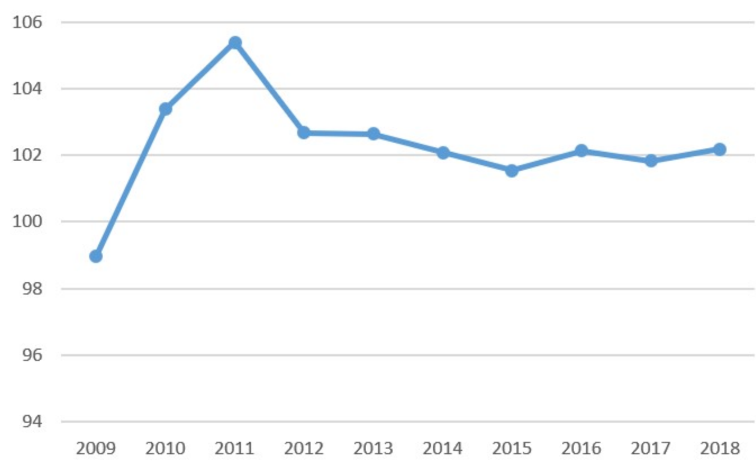

Figure 3 The line chart of the avarage number changing from 2009 to 2018

Table 4 Gross Domestic Product from 2009 to 2018 in each province of Eastern region and the total number.

Year	2018	2017	2016	2015	2014	2013	2012	2011	2010	2009
Beijing	30320	28015	25669	23015	21331	19801	17879	16252	14114	12153
Tianjin	18810	18550	17885	16538	15727	14442	12894	11307	9224	7522
Hebei	36010	34016	32070	29806	29421	28443	26575	24516	20394	17235
Shandong	76470	72634	68024	63002	59427	55230	50013	45362	39170	33897
Jiangsu	92595	85870	77388	70116	65088	59753	54058	49110	41425	34457
Shanghai	32680	30633	28179	25123	23568	21818	20182	19196	17166	15046
Zhejiang	56197	51768	47251	42886	40173	37757	34665	32319	27722	22990
Fujian	35804	32182	28811	25980	24056	21868	19702	17560	14737	12237
Guangdong	97278	89705	80855	72813	67810	62475	57068	53210	46013	39483
Hainan	4832	4463	4053	3703	3501	3178	2856	2523	2065	1654
Total	480996	447835	410186	372983	350101	324765	295892	271355	232031	196674

3.1.5 *Fixed-Asset Investment*

Fixed-Asset Investment from 2009 to 2018 in each province of Eastern region and the total number are shown in Table 5. The data of 2018 is predicted with data of previous years with an error within 5%. The data is surveyed from *National Bureau of Statistics* (unit: 100 million yuan)

We also draw the line chart of the total number changing from 2009 to 2018 in Figure 5.

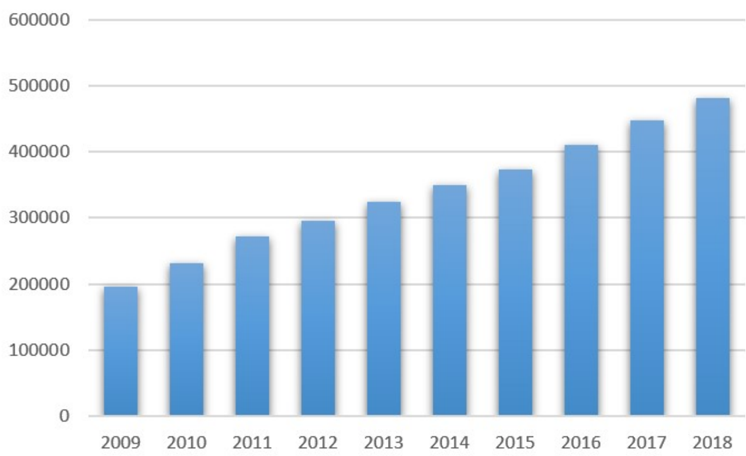

Figure 4 The bar chart of the total number changing from 2009 to 2018

Table 5 Fixed-Asset Investment from 2009 to 2018 in each province of Eastern region and the total number.

Year	2018	2017	2016	2015	2014	2013	2012	2011	2010	2009
Beijing	8862	8370	7944	7496	6924	6847	6112	5579	5403	4617
Tianjin	13881	11289	12779	11832	10518	9130	7935	7068	6278	4738
Hebei	37070	33407	31750	29448	26672	23194	19661	16389	15083	12270
Shandong	61480	55203	53323	48312	42496	36789	31256	26750	23281	19035
Jiangsu	58598	53277	49663	46247	41939	36373	30854	26693	23184	18950
Shanghai	7258	7247	6756	6353	6016	5648	5118	4962	5109	5044
Zhejiang	35233	31696	30276	27323	24263	20782	17649	14185	12376	10742
Fujian	35233	26416	23237	21301	18178	15327	12440	9911	8199	6231
Guangdong	39358	37762	33304	30343	26294	22308	18752	17069	15624	12933
Hainan	4720	4244	3890	3451	3112	2698	2145	1657	1317	988
Total	301693	268911	252923	232107	206412	179098	151923	130263	115854	95548

3.1.6 *Household Consumption*

Household Consumption from 2009 to 2018 in each province of Eastern region and the total number are shown in Table 6. The data of 2018 is predicted with data of previous years with an error within 5%. The data is surveyed from *National Bureau of Statistics* (unit: yuan)

We also draw the bar chart of the total number changing from 2009 to 2018 in Figure 6.

