For office use only	Team Control Number	For office u
T1 _____	**79002**	F1 _____
T2 _____	Problem Chosen	F2 _____
T3 _____	**B**	F3 _____
T4 _____		F4 _____

2018

MCM/ICM

How many languages?

To predict the number of the one language, we assume that native speakers are related to natural growth rate of its native speaker and number of the its second speakers. Based on the data we collected, we use **time series languages speakers difference equation model** to describe the dynamic change of both native and second language speakers, considering the influence of foreign language taught in school, social media, economics, cultural communication and so on.

The difference equation model can apply our collected indicators to the prediction of change of language distribution over time. 50 years later, top 10 languages in order of total speakers change from [Mandarin, English, Hindustani, Spanish, Arabic, Malay, Russian, Bengali, Portuguese French] to [Mandarin ,English, Spanish, Hindustani, Arabic, Bengali, Portuguese, Malay, Russian, French]; the rank of top 10 native languages speakers changes from [Mandarin, Spanish, English, Hindustani, Arabic, Bengali, Portuguese, Russian, Punjabi, Japanese] to [Mandarin, Spanish, English, Hindustani, Arabic, Bengali, Portuguese, Punjabi, Russian and Hausa]. By analyzing these changes, we find some reasonable explanations, such as the rapid natural growth rate of some developing countries and some languages' increasing speaking-power.

Given the global population growth and migration pattern, we establish **geographical distribution of difference equation model**, to predict the geographical distribution of different languages. Through the establishment of the difference equation, we consider the relationship between the distribution of languages on different continents and main migration routes. We use MATLAB to calculate language proportion changes on each continent over the next 50 years, finding some reasonable predictions. For example, Mandarin will become the No.2 native language in North America and Australia. The proportion of Mandarin and Arabic speakers in Europe will increase significantly.

In Part II, based on the requirement and the feature of service company, we choose six suitable cities based on our prediction of language speakers. Also, we find that the cities are different depending on whether the company is long-term oriented (6 suggested cities: Shanghai, New York, Calcutta, Madrid, Dubai, and Rio de Janeiro) or short-term oriented (6 cities: Shanghai, New York, Calcutta, Madrid, Dubai, and Singapore).

Moreover, we build the **cost-benefit analysis model** to calculate the suitable number of offices that this company should build. Given the level of company's profitability and cost, we set a new parameter, cost-profitability ratio. If the value of c-p ratio is less than 281, we think 6 offices should be built. If the value of this ratio is between 281 and 422, we think 5 offices should be built. If the value is between 422-527, 4 offices are best; if it is between 527 and 544, 3 office should be built; if> 544, we should only maintain two offices.

Finally, we analyze the performance of our model and the sensitivity of our model, proving that our model is relatively stable for different parameters.

Key words: Language distribution, Time Series Difference Equation Model, Dynamic simulation, site selection Cost-profit analysis

CONTENT

1   Introduction

1.1   Problem Background

In the world of globalization, number of native speakers and L2 speakers of a certain language increase or decrease over time. There are many factors that affect the increase or decrease of a certain language, including the foreign language taught in school, cultural communication and assimilation, Economic Factor, technology, social media and so on.

Our first task is to establish a model of the distribution of various language speakers over time, in which we should consider the factor listed above.

Besides, our second task is the establish a model to predict the geographic distributions of these languages over time based on the given globe population and human migration patterns for the next 50 years.

In the part II, a large multinational service company hire our team to give location options for new offices. So our third task is to consider where we should locate these offices and if opening less than six offices better.

1.2   Our work

Language is such an important topic due to its role in cultural communication, international business, migration issue and so on. Under the circumstance that we are consulted to give out 6 most suitable sites to build new office by a service company, our main work is as follows:

Firstly, based on the data we collected, we use time series languages speakers difference equation model to describe the dynamic change of both native and second language speakers, considering the influence of foreign language taught in school, social media, economics, cultural communication and so on.

Secondly, considering the global population growth model and migration patterns, we establish geographical distribution difference model, presenting the change of languages' distribution in 6 main continents over 50 years.

Thirdly, we choose the six suitable cities based on our prediction of language speakers. Also, we find that the cities are different depending on whether the company is long-term oriented or short-term oriented.

Moreover, we build the cost-benefit analysis model to calculate the suitable number of offices that this company should build.

Finally, we analyze the performance of our model and the sensitivity of our model.

2   Assumptions and Symbols

2.1   Assumptions of the initial data

1. Those languages whose current total speakers are less than 100 million won't become the top 10 languages. Thus, according to the list of languages by total number of speakers, we only use the data of top 16 languages, since they are the only languages that are used by more than 100 million people.[1]

Reason: The French ranked 10th in 2017 with a total number of 228million speakers. According to common sense, total number of Language speakers have a small possibility to decrease. So, those languages with fewer than 100million speakers are less likely to become Top10 in 50 years. At the same time, we also do this to reduce our computational load and to reduce our programming difficulty.

2. For some Languages L2 speakers number '?'in the [1], we assume it is zero.

Reason: We speculate there may be two reasons for the coming of '?'. One is because the data is too
good statistics. The other reason is that it is controversial to define who can be the second foreig
speakers. For both reasons, we can all assume the number is zero.

3. We think native speaker's growth is only related to its own natural growth rate and second language population

Reason: According to our common sense, the growth of native speakers is often associated with changes in the local population. Local population grow, native speakers also will grow. Foreigners migrate in, and the foreign language native speakers increase. Therefore, native speaker changes and population changes are very relevant. And in order to simply our model, we think native speaker's growth is only related to its own natural growth rate and second language population.[2]

4.We think L2 speaker's growth is only affected by its own feature (the languages learned in school, cultural communication) and the global situation (economics, development of technology, media use). And the relationship between them is directly proportional.

Reason: According to common sense, these factors positively affects the L2 speaker's growth. Although we are not quite sure whether the relationship between them is linear or not, in order to simplify the problem, we may think that the relationship between them is linear, and therefore, proportional.

2.2 Symbols and definitions

Table 1 Symbols and definitions

Symbols	Definitions
$x_i(t)$	The population of the native speaker of language i in year t
$y_i(t)$	The population of the L2 speaker of language i in year t
$s_i(t)$	The population of the total speaker of language i in year t
$LangT_{ij}$	the percentage of language i taught in the area where the native speakers use language j
$Cult_{i,j}$	the extent of cultural exchange between language i and language j
Eco_i	the economic power of language i
c	the velocity of technology development, c>0
Net_i	the power of language i in the internet
k_n	Scale coefficient in the difference equation, we will discuss in more detail in the text. n=1, 2,…6
$Pop(city_i)$	the number of speakers of top 10 languages except English in $city_i$

Table 2 codes for top 16 languages

i	1	2	3	4	5	6	7	8
Languages	Mandarin	English	Hindustani	Spanish	Arabic	Malay	Russian	Bengali
i	9	10	11	12	13	14	15	16
Languages	Portuguese	French	Hausa	Punjabi	German	Japanese	Persian	Swahili

3   Part I Models and Results

3.1   Model I: Various Languages Speakers Difference Equation Model

How to quantify the relationship between language growth and the various factors is a difficult problem? There are a variety of time series models we could use. If we have enough data, we can regress the functional relationship between the number of languages on various factors. But the truth is, we can't find enough credible data online. There are two reasons for this. One is because it is vague to judge whether a person has a second foreign language ability. For this reason, different data sources may not come from the same criterion. So the data between the two will be very different. In addition, the two data whose sources are the same, while years are different are still not credible. This is because the dates of the censuses vary from country to country. This makes a lot of data does not have time continuity. For these two reasons, we can't and will not use those methods of fitting forecasts.

Difference equation model can consider the impact of different factors on the size of the independent variables. Moreover, the difference equation model only requires the initial data on it. These two characteristics fit very well with our problem. The time step in which we set in the difference equation is one year. We denote that the number of native speakers of language i in next period depends on the current number of its native speakers and the current number of its 2nd language speakers. Therefore, we can construct a difference equation model to describe the change of 2 types speakers for the 16 languages.

3.1.1   The increase of the native speakers

According to our common sense, the growth of native speakers is often associated with changes in the local population. Local population grow, and native speakers also will grow. Foreigners migrate in, and the foreign language native speakers increase. Therefore, native speaker changes and population changes are very relevant.

So, we assume that native speaker's growth is only related to its natural growth and the number of second language speakers. We choose the weighted average of natural growth rate of countries whose official language is the **language i** as the increase of the native speakers of language **i.**

However, the world population growth pattern varies from countries. The most significant difference is the difference between the population growth patterns in developing and developed countries. Therefore, we divide these languages into two group. One type is mainly spoken in developed countries, and the other is mainly spoken in the developing countries. Developed countries usually have low natural population growth rates and high immigration rate while developing countries usually have higher natural population growth rates. To take this difference into consideration, we divide all 16 languages into 2 types, according to their main speakers' types—developed or developing.

Table 3   types of languages

types	*languages*
Type I: most spoken in developed countries	English, Spanish, French, German, Japanese, Mandarin
Type II: most spoken in developing countries	Hindustani, Arabic, Malay, Portuguese, Hausa, Punjabi, Persian, Swahili

Type I languages' speakers' growth model: Malthusian growth model

Type I languages are mainly spoken in developed countries, whose population usually have a very low natural

growth rate, such as 0.5%, or even a negative value. Its population growth rate mainly depe
immigration, while the next generation of those immigrants usually have to learn the local langua
mother tongue.

For this type countries, we can consider it as a constant rate, r, over the next 50 years. So we use the **Malthusian growth model** to measure the number of native speakers[4]. According to Malthusian growth model, the population growth rate, r, does not change over time. Therefore, the population in year t+1 could be written as:

$$x_i(t+1) - x_i(t) = x_i(t)r_i + R \cdot y_i(t) \tag{1}$$

$x_i(t)$ denotes the number of native speakers of language i in year t; r_i denotes the natural speaker growth rate of language i; R denotes a coefficient that explains the proportion of 2nd speakers of this language in year t that turns into native speaker in year t+1. Later, we will do sensitivity analysis for this coefficient R.