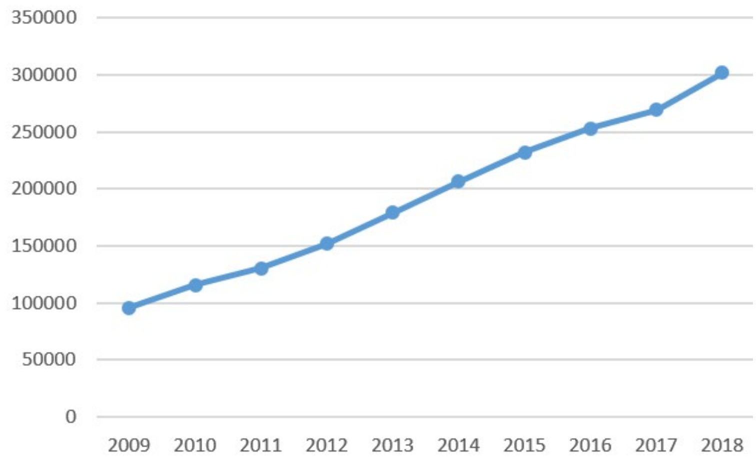

Figure 5 The line chart of the total number changing from 2009 to 2018

Table 6 Fixed-Asset Investment from 2009 to 2018 in each province of Eastern region and the total number.

Year	2018	2017	2016	2015	2014	2013	2012	2011	2010	2009
Beijing	53710	52912	48883	39200	36057	33337	30350	27760	24982	22023
Tianjin	41563	38975	36257	32595	28492	26261	22984	20624	17852	15200
Hebei	16502	15893	14328	12829	12171	11557	10749	9551	8057	7193
Shandong	28998	28353	25860	20684	19184	16728	15095	13524	11606	10494
Jiangsu	42541	39796	35875	31682	28316	23585	19452	17167	14035	11993
Shanghai	55861	53617	49617	45816	43007	39223	36893	35439	32271	26582
Zhejiang	35487	33851	30743	28712	26885	24771	22845	21346	18274	15867
Fujian	26643	25969	23355	20828	19099	17115	16144	14958	13187	11336
Guangdong	32444	30762	28495	26365	24582	23739	21823	19578	17211	15243
Hainan	21747	20939	18431	17019	12915	11712	10634	9238	7553	6695
Total	355496	341067	311844	275730	250708	228028	206969	189185	165028	142626

3.2 *1*

3.2.1

According to question *1* and the results of data process, we adopt the method PCA to build the relational model to measure the economic vitality and influencing factors in Eastern region. We give the details of model construction by steps using PCA as follows.

1) Firstly, we use the results of 6 factor variables from data process to create the

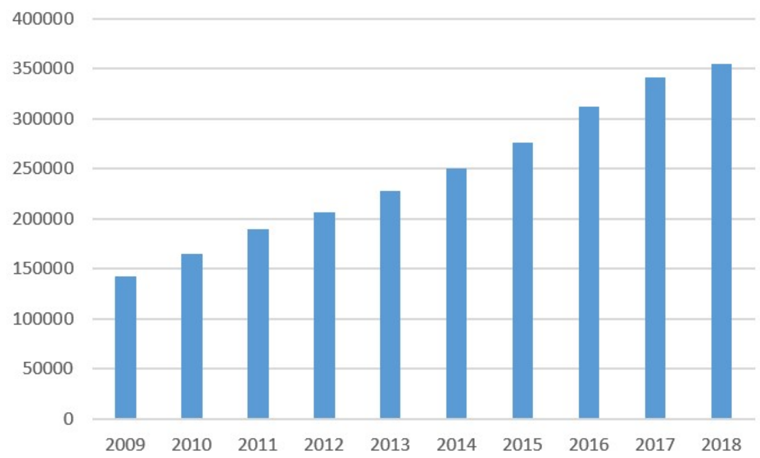

Figure 6 The bar chart of the total number changing from 2009 to 2018

factor matrix $A_{10\times6}$ which is defined in the following equation.

$$A = (a_{ij})_{10\times6} \tag{1}$$

where the integer i ranges from 1 to 10 representing year from 2009 to 2018 and the integer j ranges from 1 to 6 representing factors *household consumption, FAI, GDP, CPI, the amount of residents and the amount of enterprises* respectively. Therefore, the factor variables can be symbolized as x_j.

2) Secondly, we normalize each element value in the factor matrix using the following equation.

$$\tilde{a}_{ij} = \frac{a_{ij} - \mu_j}{s_j} \tag{2}$$

$$\mu_j = \frac{1}{10}\sum_{i=1}^{10} a_{ij} \tag{3}$$

$$s_j = \sqrt{\frac{1}{9}\sum_{i=1}^{10}(a_{ij} - \mu_j)^2} \tag{4}$$

(3) and (4) are the sample mean value and the sample standard deviation of jth factor respectively. Accordingly, the factor variables x_j can be normalized as $\tilde{x}_j$ with the following equation.

$$\tilde{x}_j = \frac{x_j - \mu_j}{s_j} \tag{5}$$

3) Thirdly, we formulate the correlation coefficient matrix R with the following equation.

$$R = (r_{j_1 j_2})_{6\times6} \tag{6}$$

$$r_{j_1 j_2} = \frac{\sum_{i=1}^{10} \tilde{a}_{ij_1} \cdot \tilde{a}_{ij_2}}{9} \tag{7}$$

where j_1 and j_2, with different value from 1 to 6, are the same as j. $r_{j_1j_2}$ is the correlation coefficient of j_1th factor and j_2th factor and also the element of correlation coefficient matrix.

4) Fourthly, we calculate the feature values λ_j ($\lambda_j \geq 0$) and corresponding normalized feature vectors $[u_{1j}, u_{2j}, \ldots, u_{6j}]^T$ of correlation coefficient matrix R. Then the principal components formed by 6 normalized feature vectors are formulated as follows.

$$\begin{cases} y_1 = u_{11}\tilde{x}_1 + u_{21}\tilde{x}_2 + \ldots + u_{61}\tilde{x}_6 \\ y_2 = u_{12}\tilde{x}_1 + u_{22}\tilde{x}_2 + \ldots + u_{62}\tilde{x}_6 \\ y_3 = u_{13}\tilde{x}_1 + u_{23}\tilde{x}_2 + \ldots + u_{63}\tilde{x}_6 \\ y_4 = u_{14}\tilde{x}_1 + u_{24}\tilde{x}_2 + \ldots + u_{64}\tilde{x}_6 \\ y_5 = u_{15}\tilde{x}_1 + u_{25}\tilde{x}_2 + \ldots + u_{65}\tilde{x}_6 \\ y_6 = u_{16}\tilde{x}_1 + u_{26}\tilde{x}_2 + \ldots + u_{66}\tilde{x}_6 \end{cases} \tag{8}$$

where y_j denotes the jth principal component.

5) Fifthly, we select p out of 6 principal components to measure their comprehensive appraisal value. Contribution rate b_j of each feature value λ_j is also the contribution rate of corresponding principal component y_j, which is formulated as follows.

$$b_j = \frac{\lambda_j}{\sum_{j=1}^{6} \lambda_j} \tag{9}$$

Then the accumulation contribution rate of p principal components can be formulated as follows.

$$\partial_p = \frac{\sum_{j=1}^{p} \lambda_j}{\sum_{j=1}^{6} \lambda_j} \tag{10}$$

If ∂_p approximates 1, it means that we can choose these p principal components to substitute for the original 6 principal components. The comprehensive appraise of principal components is then simplified.

6) Lastly, the result of comprehensive appraise Z can be acquired through computation. The following equation gives the formulation of z.

$$Z = \sum_{j=1}^{p} b_j y_j \tag{11}$$

3.2.2

Based on the model constructed using PCA, we apply Matlab to produce the model results. Primarily we select all of the 6 normalized factor variables deviated from the original factor variables and obtain the 6 principal components. Then we figure out the weight of each normalized factor variable to the 6 principal components through computation and the numbers in Table 7 show them clearly.

Table 7 Weight of each normalized factor variable to the principal components.

Principal Component	$\tilde{x}_1$	$\tilde{x}_2$	$\tilde{x}_3$	$\tilde{x}_4$	$\tilde{x}_5$	$\tilde{x}_6$
y_1	0.4502	-0.0182	-0.0287	-0.7079	0.4598	-0.2893
y_2	0.4491	-0.0504	-0.2367	0.6360	0.5734	0.0806
y_3	0.4503	0.0272	-0.1160	-0.2181	-0.2708	0.8137
y_4	-0.0052	0.9875	0.1112	0.0244	0.1058	0.0275
y_5	0.4447	0.1262	-0.4518	0.1144	-0.5834	-0.4782
y_6	0.4417	-0.0735	0.8445	0.1823	-0.1869	-0.1349

Afterwards, the contribution rates and the accumulation contribution rates of each principal component to the comprehensive appraise index Z are shown in Table 8

Table 8 Contribution rates and accumulation contribution rates of each principal component to the comprehensive appraise index.