Type II languages' speakers' growth model: Logistic Growth Model

Type II languages are mainly spoken in developing countries, whose population usually have a relatively higher natural growth rate, such as 2%. Nevertheless, we cannot use this number for the next 50 years, since their population cannot maintain growing so fast. To make a more realistic analysis, we use Logistic Growth Model. Logistic Growth Model is a slight modification of Malthus's model. It points that the population growth rate is not constant—there is a limited carrying capacity of the environment, resulting in a stable population over time[5].

According to Logistic Growth Model, the population in year t+1 could be written as:

$$x_i(t+1) - x_i(t) = \left(r_i - \frac{x_i(t)}{s_i}\right) \cdot x_i(t) \tag{2}$$

r_i denotes the current natural growth rate of Type II languages; s_i is the maximum of numbers of speakers of language i, To make it simple, we regard $s_i = 2x_i(o)$. $x_i(o)$ denotes the initial number of native speakers of language i.

3.1.2 The increase of the second language speakers

When discussing the change of 2nd language speakers, we believe it is mainly affected by two major factors. Firstly, it is affected by its own feature, such as region, difficulty, promotion by the government, cultural communication, etc. Secondly, it is affected by the global situation, such as economics, development of technology, etc.

3.1.2.1 Effect of the language own features

● **the foreign language taught in school [3]**

We assume that the language promoted by the government is the foreign language taught in school. For example, English is taught in many countries as students' main foreign language because of English's wide spread and use in the world. Although statistics are incomplete in some developing countries, we only found data of some developed countries. However, we think this promotion of language in education is related to the power of the language itself, which means the number of L2 speakers. Therefore, we use available data to make up for the missing data.

$$\Delta_{i,\text{LangT}} = k_1 \sum_{j=1}^{16} LangT_{ij} \cdot x_j(t)$$

$LangT_{ij}$ denotes the percentage of language i taught in the area where the native speakers use language j.

For example, when i refers to French and j refers to English, $LangT_{ij}$ is the percentage of Fre in school as a foreign language in English-speaking countries. k_1 is the coefficient to measure this factor (the language taught in school) will influence the increase of 2nd language speakers i. Later we will do sensitivity analysis for k_1.
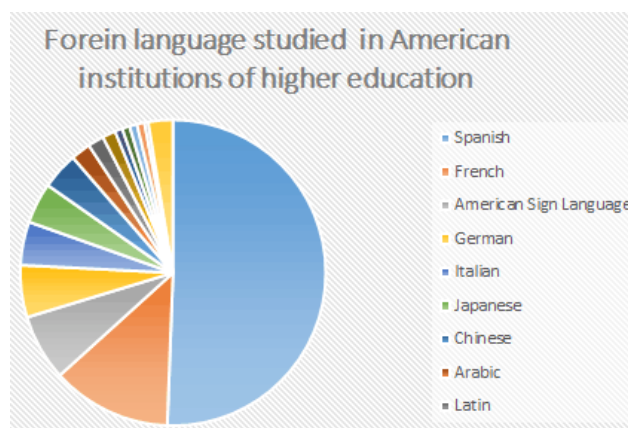

Figure 1 The foreign language studied in school, taking U.S. for example[6]

- **Cultural communication & assimilation**

Many reports state that cultural communication and assimilation play a significant role in the increase of 2nd language learners. However, how to quantify cultural exchange is a very difficult issue. In order to solve this problem, we propose two factors that will affect cultural communication. The following are two effects:

Neighborhood Effect: If the speakers of two language live near to each other, it is more likely for them to have cultural communication. And therefore, they would become more likely to learn neighbors' language as their 2nd language.
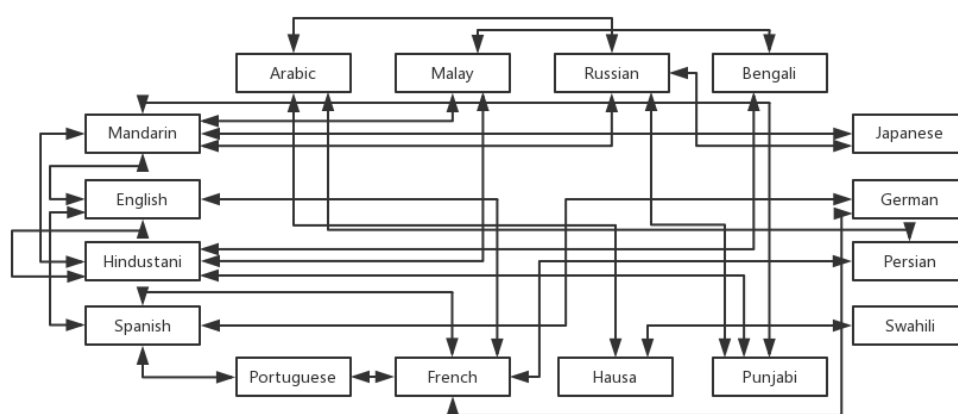
Figure 2 The neighborhood effect among the top 16 languages

(the two language having a line linked means there will be Neighborhood effect between them)

Policy-led Effect: When the home country of language speakers is strongly promoting the country's relations with some special countries, the likelihood of those languages learning from each other's languages will also increase, such as China's the Belt and Road Policy making it more motivated for the people along the policy to learn Chinese. For example, Pakistan uses Chinese as its second language.

The power of these two effect is positively related to the population of language speakers, x_j and the cultural communication between two languages. We get the following equation:

$$\Delta_{i,cult} = k_2 \cdot \sum_{j=1}^{16} Cult_{i,j} \cdot x_i(t) \cdot x_j(t)$$

$Cult_{i,j}$ describes the extent of cultural exchange between language i and language j; k_2 is the coefficient to measure how much this factor (the extent of cultural communication and assimilation) will influence the increase of 2nd language speakers of language i. Later we will do sensitivity analysis for k_2.

3.1.2.2   Effect of the global situation

● Economic Factor

Different languages have different speaking power in the global business environment. The more power it has, the more persons choose to learn it. Through the GDP contributed by different languages [See **Figure 3**], we give different weight to the top 16 language.

$$\Delta_{i,Eco} = k_3 \cdot Eco_i$$

Eco_i is the economic power of language i. k_3 describes how much the economic factor motivates people to learn language i as their 2nd language. Later we will do sensitivity analysis for k_3.
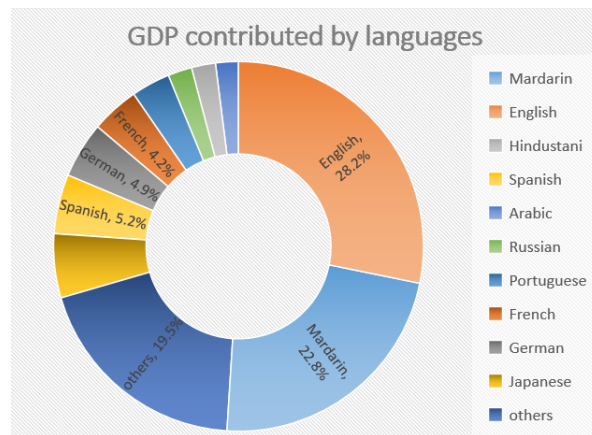

Figure 4 GDP contributed by language [7]

● **the update of translation software technology**

The development of technology will make it easier and faster to translate different languages. We assume the velocity of translation software's update rate is constant. Therefore, it will have a negative impact on the number of 2nd language learners. This factor will influence the increase of L2 speakers of language i as:

$$\Delta_{i,Tech} = k_4 c \quad (c > 0, \ k_4 < 0)$$

C is the velocity of technology development. k_4 describes the change of 2nd language learner influenced by the development of translation software. Later we will do sensitivity analysis for k_4.

● **the push of network and social media**

Though there are thousands of languages all over the world, only 5% of them are used widely in the internet. 54.5% of all web content is still in English despite huge growth in users that do not understand English or who prefer to access content in their native languages. Due to analysis of the most popular 10 million websites by W3techs, after English, the most common languages are Russian (5.9%), German (5.7%), Japanese (5.0%), and Spanish (4.7%).[8]

Data also show us that the number of languages used by mainstream social medias, such as Twitter and LinkedIn is limited to few languages.
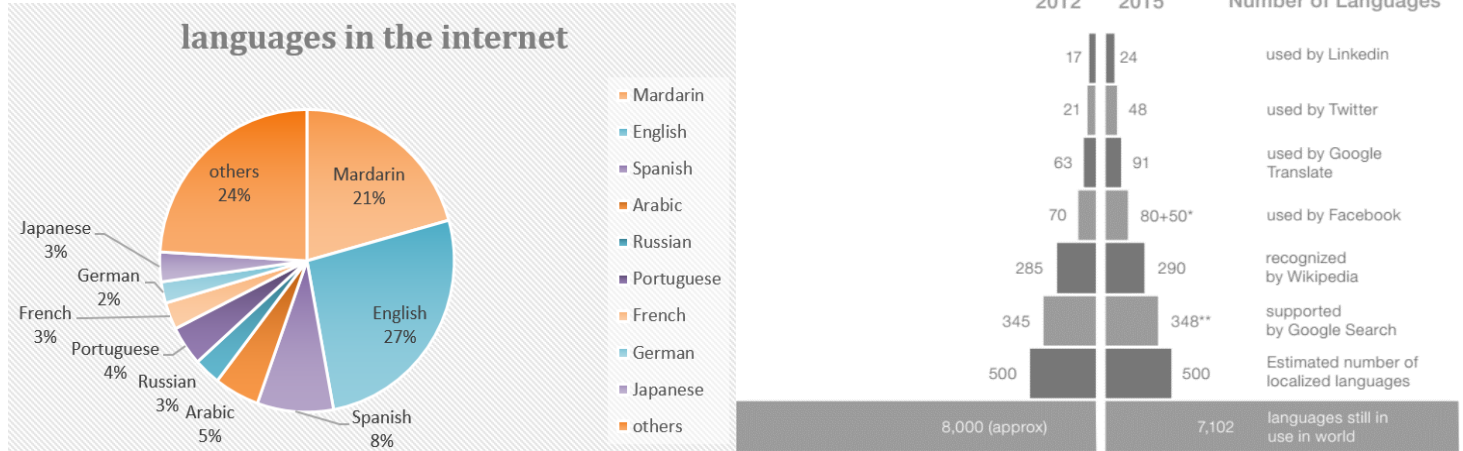

Figure 5 languages in the internet (left) [9] & the pressure of mainstream website(right)[8]

We can express the power of language in the network as follows:

$$\Delta_{i,\text{Net}} = k_5 \cdot Net_i$$

Net_i describes the power of language i in the internet. k_5 describes the change of 2nd language learner influenced by the push of internet and social media. Later we will do sensitivity analysis for k_5.