Principal Component	**Contribution Rate (%)**	**Accumulation Contribution Rate (%)**
y_1	81.9304	81.9304
y_2	17.0794	99.0098
y_3	0.8233	99.8331
y_4	0.0805	99.9136
y_5	0.0581	99.9717
y_6	0.0283	100.0000

Based on the results shown in Table 7 and Table 8, we finally select the first 3 principal components instead of the original 6 principal components according to their contribution rates and accumulation contribution rates and also we add up the weights of the 6 normalized factor variables to the first 3 principal components. The final result, also the relational model of influencing factors, is formulated as follows.

$$Z = 0.4824\tilde{x}_1 - 0.0213\tilde{x}_2 - 0.0735\tilde{x}_3 - 0.4889\tilde{x}_4 + 0.4524\tilde{x}_5 - 0.0497\tilde{x}_6 \qquad (12)$$

where $\tilde{x}_1$, $\tilde{x}_2$, $\tilde{x}_3$, $\tilde{x}_4$, $\tilde{x}_5$ and $\tilde{x}_6$ denote to household consumption, FAI, GDP, CPI, the amount of residents and the amount of enterprises respectively.

3.3

3.3.1 *1*

In a short term, the growth of residents will benefit the economy in terms of labour resource and skilled talents. However, the restriction from limited living room and finite natural resources will constrain the resident growth. Therefore, we propose to use Logistic Model to predict the population growth of residents and give the detailed steps.

1) The population is a function of year i and can be formed as $n(i)$. The growth rate r is a function of population and can then be described as a decrement function $r(n)$. We suppose $r(n)$ is a linear function and is formulated as follows.

$$r(n) = r - sn \tag{13}$$

where s is a coefficient. Assumption is made that the maximum population which Eastern region can hold is n_{max}. That means when the population reaches the maximum the growth rate reduces to nearly 0.

2) Based on the above model assumption, we further formulate the growth rate as follows.

$$r(n) = r(1 - \frac{n}{n_{max}}) \tag{14}$$

Logistic model is built and shown as follows.

$$\begin{cases} \frac{dn}{di} = n \cdot r(1 - \frac{n}{n_{\max}}) \\ n(i_0) = n_0 \end{cases} \tag{15}$$

where i_0 and n_0 are the initial value of year and population respectively. The solution to Equations (15) is as follows.

$$n(i) = \frac{n_{\max}}{1 + (\frac{n_{\max}}{n_0} - 1)e^{-r(i-i_0)}} \tag{16}$$

3) Based on the solution given above, we import data in *Excel* to *Matlab* and get the prediction results of resident population in next years.

3.3.2 *2*

Based on the background, we plan to apply commonly used *GM(1,1)* to predict the quantity of enterprises in next years. We give the steps of model construction.

1) Firstly, we test the data of enterprise quantity. The primary data sequence is $m^{(0)}$ and

$$m^{(0)} = [m^{(0)}(1), m^{(0)}(2), m^{(0)}(3), ..., m^{(0)}(i)] \tag{17}$$

where i is year. The sequence ratio $\lambda(k)$ is define as follows.

$$\lambda(k) = \frac{m^{(0)}(k-1)}{m^{(0)}(k)} \tag{18}$$

where k equals i and starts from 2. The range of sequence ratio $(1.0040, 1.0158)$ is obtained through *Matlab* and justifies the use of *GM(1,1)* by covering the following range.

$$\Theta = (e^{-\frac{2}{\pi}}, e^{\frac{1}{6}}) = (0.834, 1.181) \tag{19}$$

2) Secondly, we obtain $m^{(1)}$ through the following equation.

$$m^{(1)} = [m^{(0)}(1), \sum_{k=1}^{2} m^{(0)}(k), \sum_{k=1}^{3} m^{(0)}(k), ..., \sum_{k=1}^{10} m^{(0)}(k)] \tag{20}$$

Then the data matrix B and the data column vector Y are formulated as

$$B = \begin{bmatrix} -\frac{1}{2}[m^{(1)}(1) + m^{(1)}(2)] & 1 \\ -\frac{1}{2}[m^{(1)}(2) + m^{(1)}(3)] & 1 \\ \vdots & \vdots \\ -\frac{1}{2}[m^{(1)}(9) + m^{(1)}(10)] & 1 \end{bmatrix}, Y = \begin{bmatrix} m^{(0)}(2) \\ m^{(0)}(3) \\ \vdots \\ m^{(0)}(10) \end{bmatrix} \tag{21}$$

3) Thirdly, *GM(1,1)* is built and shown as follows.

$$\begin{cases} \frac{dm^{(1)}}{di} + \hat{a}m^{(1)} = \hat{b} \\ \hat{u} = \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = (B^T B)^{-1} B^T Y \end{cases} \tag{22}$$

The solution is

$$\hat{m}^{(1)}(k+1) = [m^{(0)}(1) - \frac{\hat{b}}{\hat{a}}]e^{-\hat{a}k} + \frac{\hat{b}}{\hat{a}} \tag{23}$$

4) Fourthly, We import the data of enterprises to *Matlab* and get the fitting results shown in Figure **??**.