3.1.3   The total difference equation of model I

In this way, we get the general formula describing the change in the number of speakers learning i as their 2nd languages:

$$y_i(t + 1) - y_i(t) = \Delta_{i,LangT} + \Delta_{i,cult} + \Delta_{i,Eco} + \Delta_{i,Tech} + \Delta_{i,Net} \tag{3}$$

Adding (1) and (3), we get **the change of total speakers for Type I language**:

$$S_i(t + 1) - S_i(t) = x_i(t)r_i + R \cdot y_i(t) + \Delta_{i,LangT} + \Delta_{i,cult} + \Delta_{i,Eco} + \Delta_{i,Tech} + \Delta_{i,Net} \tag{4}$$

Adding (2) and (3), we get **the change of total speakers for Type II language**:

$$S_i(t + 1) - S_i(t) = \left(r_i - \frac{x_i(t)}{s_i}\right) \cdot x_i(t) + \Delta_{i,LangT} + \Delta_{i,cult} + \Delta_{i,Eco} + \Delta_{i,Tech} + \Delta_{i,Net} \tag{5}$$

3.2   Model I Results & Analysis

3.2.1   Initial rank and parameter setting

In the section above, we build a difference equation model, which could predict the number of native speakers and second language speakers in the following years. We get the initial rank by total language speaker[1]. It represents the numbers of native speakers and second language speakers of different languages in 2017. And here is the initial data:

Table 4 The initial rank by language size [1]

Rank	Language	L1speakers	L1Rank	L2speakers	L2Rank	Total
1	Mandarin	897	1	193	4	1090
2	English	371	3	611	1	982

3	Hindustani	329	4	215	2	544
4	Spanish	436	2	91	8	527
5	Arabic	290	5	132	6	422
6	Malay	77	15	204	3	281
7	Russian	153	8	113	7	267
8	Bengali	242	6	19	13	261
9	Portuguese	218	7	11	15	229
10	French	76	17	153	5	229
11	Hausa	85	11	65	10	150
12	Punjabi	148	9	?	?	148
13	German	76	18	52	12	129
14	Japanese	128	10	1	19	129
15	Persian	60	25	61	11	121
16	Swahili	16	26	91	8	107

According to the data we collected in [1, 4-7, 9], we can set indices in the formula (4) and formula (5). The indices are given as follows:

Table 5 Indices setting

Index	k_1	k_2	k_3	k_4	k_5
value	$\dfrac{1}{300}$	$\dfrac{1}{90000}$	0.2	-0.3	0.2
Explanation	L2 learner condition	Cultural communication	Language's power in business	Technology factor	Language's power in internet

These parameters in our various languages speakers difference equation model are given by our estimate. We will do sensitivity analysis for them to judge if the change of L2 speakers will be sensitive to these indices.

3.2.2 Results & Analysis

After inputting the initial value and indices in MATLAB, we got the new rank:

Table 6 The predicted rank by language size after 50 years, Unit: million

rank	Languages	L1speakers	L1 rank	L2speakers	L2 rank	Total
1	Mandarin	1119.9556	1	397.2715	2	1517.227
2	English	625.07569	3	752.60658	1	1377.682
3	Spanish	714.54914	2	232.09149	7	946.6406
4	Hindustani	437.73828	4	322.65571	3	760.394
5	Arabic	410.4004	5	235.69087	5	646.0913
6	Bengali	337.48877	6	112.72084	13	450.2096
7	Portuguese	306.72269	7	101.83735	14	408.56
8	Malay	121.75222	11	269.89281	4	391.645
9	Russian	159.65859	9	187.58738	8	347.246
10	French	109.24506	13	234.28833	6	343.5334

11	Punjabi	226.42109	8	78.858666	15	305.279
12	Hausa	133.98132	10	132.39229	10	266.373
13	Persian	87.30792	14	122.95997	12	210.267
14	German	81.926176	15	125.04862	11	206.9748
15	Japanese	121.48183	12	71.268763	16	192.7506
16	Swahili	26.054248	16	145.23578	9	171.29

The top ten of total speakers vary from Mandarin、English、Hindustani、Spanish、Arabic、Malay、Russian、Bengali、Portuguese、French to Mandarin、English、Spanish、Hindustani、Arabic、Bengali、Portuguese、Malay、Russian、French.

The top ten of native speakers vary from Mandarin、Spanish、English、Hindustani、Arabic、Bengali、Portuguese、Russian、Punjabi、Japanese to Mandarin、Spanish、English、Hindustani、Arabic、Bengali、Portuguese、Punjabi、Russian、Hausa.

Comparing the two rankings, we can draw 5 main conclusions:

The fast increase of Mandarin L2 speakers: As the official language of China, a typical representative of the fast-growing countries, Mandarin has attracted many people to choose it as their second language. At the same time, China has also adopted a series of exchange policies with other countries, such as the aid to Africa and the Belt and Road Policy, which has also enhanced the attractiveness of Chinese.

Russian total rank declines: due to the very low natural growth rate in Russia, Russian native speakers' rank declines, directly leading the decline in Russian's total rank.

Bengali total rank increase: Bangladesh has a large national population base and a high natural increase rate, resulting in a rapid increase in Bengali native speakers and its rank.

The list of top 10 languages only have an intra-group change: Because there is a huge gap between the 10th language (French, total speakers:229 million) and the 11th language (Hausa, total speaker: 150 million). The gap doesn't disappear completely over time. But we can see from Table 7 that the top 16 list changes during 50 years.

Native speaker of Japanese decline from the top 10 list: This is mainly because Japan's natural growth rate has been negative recently. In addition, Hausa in Africa has become the 10th largest native language speaker due to its rapid rate of natural increase

3.3 Model II: Geographical Distribution Difference Model

As the model above established, we have established a model to measure the speakers' numbers of different languages over time. The model above quantifies the relationship between the trend of native speakers and second language speakers and School, migration of cultural group, economics, the use of translation technology and social media. But obviously, we did not consider the impact of geographical distribution on the language. The following model specifically addresses this issue.

However, how to put the geographical distribution into our model is a huge challenge. We did not know at the outset what amount of language should be used to measure the geographic distribution of languages. But as we looked up the data, we found that the language distribution across all continents varied greatly. Below is the language distribution in 2017:
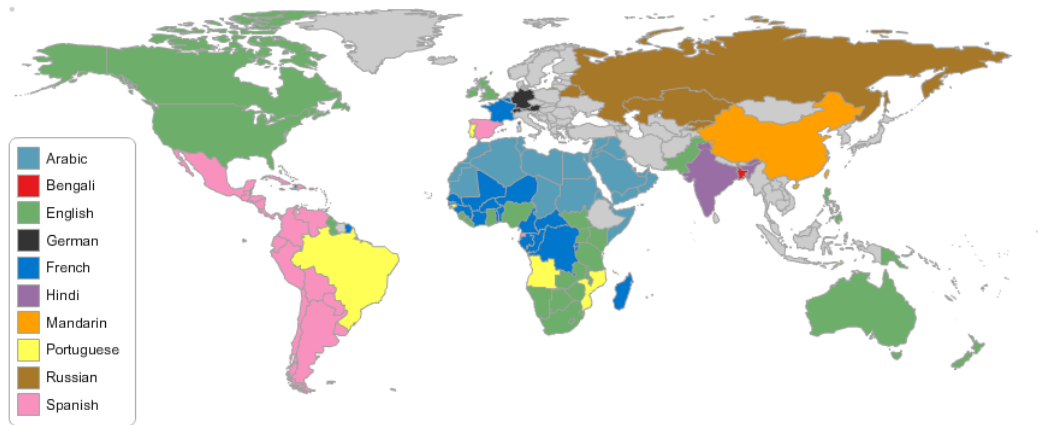
Figure 6 Language distribution in 2017

Therefore, we divide Earth language data into six continents, and each different continent has a different language distribution. Then, based on the modeling ideas of the above model, we consider the natural increase of population on all continents on the one hand, and population migration across the other continents on the other hand. Then we create a difference equation for each language of each continent over time. This equation can consider the impact of the annual immigrant population on the local language and the local language's own internal growth, which is what we want.

The following section considers assumptions about the world's population growth and the pattern of migration.

3.4 Population growth fitting and current migration pattern

In the following section, we will fit the world's population data and find the main migration path of the world's population.

Although we can easily find the world demographic data from the World Bank[10] from 1960 to 2016, how to fit the 56-year data to make their errors smaller is still a problem. We tried exponential fitting, polynomial fitting, logistic equation fitting and so on, and finally found that Gaussian function fitting the best. Below is our fitting effect and function expression.
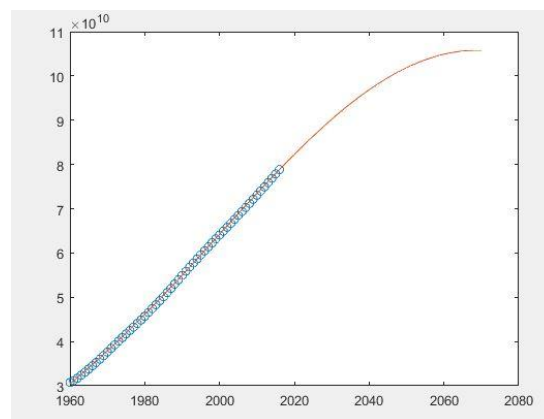

Figure 7 The regression on the world demographic data

$$N(t) = 1.058 * 10^{11} * e^{\left((t-2069)/97.67\right)^2}$$

N(t) denotes the world population in t years.

Current global migration model is very complicated. If we focus on the population transfer data in various

countries, we can easily fall into too many data and can't get the result. So, in order to simplify migration pattern, we may wish to consider only major migration routes in the current world. Then v the main migration patterns, and the image below shows some of the routes we found.
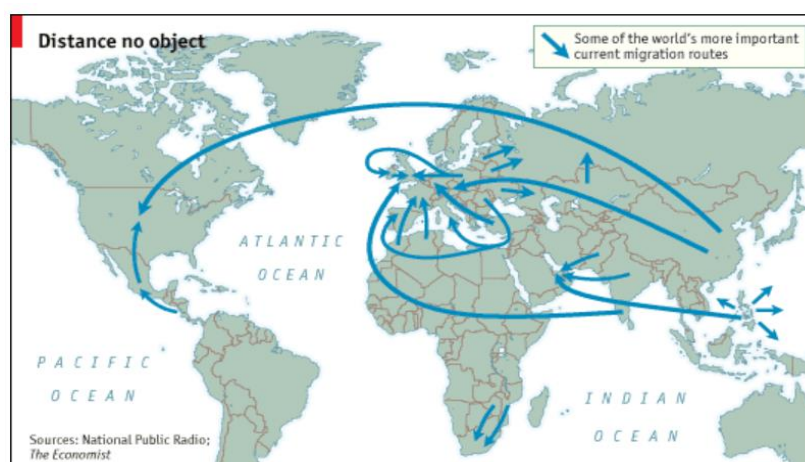

Figure 8 Some main routes of global migration[11]

Based on the current global migration patterns, we propose to assume the following seven major migration routes. These seven migration routes are the most promising migration path we believe will be in the next 50 years. Based on the average annual data and our understanding of these routes, we make the following assumptions as to the proportion of these seven routes in all the routes.

Table 8   7 main migration routes

route	details	proportion
1	China to US and Canada	20%
2	China to EU	15%
3	India to EU	15%
4	West Asia to EU	15%
5	north Africa to EU	15%
6	Latin America to US and Canada	15%
7	China to Australia	5%

3.5 The increase speakers of each language on each continent

Because we mainly consider the impact of immigration on the geographical distribution of language, we do not intend to consider the second language speakers here. All the following language-related data refers to the native speakers. Just like the model we built above, we also regard one year as the variation time of the difference equation. We denote $X_j^i(t)$ as the number of ith language speakers in j^{th} continent in the year t. Also, we denote

L_i as the number of speakers in the ith migration route, which is marked above.

When we think about the ith language native speakers' increase, the first thing we think of is the natural increase of the population. But different languages generally have different natural growth rates of its native speakers.

So we denote r_j^i is the natural increase rate of ith language in j^{th} continent.

When we think of the impact of migration route on native speakers, what we think first is that the population of immigrants will increase the number of native speakers of the language which the migration group speaks.

Besides, the population of immigrants will have the next generation in the destination country. B
difficult to measure the impact of this factor on the number of native speakers in that country, we a
the annual growth of native speakers in local languages because of migration is proportional to the
immigrants, and we set the index is k_6.

Thus, we have the following difference equation:

$$X_j^i(t+1) = X_j^i(t) + r_j^i \cdot X_j^i(t) + \sum_{l=1}^{6} k_6 \cdot L_l \cdot p_j^i * I_{\{destination==j\}} + \sum_{l=1}^{6} L_l \cdot I_{\{destination==j\}} \cdot I_{\{Language==i\}}$$

p_j^i denotes the ith language's proportion in jth continent; $I_{\{destination==j\}}$ judges whether the migration's destination is jth continent; $I_{\{Language==i\}}$ judges whether the migration's language is ith language.

After we calculated every language of the six continents, we can get the language distribution of each continent one year later. After 50 times of this process, we can so predict the geographical distribution of every continent 50 years later.

It should be noted that this model can only consider the impact of immigrants on the geographical distribution of language. There is no other factor considered in this model. In addition, we did not consider the changes of second language learners. The above model can only estimate the changes of native speakers on all continents over the next 50 years.

3.6 Result and Analysis

We use MATLAB to calculate that difference equations. And the chart below is our result. In order to easily differentiate data changes on all continents, we specialized in converting data of different continents to pie charts.
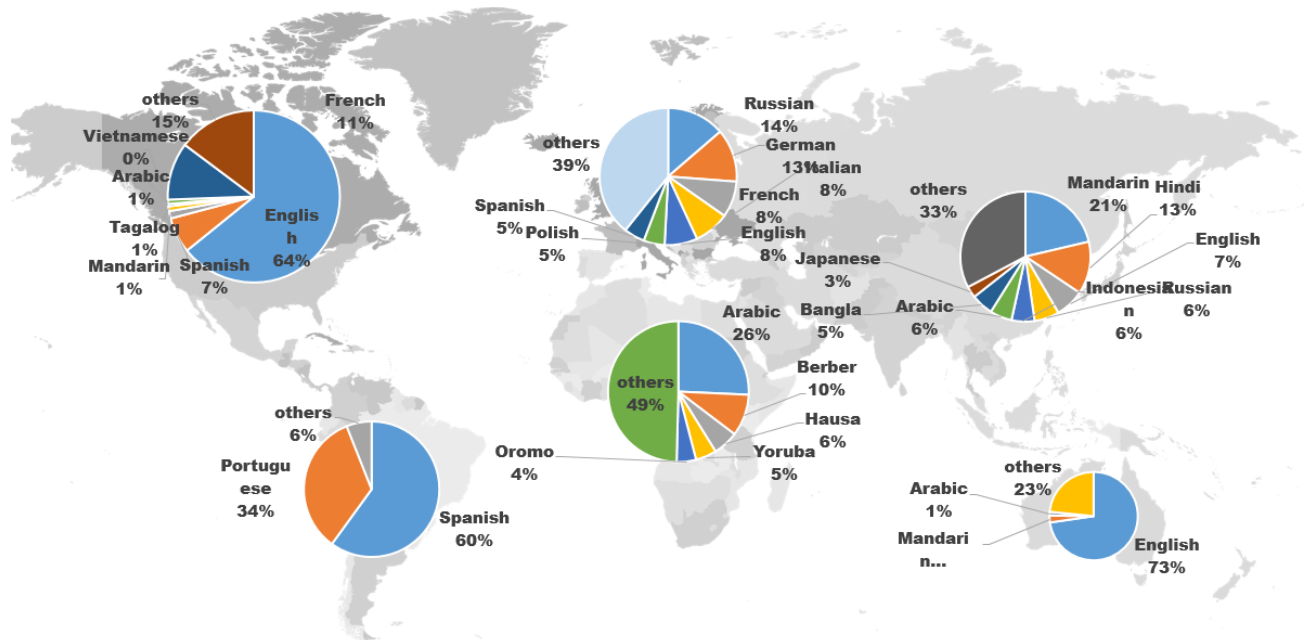

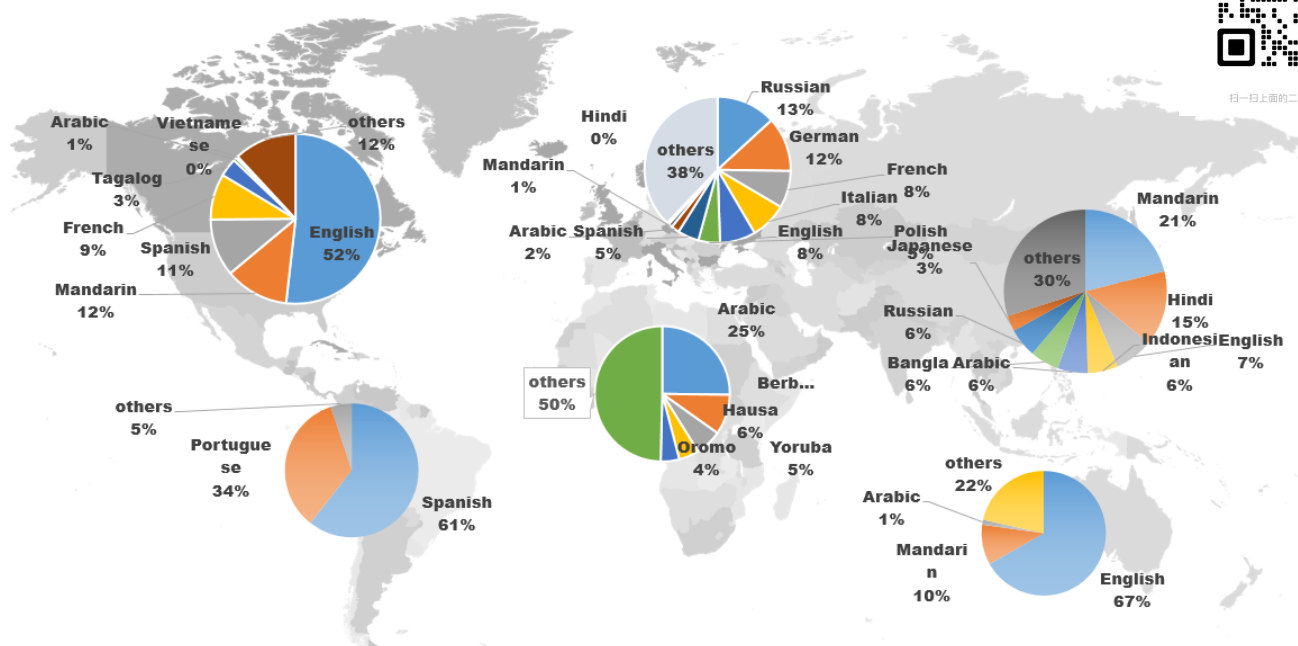Figure 9 The distribution of languages in 6 main continents in 2017.