5) Finally, we test our model using sequence ratio deviation $p(k)$. We calculate the sequence ratio $\lambda(k)$ with reference data $m^{(0)}(k-1)$ and $m^{(0)}(k)$. Therefore, the sequence ratio deviation is formulated as follows.

$$p(k) = 1 - (\frac{1 - 0.5a}{1 + 0.5a}) \cdot \lambda(k) \tag{24}$$

The final sequence ratio deviation is 0.18261 and less than 0.2. The prediction results meet the requirement.

3.3.3 *she'ji*

Based on previous analysis, we classify CPI, GDP, FAI and household consumption as economic category , compared with the other two factors. Commonly, the effect of history data on prediction usually fades away as time proceeds. Therefore, a better and more practical way to predict the value of these factors is applying weight which decreases with time to history data in the prediction work. We then choose to apply second-exponential smoothing model and discuss the details in model construction by steps.

1) Firstly, the exponential smoothing model is formulated as

$$S_i^{(1)} = \alpha e_i + (1-\alpha)e_{i-1}^{(1)} = e_{i-1}^{(1)} + \alpha(e_i - S_{i-1}^{(1)}) \tag{25}$$

where e_i and α are time sequence of factors and weighted coefficient respectively. The integer i ranges from 1 to 10 representing year from 2009 to 2018 and α is from 0 to 1. The iteration of moving average is

$$\mu_i^{(1)} = \mu_{i-1}^{(1)} + \frac{e_i - e_{i-N}}{N} \tag{26}$$

where α equals $\frac{1}{N}$. We then choose $\mu_{i-1}^{(1)}$ to be the best estimation of e_{i-N} and get the following equation.

$$\mu_i^{(1)} = \mu_{i-1}^{(1)} + \frac{e_i - \mu_{i-1}^{(1)}}{N} = \frac{e_i}{N} + (1 - \frac{1}{N})\mu_{i-1}^{(1)} \tag{27}$$

We replace $\mu_i^{(1)}$ with S_i and get the following equation.

$$S_i^{(1)} = \alpha e_i + (1-\alpha)S_{i-1}^{(1)} \tag{28}$$

That means $S_i^{(1)}$ is the weighted average of all history data with the weighted coefficient to be α, $\alpha(1-\alpha)$, $\alpha(1-\alpha)^2$ and so on. Obviously, the following equation is established.

$$\sum_{k=0}^{\infty} \alpha(1-\alpha)^k = \frac{\alpha}{1-(1-\alpha)} = 1 \tag{29}$$

Then the prediction model can be further formulated as

$$\hat{e}_{i+1} = S_i^{(1)} = \alpha e_i + (1-\alpha)\hat{e}_i \tag{30}$$

2) Secondly, we transform the model equation to

$$\hat{e}_{i+1} = \hat{e}_i + \alpha(e_i - \hat{e}_i) \tag{31}$$

We can see from the above equation that the selection of weighted coefficient α directly affects the prediction results. Therefore, the selection principle that we stick to is to minimize the prediction error.

3)

3.4 Model Built to Solve Question *3*

3.4.1 *Economic Vitality Measurement and Ranking Model Based on PCA*

According to question *3* and the surveyed data from *National Bureau of Statistics* and *Local Bureau of Statistics*, we adopt the method PCA to build the economic vitality measurement and ranking model. We give the details of model construction using PCA by steps as follows.

1) Firstly, we use the data of 10 indexes to build the index system to create the index matrix $IN_{19\times10}$ which is defined in the following equation.

$$IN = (a_{ij})_{19\times10} \tag{32}$$

where the integer i ranges from 1 to 19 representing each city in *Attachment 3* and the integer j ranges from 1 to 10 representing indexes *GDP (100 million), gross value of the primary industry (100 million), gross value of the secondary industry (100 million), gross value of the tertiary industry (100 million), the amount of residents (10000), average salary of employees (yuan), household saving balance at the end of the year (100 million), total retail sales of goods (100 million), total imports and exports (million dollars) and the quantity of enterprises (10000)* respectively. Therefore, the indexes can be symbolized as x_j.