Figure 10 The distribution of languages in 6 main continents in 2037
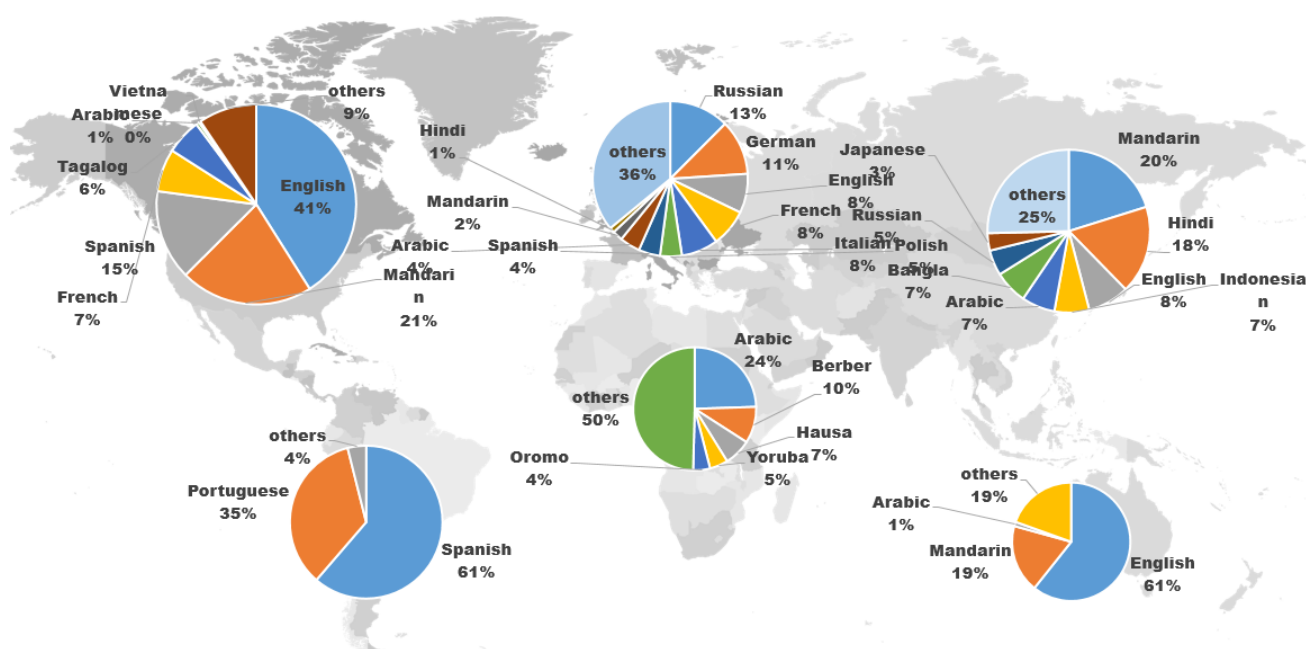
Figure 11 The distribution of languages in 6 main continents in 2067

According to the above changes show, we could find:

1. The distribution of languages in Africa, Australia, Asia and Latin America did not change much. This is because the linguistic changes in these two continents are mainly determined by the natural population growth. The mode of population migration has little effect on these two continents. (Although Africa is one of the immigrants' exit points and Australia is one of the destinations for immigrants).

2. The proportion of Mandarin and Spanish in the United States and Canada is on the rise. Portuguese native speakers appear and account for a certain percentage. This is because Route 1 brings a large number of Mandarin speakers to the United States and Canada, and Route 6 brings a large number of

Spanish and Portuguese speakers.

3. Native speakers of Arabic, Mandarin, Hindi appear in Europe and continue to increase. Ho to the small number of immigrants, the proportion is still low. The language of Europe still contains a lot of varieties.

4 Part II Models & Results

4.1 Assumptions about the service company

A service company is a business that generates income by providing services instead of selling physical products. A good example of a service company is a public accounting firm. They earn revenues by preparing income tax returns, performing audit and asset services, and even doing bookkeeping work.[12]

Based on our understanding of the service company, we make the following assumptions about the location of the new international offices.

1. The service company's profit is in direct proportion to the total number of languages it serves. The more languages it serves, the higher the profit it earns. This is also the main profit pattern of this service company.

2. The offices should be located in the places where English is widely used , in consideration of the crucial role English plays in the communication between different branch offices.

3. The offices tend to be located in a densely populated and easily accessible metropolitan area. That is, if there are two cities in the same language area, we prefer to choose a city with a large population and convenient traffic.

4.2 Explanation about our choices

There will not be major changes in language population in the short term. So, we refer to the number of language speakers in 2017 when considering the short-term site selection. The top-six used languages are Mandarin, English, Hindustani, Spanish, Arabic and Malay. The six locations are correspondingly Shanghai, New York, Calcutta, Madrid, Dubai, Singapore. The population density map and the locations of offices are shown below:
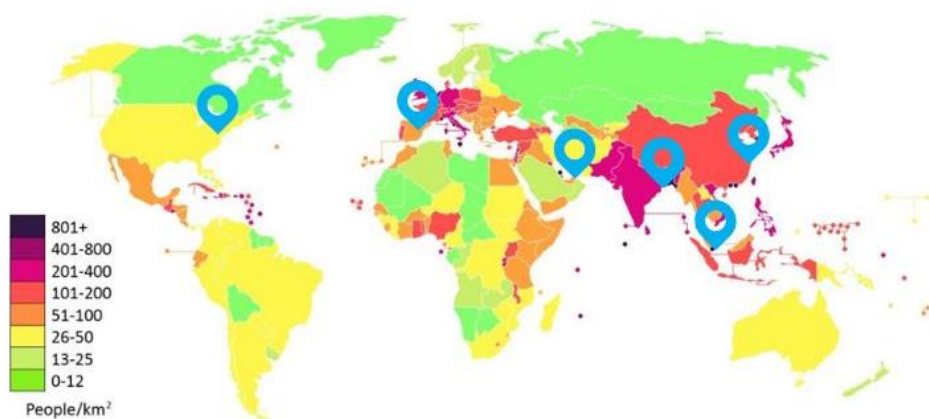

Figure 12 population density map and locations of 6 offices based on data in 2017

However, when we consider the site selection in the long term, we refer to the results of our model I. At this point, the top-six used languages have changed. They are Mandarin, English, Spanish, Hindustani, Arabid and

Bengali. However, due to its large population but underdeveloped economy in Bangladesh, the service projects lack local consumer groups. So, we consider the seventh language in the projected rankings –Portuguese. Therefore, we choose Shanghai, New York, Madrid, Calcutta, Dubai, and Rio to locate these offices, arranged in order.
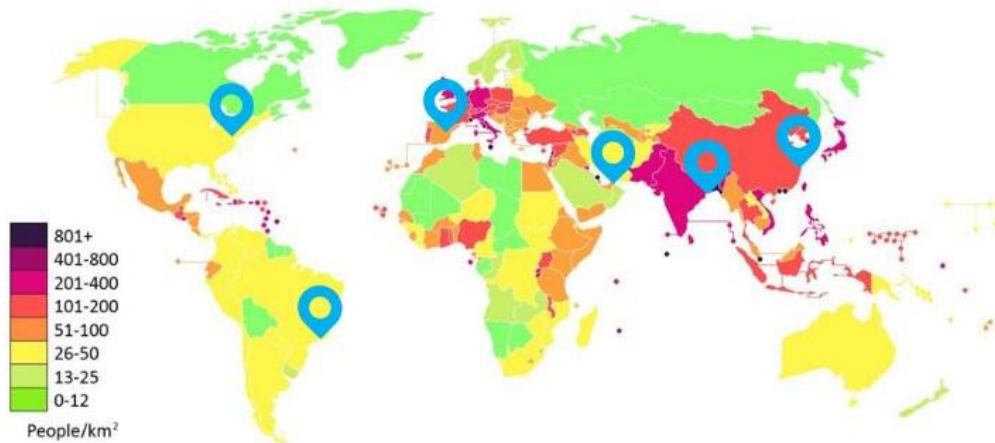

Figure 13 population density map and locations of 6 offices based on data in 2067

As we can see in Figure 14, Singapore is replaced by Rio de Janeiro since the rapid growth of Portuguese speakers.

Based on our model from Part I, here are our results:

In the short term, we choose **Shanghai, New York, Calcutta, Madrid, Dubai and Singapore** to locate these international offices. The top-six used languages-- Mandarin, English, Hindustani, Spanish, Arabic and Malay would be spoken correspondingly in these offices.

In the long term, we choose **Shanghai, New York, Madrid, Calcutta, Dubai, and Rio de Janeiro**. Mandarin, English, Spanish, Hindustani, Arabid and Portuguese would be spoken correspondingly in these offices. Cost-Benefit Analysis Model

Whenever establishing a new office in a new location, the company will have a wider range of consumer groups, and thus more revenue. But at the same time, building a new office also costs a lot. According to the general Benefit = Revenue − Cost equation, we need to clarify the source of company's revenue R and cost C.

Since English is a necessary language, we used the number of the most popular language speakers except English in that region, to judge the company's profitability when we set up our Cost-Benefit Analysis model. Here we give some Assumptions as follows:

- **R (Total Revenue)** is positively related to the company's profitability, which is α **C (Total Cost)** is a fix number in each place. We assume the **cost of building** a new office is a constant **C**.

So, we have the equation:

$$\text{Profit } (\text{city}_i) = \alpha \cdot \text{pop}(\text{city}_i) - C$$

$\text{Pop}(\text{city}_i)$ is the number of speakers of top 10 languages except English in city_i. Because we use $\text{pop}(\text{city}_i)$ to describe the power and popularity of language i, we could let the number of total speakers to roughly represent it.

We have calculated the most suitable cities in 4.1. The rank is in Table 9

Table 10 The most suitable cities

Rank	1	2	3	4	5	6
City	Shanghai	New York	Madrid	Calcutta	Dubai	Singapore

Note: 1. Shanghai and New York is the two office we already have.

2. We use the data collected in 2017 rather than the predicted value for 50 years later, because th will construct the 6 new office now but not 50 years later.

To calculate the profit-maximize office amount, we need to ask some additional information from our client company, which is the exact value of α and C, from the company to measure its profitability.

For $city_i$, we have a judgement to determine whether we should set up an office here:

$$
\begin{cases}
If \ \dfrac{C}{\alpha} > pop(city_i), & set \ up \ an \ office \ in \ city_i \\
If \ \dfrac{C}{\alpha} \le pop(city_i), & do \ not \ set \ up \ an \ office \ in \ city_i
\end{cases}
$$

C/α measures the company's ability to turn the cost into profit, we call it **cost-profitability ratio**, or **c-p ratio** in short.

Given different value of C/α, we can give the following suggestions:

Table 11 How many offices should the company set up?

C-p ratio (unit: million)	Number of offices	City
[544.982)	2	Shanghai, New York
[527,544)	3	Shanghai, New York, Madrid,
[422,527)	4	Shanghai, New York, Madrid, Calcutta
[281, 422)	5	Shanghai, New York, Madrid, Calcutta, Dubai
[0, 281)	6	Shanghai, New York, Madrid, Calcutta, Dubai, Singapore

Therefore, given the additional information of α and C, we can help the company to decide how many offices they should build.

5   Sensitivity Analysis

5.1   sensitivity analysis of Model I

The purpose of our Model I was to get the change of Top10 language of the most native speakers and total speakers for the next 50 years. But if you use rankings as the dependent variable for our sensitivity analysis, we think the result we got must be insensitive, and the rank remain the same, because the Independent variables change little, and the rank is a discrete data. So we want to choose one of the model I's result data as our other dependent variable. For example, we choose the number of Mandarin second language speakers in 50 years later as our observed variables. The following analysis is to consider the sensitivity of this variable under small changes in below parameters.

1.   natural increase rate

We enter different natural growth rates of Mandarin into the program, and we get different Mandarin L2 speakers number. Here is the result:

Table 12 Sensitivity analysis on natural increase rate

natural increase rate of Mandarin	*0.34*	*0.44*	*0.54*
the number of Mandarin L2 speakers	386.34	397.27	403.18
sensitivity	0.093256475		

The Sensitivity Index shown above means that when the natural increase rate increased by 1%, the Mandarin L2 speakers will in turn increase by 0.09%. So the variation of this index has minor influ... results.

2. k_1

k_1 is the proportional coefficient of the second language growth rate and school learning rate. We enter different k_1 into the program, and we get different Mandarin L2 speakers number. Here is the result:

Table 13 Sensitivity analysis on k1

k_1	0.0025	0.003333	0.005
the number of Mandarin L2 speakers	383.14	397.27	410.5
sensitivity		0.104437788	

The Sensitivity Index shown above means that when the k_1 increased by 1%, the number of Mandarin L2 speakers will in turn increase by 0.10%. So the variation of this index has minor influence on the results.

3. k_2

k_2 is the proportional coefficient of the second language growth rate and cultural communication. we enter different k_2 into the program, and we get different Mandarin L2 speakers number. Here is the result:

Table 14 Sensitivity analysis on k2

k2	0.00001	0.0000111	0.0000125
the number of Mandarin L2 speakers	393.11	397.27	401.38
sensitivity		0.093739774	

The Sensitivity Index shown above means that when the k_2 increased by 1%, the number of Mandarin L2 speakers will in turn increase by 0.09%. So the variation of this index has minor influence on the results.

4. k_3

k_3 is the proportional coefficient of the second language growth rate and economics. we enter different k_3 into the program, and we get different Mandarin L2 speakers number. Here is the result:

Table 15 Sensitivity analysis on k3

k3	0.2	0.3	0.4
the number of Mandarin L2 speakers	387.12	397.27	408.2
sensitivity		0.079593224	

The Sensitivity Index shown above means that when the k_3 increased by 1%, the number of Mandarin L2 speakers will in turn increase by 0.079%. So the variation of this index has minor influence on the results.

5. k_4

k_4 is the proportional coefficient of the second language growth rate and technology. we enter different k_4 into the program, and we get different Mandarin L2 speakers number. Here is the result:

Table 16 Sensitivity analysis on k4

k4	0.1	0.2	0.3
the number of Mandarin L2 speakers	391.82	397.27	403.1
sensitivity		0.028393788	

The Sensitivity Index shown above means that when the k_4 increased by 1%, the number of Mandarin L2 speakers will in turn increase by 0.028%. So, the variation of this index has minor influence on the results.

6. k_5

k_5 is the proportional coefficient of the second language growth rate and media. we enter different program, and we get different Mandarin L2 speakers number. Here is the result:

Table 17 Sensitivity analysis on k_5

k_5	0.1	0.2	0.3
the number of Mandarin L2 speakers	394.88	397.27	400.12
sensitivity	0.013190022		

The Sensitivity Index shown above means that when the k_5 increased by 1%, the number of Mandarin L2 speakers will in turn increase by 0.013%. So, the variation of this index has minor influence on the results.

5.2   sensitivity analysis of model II

The purpose of Model II is to show the relationship between the change in the geographical distribution of language and the immigrants over time. The man-made parameters in model II are the natural increase rates and k_6. Below we will conduct a sensitivity analysis of these two parameters. We will use different result to measure the change of the two indices.

7.   Natural growth rate

In our model II, we set a very large number of natural growth rates. Almost every continent has its own natural rate of growth in every language. But we can't analysis every parameter. Therefore, we only change the natural growth rate of Hindi in Asia to see what changes it will bring.
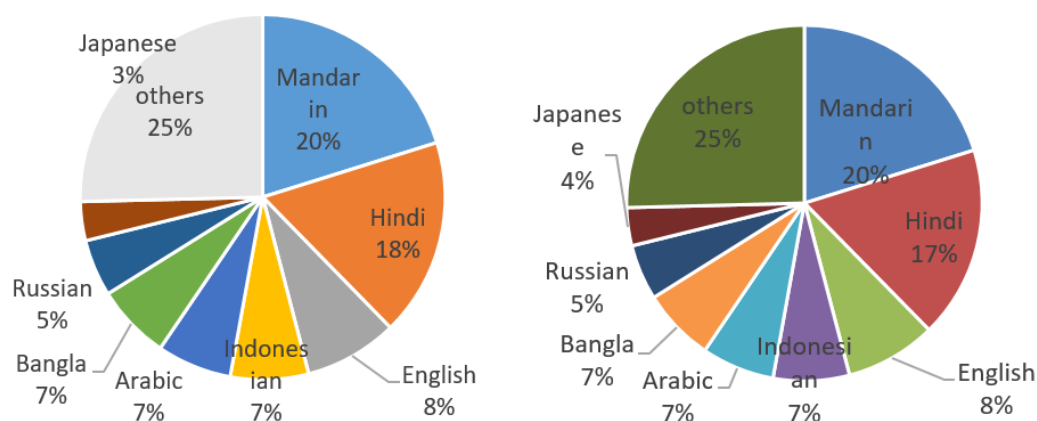
Figure 15 Sensitivity analysis on the natural growth rate

the given Hindi growth rate is 0.8%(right) and 1.2%(left)

As the pie chart shows, subtle changes in the rate of natural increase do not change the distribution in Asia. Therefore, we say that the selection of our natural growth rate is insensitive to the result.

8. k_6

k_6 is the proportional coefficient of migration to the growth of native language speakers. If we change the number of k_6, all results may change. The k_6 we set first is 0.5%. Then we slightly lower this value to 0.3% to see what changes it will produce.