2) Secondly, we normalize each element in the index matrix using the following equation.

$$\tilde{a}_{ij} = \frac{a_{ij} - \mu_j}{s_j} \tag{33}$$

$$\mu_j = \frac{1}{19}\sum_{i=1}^{19} a_{ij} \tag{34}$$

$$s_j = \sqrt{\frac{1}{18}\sum_{i=1}^{19}(a_{ij} - \mu_j)^2} \tag{35}$$

where μ_j and s_j are the sample mean value and the sample standard deviation of jth index respectively. Accordingly, the indexes x_j can be normalized as $\tilde{x}_j$ with the following equation.

$$\tilde{x}_j = \frac{x_j - \mu_j}{s_j} \tag{36}$$

3) Thirdly, we formulate the correlation coefficient matrix R with the following equation.

$$R = (r_{j_1 j_2})_{10\times10} \tag{37}$$

$$r_{j_1 j_2} = \frac{\sum_{i=1}^{19} \tilde{a}_{ij_1} \cdot \tilde{a}_{ij_2}}{18} \tag{38}$$

where j_1 and j_2, with different value from 1 to 10, are the same as j. $r_{j_1 j_2}$ is the correlation coefficient of j_1th index and j_2th index and also the element of correlation coefficient matrix.

4) Fourthly, we calculate the feature values λ_j ($\lambda_j \geq 0$) and corresponding normalized feature vectors $[u_{1j}, u_{2j}, \ldots, u_{10j}]^T$ of correlation coefficient matrix R. Then the principal components formed by 10 normalized feature vectors are formulated as follows.

$$\begin{cases} y_1 = u_{11}\tilde{x}_1 + u_{21}\tilde{x}_2 + \ldots + u_{101}\tilde{x}_{10} \\ y_2 = u_{12}\tilde{x}_1 + u_{22}\tilde{x}_2 + \ldots + u_{102}\tilde{x}_{10} \\ y_3 = u_{13}\tilde{x}_1 + u_{23}\tilde{x}_2 + \ldots + u_{103}\tilde{x}_{10} \\ y_4 = u_{14}\tilde{x}_1 + u_{24}\tilde{x}_2 + \ldots + u_{104}\tilde{x}_{10} \\ y_5 = u_{15}\tilde{x}_1 + u_{25}\tilde{x}_2 + \ldots + u_{105}\tilde{x}_{10} \\ y_6 = u_{16}\tilde{x}_1 + u_{26}\tilde{x}_2 + \ldots + u_{106}\tilde{x}_{10} \\ y_7 = u_{17}\tilde{x}_1 + u_{27}\tilde{x}_2 + \ldots + u_{107}\tilde{x}_{10} \\ y_8 = u_{18}\tilde{x}_1 + u_{28}\tilde{x}_2 + \ldots + u_{108}\tilde{x}_{10} \\ y_9 = u_{19}\tilde{x}_1 + u_{29}\tilde{x}_2 + \ldots + u_{109}\tilde{x}_{10} \\ y_{10} = u_{110}\tilde{x}_1 + u_{210}\tilde{x}_2 + \ldots + u_{1010}\tilde{x}_{10} \end{cases} \tag{39}$$

where y_j denotes the jth principal component.

5) Fifthly, we select p out of 10 principal components to measure their comprehensive appraisal value. Contribution rate b_j of each feature value λ_j is also the contribution rate of corresponding principal component y_j, which is formulated as follows.

$$b_j = \frac{\lambda_j}{\sum_{j=1}^{10} \lambda_j} \tag{40}$$

Then the accumulation contribution rate of p principal components can be formulated as follows.

$$\partial_p = \frac{\sum_{j=1}^{p} \lambda_j}{\sum_{j=1}^{10} \lambda_j} \tag{41}$$

If ∂_p approximates 1, it means that we can choose these p principal components to substitute for the original 10 principal components to simplify the appraise.

6) Lastly, the result of comprehensive appraise Z can be acquired through computation. The following equation gives the formulation of z.

$$Z = \sum_{j=1}^{p} b_j y_j \tag{42}$$

3.4.2 *Data and Model Results*

Table 9 The surveyed data of 10 indexes for city ranking.

City	Shanghai	Shenzhen	Beijing	Guangzhou	Chongqing	Chengdu	Nanjing	Hangzhou	Suzhou
GDP	30633	22490	28015	21503	19425	13889	11715	12603	18597
Gross value of the primary industry	111	20	120	233	1276	501	263	311	214
Gross value of the secondary industry	9331	9320	5327	6015	8585	5998	4455	4362	8933
Gross value of the tertiary industry	21192	13150	22568	15254	9564	7390	6997	7930	9450
The amount of residents	1455	563	1359	898	3390	1435	681	754	1068
Average salary of employees	130765	100173	134994	98612	73272	79292	101502	96670	87350
Household saving balance at the end of the year	6642	10838	5431	1537	2252	1276	1272	1567	8166
Total retail sales of goods	11830	6016	11575	9403	8068	6404	5605	5717	4956
Total imports and exports	476197	414146	324017	143250	66601	58316	61187	69013	316079
The quantity of enterprises	157	174	118	90	70	61	56	49	54