Figure 16 The sensitivity analysis on migration coefficient k_6
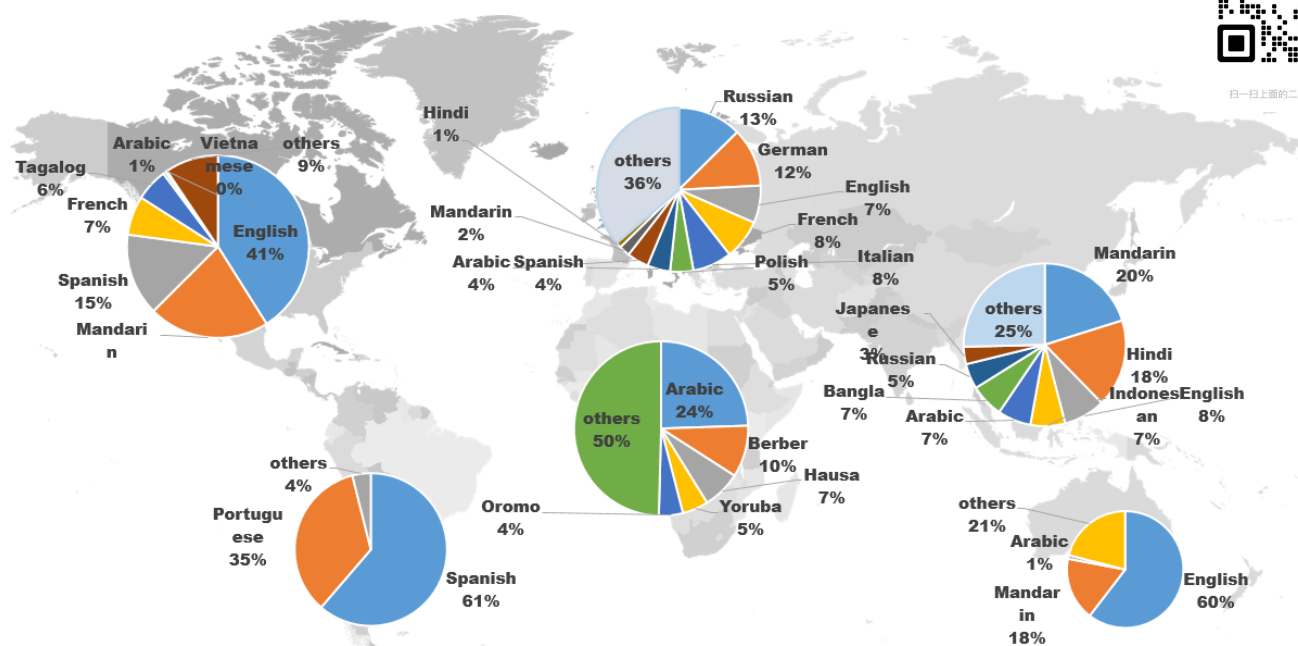
We now have a slight difference between this result and in Figure 17 the original result in Figure 18. English and French speakers in the U.S. and Canada dropped slightly. Besides, Australian English users dropped slightly. We can see that these changes are very small and do not affect the overall situation. So, we think the change of this parameter has little effect on the result.

6 Strength and Weakness

Strengths:

(1) We do plenty of research, and collect plenty of data which make our model close to reality.

(2) We consider various factors in terms of second language increase, such as school teaching, cultural migration and assimilation, the use of technology, social media and economics.

(3) We do a full sensitivity analysis.

Weakness:

(1) We do not include all influence to the total number of speakers of a language, such as the use of electronic communication for lack of data.

(2) We assume the second language speakers is in direct proportion to its influences. But the fact may be not. For example, we assume the second language speakers is in direct proportion to school teaching, but the fact may be exponential relationship or paternity relationship.

(3) Our model II does not consider age and sex ratio.

(4) We do not consider the second language in respect of geographic distributions of these languages. Because we do not find efficient data to analyze this.

7 Memo

MEMORANDUM
TO: Chief Operating Officer
FROM: Team#79002

About half the world's total population are saying that one of the top 10 languages (in order of most speakers) as their mother tongue. In addition, many people learn a language as their second language because of government promotion, schooling, neighborhood effect, social media trends, international business and immigration. Therefore, the number of native speakers and 2nd language speakers in each language is dynamically changing over time.

Prediction of top 10 languages over 50 years:

Based on the various data collected from Ethnologue, World Bank and many other resources, our team has taken into account the factors mentioned above and has built a model of the number of speakers of each language over time to predict the change in language rank over the 50-year period.

We found the top 10 languages' ranks change from Mandarin、English、Hindustani、Spanish、Arabic、Malay、Russian、Bengali、Portuguese、French in 2017 to Mandarin、English、Spanish、Hindustani、Arabic、Bengali、Portuguese、Malay、Russian、French in 2067.

As a service company, when choosing whether to set up local branches, your company should focus on how many potential local clients, and then to conduct a wide range of professional service, to achieve higher profitability.

In the meantime, a large percentage of these cities are either native or second foreign languages, and employees of more than two languages are easily admitted to the company's branch office.

Six office sites:

Our recommendations are different in the short term versus the long term because the rapid growth in both population and economic of Portuguese speaker.

In the short term, the six office sites recommended are: Shanghai, New York, Calcutta, Madrid, Dubai, and Singapore. While in the long term, the six offices sites recommended are: Shanghai, New York, Calcutta, Madrid, Dubai, and Rio de Janeiro.

As we can see, the Malay speakers' growth is slightly slower than the growth of Portuguese speakers.

The best number of offices:

To determine the best number of offices, we set up a cost-benefit analysis model, and when your company has provided us with your profitability and office-building costs, we can figure out the suitable company number. When the company's profitability and cost levels are different, the number of plants best suited for construction is different.

If the value of c-p ratio is less than 281, we think 6 offices should be built. If the value of this ratio is between 281 and 422, we think 5 offices should be built. If the value is between 422-527, 4 offices are best; if it is between 527 and 544, 3 office should be built; if> 544, we should only maintain two original offices.

Thank you for your consultation.

Best,

Team#79002

References:

[1]. https://en.wikipedia.org/wiki/List_of_languages_by_total_number_of_speakers.

[2]. Languages for the future. 2013: British Council.

[3]. Sarah Elaine Eaton, P.D., global trends in language learning in 21st century. 2010.

[4]. https://en.wikipedia.org/wiki/Malthusian_growth_model.

[5]. http://www.stolaf.edu/people/mckelvey/envision.dir/logistic.html.

[6]. https://en.wikipedia.org/wiki/List_of_most_commonly_learned_foreign_languages_in_the_United_States.

[7]. http://unicode.org/notes/tn13/.

[8]. https://www.weforum.org/agenda/2015/10/is-the-internet-killing-off-the-worlds-languages/.

[9]. http://www.internetworldstats.com/stats7.html.

[10]. https://data.worldbank.org.cn/indicator/SP.POP.TOTL.

[11]. https://faculty.washington.edu/sis/.

[12]. https://www.myaccountingcourse.com/accounting-dictionary/service-company.

[13]. https://en.wikipedia.org/wiki/List_of_most_commonly_learned_foreign_languages_in_the_United_States.

8   Appendix

8.1   data

(1) language distribution of various continents:

		Asia	Europe	North America	Latin America	Africa	Australia
1	Mandarin	900	0	0	0	0	0.62
2	English	301.6	60	347.8	0	6.5	18
3	Hindustani	0	0	0	0	0	0
4	Spanish	0	38	0	314.2	4.1	0
5	Arabic	230	0	0	0	150	0.3472
6	Malay	30	0	0	0	0	0
7	Russian	260	106	0	0	0	0
8	Bengali	0	0	0	0	0	0
9	Portuguese	1.2	10	0	201	13.7	0
10	French	0	66	11.2	0.2	0.7	0
11	Hausa	0	0	0	0	34	0
12	Punjabi	100	0	0	0	0	0
13	Japanese	120	0	0	0	0	0
14	German	0	97	0	0	0	0
15	Persian	50	0	0	0	0	0
16	Swahili	0	0	0	0	15	0
				0	0		
	total	4436	741	579	517	1216	24.8
	others	2443.2	364	200	1.6	992	5.8328

Source:

https://www.worldatlas.com/articles/the-most-spoken-languages-in-america.html

https://en.wikipedia.org/wiki/Teaching_English_as_a_second_or_foreign_language#Asia

https://en.wikipedia.org/wiki/List_of_countries_by_natural_increase

http://www.myeses.com/news/view.asp?id=3457

https://www.douban.com/note/635706471/

https://en.wikipedia.org/wiki/List_of_countries_by_population_growth_rate

https://en.wikipedia.org/wiki/List_of_countries_and_dependencies_by_population

(2) languages On the Internet[9]

Mardarin	771	20.68133	0.206813			
English	985	26.42167	0.264217			
Hindustani	0	0	0			
Spanish	312	8.369099	0.083691	Mardarin	771	20.68133
Arabic	185	4.962446	0.049624	English	985	26.42167
Malay	0	0	0	Spanish	312	8.369099
Russian	109	2.92382	0.029238	Arabic	185	4.962446
Bengali	0	0	0	Russian	109	2.92382
Portuguese	158	4.238197	0.042382	Portuguese	158	4.238197
French	108	2.896996	0.02897	French	108	2.896996
Hausa	0	0	0	German	85	2.280043
Punjabi	0	0	0	Japanese	118	3.165236
German	85	2.280043	0.0228	others	897	24.06116
Japanese	118	3.165236	0.031652			
Persian	0	0	0			
swahili	0	0	0			
others	897	24.06116				
total	3728					

(3) GDP

Mardarin	22.8	0.228			
English	28.2	0.282			
Hindustani	2.1	0.021			
Spanish	5.2	0.052			
Arabic	2	0.02		Mardarin	22.8
Malay	0	0		English	28.2
Russian	2.1	0.021		Hindustani	2.1
Bengali	0	0		Spanish	5.2
Portuguese	3.4	0.034		Arabic	2
French	4.2	0.042		Russian	2.1
Hausa	0	0		Portuguese	3.4
Punjabi	0	0		French	4.2
German	4.9	0.049		German	4.9
Japanese	5.6	0.056		Japanese	5.6
Persian	0	0		others	19.5
swahili	0	0			
others	19.5				

[7]

(4) foreigh language taught in school:[13]

Source:

https://en.wikipedia.org/wiki/List_of_most_commonly_learned_foreign_languages_in_the_United_States

http://ec.europa.eu/eurostat/statistics-explained/index.php/File:Foreign_languages_learnt_per_pupil_in_upper_secondary_education_(general),_2010_and_2015_(%25)_ET2017.png