City	Tianjin	Qingdao	Dongguan	Zhengzhou	Wuhan	Xi'an	Ningbo	Changsha	Shengyang	Kunming
GDP	18549	11037	7582	9130	13410	7470	9842	10536	5865	4858
Gross value of the primary industry	169	381	24	159	408	281	306	379	268	210
Gross value of the secondary industry	7594	4546	3594	4248	5861	2596	5119	4998	2261	1866
Gross value of the tertiary industry	10787	6110	3965	4724	7141	4593	4417	5158	3335	2782
The amount of residents	1050	803	834	842	854	906	597	709	737	563
Average salary of employees	96965	83539	53446	70486	79684	77774	91705	85187	74181	76350
Household saving balance at the end of the year	2310	1157	5103	1057	1403	655	1245	800	656	561
Total retail sales of goods	5730	4541	2688	4057	6196	4330	4048	4548	3990	2591
Total imports and exports	112919	74124	12264	59635	28597	37700	112197	13886	12846	7818
The quantity of enterprises	44	41	43	43	40	38	31	29	22	24

Based on the model constructed using PCA, we apply *Matlab* to process surveyed data and produce model results. The surveyed data of the 10 indexes for question *3* is shown in Table 9. (*Sources: National Bureau of Statistics and local bureau of statistics*)

Primarily we select all of the 10 normalized indexes deviated from the original indexes and obtain the 10 principal components. Then we figure out the feature values of each normalized index to the 10 principal components through computation and the numbers in Table 10 show them clearly.

Table 10 Feature values of each normalized index to the principal components.

Principal Component	$\tilde{x}_1$	$\tilde{x}_2$	$\tilde{x}_3$	$\tilde{x}_4$	$\tilde{x}_5$	$\tilde{x}_6$	$\tilde{x}_7$	$\tilde{x}_8$	$\tilde{x}_9$	$\tilde{x}_{10}$
y_1	0.392416	0.084282	-0.06727	0.09987	-0.28189	0.081519	-0.13994	-0.0376	-0.35948	0.766864
y_2	-0.05294	0.66503	0.051769	0.1299	0.496243	-0.01219	-0.34922	0.338802	-0.2268	-0.02809
y_3	0.315398	0.17511	0.417432	0.664158	-0.17558	0.324296	0.098946	-0.15247	0.163802	-0.24216
y_4	0.380725	0.005971	-0.25961	-0.14805	-0.31597	-0.0264	-0.20459	-0.00241	-0.5204	-0.59371
y_5	0.128161	0.62569	0.1067	-0.3409	-0.17424	-0.36302	0.333918	-0.41397	0.139556	6.79E-07
y_6	0.325207	-0.1309	-0.46454	0.352311	0.412176	-0.38635	-0.14953	-0.39561	0.197891	-1.1E-06
y_7	0.293202	-0.20703	0.568556	-0.23893	-0.02374	-0.34884	-0.55047	0.075506	0.244643	7.24E-07
y_8	0.345296	0.17889	-0.39818	-0.14414	-0.18538	0.20288	-0.01918	0.480908	0.601894	2.84E-07
y_9	0.37035	-0.17199	0.173384	0.060711	0.237163	-0.34733	0.600566	0.471349	-0.19945	-2.3E-06
y_{10}	0.366386	-0.08286	0.109983	-0.43387	0.502184	0.568158	0.094276	-0.27116	-0.03388	5.02E-07

Table 11 Contribution rates and accumulation contribution rates of each principal component to the comprehensive appraise index.

Principal Component	**Contribution Rate (%)**	**Accumulation Contribution Rate (%)**
y_1	62.51726	62.517263
y_2	20.98678	83.5040424
y_3	10.08538	93.5894205
y_4	2.671435	96.260856
y_5	1.50563	97.7664857
y_6	1.334661	99.1011462
y_7	0.468751	99.5698969
y_8	0.326133	99.8960302
y_9	0.10397	99.9999999
y_{10}	4.12E-11	100

Afterwards, the contribution rates and the accumulation contribution rates of each principal component to the comprehensive appraise Z are shown in Table 11

Based on the results shown in Table 10 and Table 11, we finally select the first 4 principal components y_1, y_2, y_3 and y_4 instead of the original 10 principal components according to their contribution rates and accumulation contribution rates. The result of measurement and ranking model is formulated as follows.

$$Z = 0.6252\tilde{y}_1 + 0.2099\tilde{y}_2 + 0.1009\tilde{y}_3 + 0.0267\tilde{y}_4 \tag{43}$$

Finally, we can easily rank the 19 cities in *Attachment 3* according to appraisal index Z and the ranking results are shown in Table 12.

Table 12 Ranking and appraisal of cities.