语言	1 Mandarin	2 English	3 Hindustani	4 Spanish	5 Arabic	6 Malay	7 Russian	8 Bengali	9 Portuguese	10 French	11 Hausa	12 Punjabi	13 Japanese	14 German	15 Persian	16 Swahili
1 Mandarin	9.59	30.37	10.69	4.52	6.56	10.14	5.62	0.94	0.55	7.60	3.23	0.00	2.58	0.05	3.03	4.52
2 English	3.90%	0		50.60%	2.10%		1.40%		0.80%	12.70%				4.30%	5.50%	
3 Hindustani	9.59	30.37	10.69	4.52	6.56	10.14	5.62	0.94	0.55	7.60	3.23	0.00	2.58	0.05	3.03	4.52
4 Spanish		79.8								19.4				0.08		
5 Arabic	9.59	30.37	10.69	4.52	6.56	10.14	5.62	0.94	0.55	7.60	3.23	0.00	2.58	0.05	3.03	4.52
6 Malay	9.59	30.37	10.69	4.52	6.56	10.14	5.62	0.94	0.55	7.60	3.23	0.00	2.58	0.05	3.03	4.52
7 Russian	9.59	30.37	10.69	4.52	6.56	10.14	5.62	0.94	0.55	7.60	3.23	0.00	2.58	0.05	3.03	4.52
8 Bengali	9.59	30.37	10.69	4.52	6.56	10.14	5.62	0.94	0.55	7.60	3.23	0.00	2.58	0.05	3.03	4.52
9 Portuguese	9.59	30.37	10.69	4.52	6.56	10.14	5.62	0.94	0.55	7.60	3.23	0.00	2.58	0.05	3.03	4.52
10 French		51.6		37.5										10.9		
11 Hausa	9.59	30.37	10.69	4.52	6.56	10.14	5.62	0.94	0.55	7.60	3.23	0.00	2.58	0.05	3.03	4.52
12 Punjabi	9.59	30.37	10.69	4.52	6.56	10.14	5.62	0.94	0.55	7.60	3.23	0.00	2.58	0.05	3.03	4.52
13 Japanese	9.59	30.37	10.69	4.52	6.56	10.14	5.62	0.94	0.55	7.60	3.23	0.00	2.58	0.05	3.03	4.52
14 German		69.10		14.00						16.90						
15 Persian	9.59	30.37	10.69	4.52	6.56	10.14	5.62	0.94	0.55	7.60	3.23	0.00	2.58	0.05	3.03	4.52
16 Swahili	9.59	30.37	10.69	4.52	6.56	10.14	5.62	0.94	0.55	7.60	3.23	0.00	2.58	0.05	3.03	4.52

(5)natural growth rate:

NO.	language	natural growth rate（%）	Main coutries
1	Mandarin	0.44	China
2	English	1	Australia;Canada;Ireland;Jamaica;New Zealand;;United Kingdom;United States of America
3	Hindustani	1.4	India
4	Spanish	1	Spain;Argentina;Peru;Colombia;Venezuela
5	Arabic	1.8	Algeria;Bahrain;Egypt;Iraq;Jordan;Yemen
6	Malay	2.7	Brunei, Indonesia, Malaysia and Singapore
7	Russian	0.04	Russia, Belarus, Kazakhstan, Kyrgyzstan
8	Bengali	1.7	Bangladesh
9	Portuguese	1.76	Brazil, Mozambique, Angola, Portugal, Guinea-Bissau, East Timor
10	French	0.64	Belgium, Canada, Congo, France, Haiti,, Mali,Niger, Switzerland, Togo, and Vanuatu.
11	Hausa	2.67	Nigeria
12	Punjabi	2.4	Pakistan
13	Japanese	-0.12	Japan
14	German	0.1	Germany;Belgium;Austria;Switzerland
15	Persian	2	Iran,Afghanistan,Tajikistan
16	Swahili	3	Tanzania, Kenya, Uganda,Comoros

Source:

https://en.wikipedia.org/wiki/List_of_countries_by_natural_increase

8.2 program

1.module I

%r1 average natural popurlation of people

```
z=xlsread('data.xlsx');
x0=z(:,1);
y0=z(:,2);
r1=z(:,3);
c=[1;1;0;1;0;0;1;0;0;1;0;0;1;1;0;0];%developed or not
n=length(x0);
x=zeros(n,60);
y=zeros(n,60);
x(:,1)=x0;
y(:,1)=y0;
r2=0.0005;% increase contributed by the second language speakers
k1=1/300;% school factor
k2=1/90000;% cultural communication
k3=0.3;% international bussiness
```

```
k4=0.2;% the use of translation technology
k5=0.3;% socia media
a=xlsread('a.xlsx');%different language use percentage in school of different region
b=xlsread('b.xlsx');%comunacation percentage
q=xlsread('q.xlsx');%GDP communicated by languages
p=xlsread('p.xlsx');%language used in internet
for i=2:60
    %the next year native speakers
    x(:,i)=x(:,i-1)+native1(x(:,i-1).*c,y(:,i-1).*c,r1,r2)+native2(x(:,i-1).*(c==0),r1,2*x0);
    %the increase of second language speakers because of school use
    delta1=school(x(:,i-1),a,k1);
    %the increase of second language speakers because of cultural communication
    delta2=culturecom(x(:,i-1),b,k2);
    %the increase because of international bussiness
    delta3=interbussiness(q,k3);
    %the decrease because of the use of translation technology
    delta4=technology(k4);
    %the increase because of socia media
    delta5=media(p,k5);
    y(:,i)=y(:,i-1)+delta1+delta2+delta3+delta4+delta5;
end
xlswrite('answer.xlsx',[x(:,50),y(:,50)]);


function z=native1(x,y,r1,r2)
% the increase of native speaker in developed countries
z=r1.*x/100+r2.*y;
end


function z=native2(x,r,s)
% the increase of native speaker in developing countries
z=r/100.*(1-x./s).*x;
end


function z=school(x,a,k)
%the increase of second language speakers because of school use
%k is a properation index
n=length(x);
z=zeros(n,1);
for i=1:n
    z(i)=k*a(i,:)*x;
```

```
end
end

function z=technology(k)
% the increase because of the use of translation technology
z=-k;
end

function z=media(p,k)
z=k*p;
end

function z=culturecom(x,b,k)
%the increase of second language speakers because of cultural communication
n=length(x);
z=zeros(n,1);
for i=1:n
    z(i)=k*x(i)*b(i,:)*x;
end
end
```

2. module II

```
#include<stdio.h>
#include<math.h>
#include<cstring>
#include<iostream>
#include<algorithm>
using namespace std;
float r[50][50],sum[50];
float x[50][50],y[50],ry[50],p[50][50];

struct node{
    float people;
    int language;
}aa[50];

bool cmp(node a,node b){
    return a.people>b.people;
}
int T=50;
float k=0.003;
```

```c
int main()
{

    //x[i][j]第 i 洲说 j 语言人数
    //r[i][j]表示第 i 洲说语言 j 的自然增长率
    //y[i]第 i 洲 others 的人数
    //ry[i]第 i 洲 others  的自然增长率
    //p[i][j]第 i 洲说 j 语言的比例
    x[1][1]=150,x[1][2]=56,x[1][3]=34,x[1][4]=28,x[1][5]=26;y[1]=289.25692;

x[2][6]=897,x[2][7]=550,x[2][8]=301.625412,x[2][9]=260,x[2][10]=240,x[2][1]=230,x[2][11]=230,x[2][12]=120;
    y[2]=1377.734272;
    x[3][8]=18.175,x[3][6]=0.625,x[3][1]=0.35,y[3]=5.85;

x[4][9]=106,x[4][13]=97,x[4][14]=66,x[4][15]=65,x[4][8]=60,x[4][16]=38.5,x[4][17]=38,y[4]=304.382487;
    x[5][17]=383.4,x[5][18]=217.26,y[5]=38.34;

x[6][8]=43.240855,x[6][17]=4.508850,x[6][6]=0.932035,x[6][18]=0.601385,x[6][19]=0.30643,x[6][1]=0.539895,x[6][14]=7.2867;
    y[6]=9.9511;
    r[1][1]=0.015;r[1][3]=0.02;for(int i=1;i<=19;i++)if(r[1][i]==0)r[1][i]=0.016;
    r[2][6]=0.005;r[2][7]=0.012;r[2][8]=0.009;r[2][9]=0.002;for(int i=1;i<=19;i++)if(r[2][i]==0)r[2][i]=0.01;
    for(int i=1;i<=19;i++)r[5][i]=0.01,r[6][i]=0.004,r[3][i]=0.008;
    ry[1]=0.016;ry[2]=0.001;ry[3]=0.008;ry[5]=0.001;ry[6]=0.004;

    printf("time      continent    language          people(million)      others(million)\n");
    for(int t=1;t<=T;t++){
            for(int i=1;i<=6;i++){
                sum[i]=0;
            for(int j=1;j<=19;j++)
                sum[i]=sum[i]+x[i][j];
                sum[i]+=y[i];
            for(int j=1;j<=19;j++)
                p[i][j]=x[i][j]/sum[i];
            }
        for(int i=1;i<=6;i++){
            for(int j=1;j<=19;j++)
        x[i][j]=x[i][j]*(1+r[i][j]);
        y[i]=y[i]*(1+ry[i]);
        }
```

```
        x[4][6]+=0.36+0.36*k*p[4][6];x[4][1]+=0.72*(1+k*p[4][1]);
        x[4][7]+=0.18*(1+k*p[4][7]);x[4][8]+=0.18*(1+k*p[4][8]);
        x[6][6]+=0.48*(1+k*p[6][6]);x[6][17]+=0.24*(1+k*p[6][17]);
        x[6][18]+=0.12*(1+k*p[6][18]);x[3][6]+=0.12*(1+k*p[3][6]);
//      x[2][6]-=8.4;x[2][1]-=3.6;
//      x[2][7]-=1.8;x[2][8]-=1.8;
//      x[5][17]-=2.4;x[5][18]-=1.2;
//      x[1][1]-=4.8;


        for(int i=1;i<=6;i++){
            for(int j=1;j<=19;j++){
                aa[j].people=x[i][j];aa[j].language=j;
            }

            sort(aa+1,aa+20,cmp);

        if(t==20||t==50){

            for(int j=1;j<=19;j++)

printf("%d         %d                 %d                 %lf                 %lf\n",t,i,aa[j].language,aa[j].people,y[i]);
            printf("\n");
        }
        }
        }

        return 0;
}
```