City	Rankings	Appraisal Index
Shanghai	1	3.56886
Beijing	2	2.505138
Shenzhen	3	2.22517
Chongqing	4	1.774264
Suzhou	5	1.074082
Guangzhou	6	0.896069
Tianjin	7	0.326718
Chengdu	8	-0.02949
Wuhan	9	-0.45961
Hangzhou	10	-0.52865
Nanjing	11	-0.59889
Qingdao	12	-0.824
Ningbo	13	-0.9253
Changsha	14	-1.01348
Zhengzhou	15	-1.21076
Xi'an	16	-1.42262
Dongguan	17	-1.47376
Shengyang	18	-1.79621
Kunming	19	-2.08752

IV. Conclusions

4.1 Conclusions of Question *1*

-
-
-
-

4.2 Conclusions of Question *2*

-
-
-
-

4.3 Conclusions of Question *3*

- We select 10 indexes to especially measure economic vitality of cities. The results show that these 10 indexes perform well.
- The reason why we choose the first 4 principal components is that their accumulation contribution rates reach 96.26% and go beyond 95%. Besides, each contribution rate of them is higher than the other.
- The appraisal index Z is the mathematical model that analyses and measures economic vitality.
- From the ranking and appraisal result, Shanghai ranks 1st, followed by Beijing and Shenzhen. Shanghai is approximately 1.06 higher than the second Beijing and the following cities differ from the previous one by 0.2-0.5. Cities that rank middle like Wuhan, Hangzhou, Nanjing and so on have little difference in terms of economic vitality. Cities like Xi'an, Dongguan, Shengyang and Kunming have a low ranking, which means they have relatively low economic vitality and need improvement.
- It is worth noting that Chongqing and Wuhan rank high in the results. That is probably because Chongqing has a large population and benefits from national policies and Wuhan is attracting the talents in recent years for better development.
- As a whole, the rankings and appraisal largely share similarity with our common sense and reality, which means our proposed model using PCA has satisfying

performance.

4.4 The Development Proposal for Eastern region

the economic vitality in this region presents the benign sustainable development and the regional competitiveness is stronger.

V. Future Work

5.1 Strength and Weakness of the Whole Model

Strength: The Improved Model aims to make up for the neglect of . The result seems to declare that this model is more reasonable than the Basic Model and much more effective than the existing design.

Weakness: Thus the model is still an approximate on a large scale. This has doomed to limit the applications of it.

5.2 Possible Optimization

VI. References

[1] Author, Title, Place of Publication: Press, Year of publication.

[2] author, paper name, magazine name, volume number: starting and ending page number, year of publication.

[3] author, resource title, web site, visit time (year, month, day).

[4] LaTeX资源和技巧学习 `https://www.latexstudio.net`

[5] LaTeX问题交流网站 `https://wenda.latexstudio.net`

[6] 模板库维护 `https://github.com/latexstudio/APMCMThesis`

VII. Appendix

Listing 1: The matlab Source code of Algorithm

```
kk=2;[mdd,ndd]=size(dd);
while ~isempty(V)
[tmpd,j]=min(W(i,V));tmpj=V(j);
for k=2:ndd
[tmp1,jj]=min(dd(1,k)+W(dd(2,k),V));
tmp2=V(jj);tt(k-1,:)=[tmp1,tmp2,jj];
end
tmp=[tmpd,tmpj,j;tt];[tmp3,tmp4]=min(tmp(:,1));
if tmp3==tmpd, ss(1:2,kk)=[i;tmp(tmp4,2)];
else,tmp5=find(ss(:,tmp4)~=0);tmp6=length(tmp5);
if dd(2,tmp4)==ss(tmp6,tmp4)
ss(1:tmp6+1,kk)=[ss(tmp5,tmp4);tmp(tmp4,2)];
else, ss(1:3,kk)=[i;dd(2,tmp4);tmp(tmp4,2)];
end;end
dd=[dd,[tmp3;tmp(tmp4,2)]];V(tmp(tmp4,3))=[];
[mdd,ndd]=size(dd);kk=kk+1;
end; S=ss; D=dd(1,:);
```

Listing 2: The lingo source code

```
kk=2;
[mdd,ndd]=size(dd);
while ~isempty(V)
   [tmpd,j]=min(W(i,V));tmpj=V(j);
for k=2:ndd
   [tmp1,jj]=min(dd(1,k)+W(dd(2,k),V));
   tmp2=V(jj);tt(k-1,:)=[tmp1,tmp2,jj];
end
   tmp=[tmpd,tmpj,j;tt];[tmp3,tmp4]=min(tmp(:,1));
if tmp3==tmpd, ss(1:2,kk)=[i;tmp(tmp4,2)];
else,tmp5=find(ss(:,tmp4)~=0);tmp6=length(tmp5);
if dd(2,tmp4)==ss(tmp6,tmp4)
   ss(1:tmp6+1,kk)=[ss(tmp5,tmp4);tmp(tmp4,2)];
else, ss(1:3,kk)=[i;dd(2,tmp4);tmp(tmp4,2)];
```

```
end;
end
    dd=[dd,[tmp3;tmp(tmp4,2)]];V(tmp(tmp4,3))=[];
    [mdd,ndd]=size(dd);
    kk=kk+1;
end;
S=ss;
D=dd(1,:);
```