WEP Warwick Evans Publishing

International Journal of Social Sciences and Public Administration

ISSN: 3005-9836 (Print), ISSN: 3005-9585 (Online) | Volume 4, Number 2, Year 2024
DOI: https://doi.org/10.62051/ijsspa.v4n2.25
Journal homepage: https://wepub.org/index.php/IJSSPA/index

Machine Learning-based Analysis of Chinese Population Structure

Meng Wu*

NorthWest University, Xi'an, Shaanxi 710000, China
*E-mail: wmzzzm11@aliyun.com

ABSTRACT

Using in-depth data analyses and scientific methods, this paper provides an exhaustive discussion of the gender composition, urban-rural distribution, and age levels of the population in various provinces of China, thus revealing the far-reaching impact of the evolution of the population structure on social and economic progress. Firstly, using data visualisation techniques combined with K-mean cluster analysis, this study reveals the phenomenon of gender imbalance and proposes a series of countermeasures, including strengthening education on gender equality, promoting gender-equal social participation, and paying attention to gender-selective behaviours. Secondly, the study provides an in-depth analysis of the trend of population migration in the process of urbanisation and the problems derived therefrom, and advocates the government to adopt scientific planning and policy guidance measures in order to narrow the development gap between urban and rural areas, and to promote integrated urban and rural development. Finally, the paper provides a detailed analysis of the trend of population aging and insights into the challenges it poses to the pension service system, health care system and labour market, suggesting the formulation of a more flexible fertility policy and the vigorous development of the pension service industry to cope with the challenges posed by aging.

KEYWORDS

Population Structure; K-means Cluster Analysis; Sex Ratio; Population Migration.

1. INTRODUCTION

1.1. Research Background

China's demographics are undergoing significant changes, mainly in a number of ways. First of all, the growth of the total population has slowed down and gradually stabilized, showing a low-speed growth trend. Secondly, the birth rate of newborns continued to decline, and the effect of the two-child policy was not obvious. Third, the aging of the population has intensified, the proportion of the elderly population has risen, and the pressure on the elderly has increased. Fourth, the process of urbanization is rapid, but the transformation from rural to urban is slow. Fifth, the education level of the overall population has improved, and the demographic dividend has changed from quantitative to qualitative. Sixth, the size of households has decreased, and the average number of people per household has decreased. Finally, the gender ratio is gradually becoming more even, but the male population still outnumbers the female population. General Secretary Xi Jinping stressed that demographic change is crucial to economic development, and detailed statistics are needed to support high-quality economic development planning. Research shows that the Chinese population and the economy entered the new normal at about the same time, and the two are closely related. Therefore,

this paper will deeply analyze the trend of demographic change, provide a basis for formulating scientific population policies, and promote sustained and stable economic growth.

1.2. Research Significance

The theoretical significance of this paper is to analyze the panel data of 28 provinces (autonomous regions and municipalities directly under the central government) in China from 2013 to 2022, and divide the population structure into five categories: age, gender, education, family size, and urban and rural areas, and study the impact of population structure on the economy. This approach complements the limitations of previous studies using cross-sectional data or single-province time series data, and provides a more representative analysis of national and regional differences. The practical significance lies in the fact that through data analysis, the trend of demographic change can be clarified, so as to provide a direction for the country to improve the population policy and promote sustained and stable economic growth. Since the implementation of the family planning policy in 1971, the growth rate of the Chinese population has slowed, but it has also brought about the problem of aging. This paper uses Python software to analyze population structure changes and put forward scientific policy recommendations.

2. DATA COLLECTION AND PROCESSING

2.1. Data Collection

This paper studies and discusses the structure of Chinese population from three dimensions: population age structure, population urban-rural structure, and population gender structure, so this paper establishes an index system, as shown in Table 1 below.

Table 1. indicator systems

Primary index	Secondary indicators	Three-level index	Unit
Population structure	Overall structure of population	Total population	Ten thousand
	Gender structure of population	Male population	Ten thousand
		Male population proportion	%
		Female population	Ten thousand
		Proportion of female population	%
	Urban-rural structure of population	Urban population	Ten thousand
		Proportion of urban population	%
		Rural population	Ten thousand
		Proportion of rural population	%
	Age structure of population	Population aged 0-14	Ten thousand
		Proportion of population aged 0-14	%
		Population aged 15-64	Ten thousand
		Proportion of population aged 15-64	%
		Population over 65 years old	Ten thousand
		Proportion of population over 65 years old	%

This paper takes 28 provincial-level administrative regions (excluding Xinjiang, Tibet, Shanghai, Hong Kong, Macao and Taiwan) as research samples, considers the feasibility and timeliness of the data, and selects 2013-2022 as the time span for research. In order to ensure the reliability of the data and the subsequent model training process, the data needed in this paper were obtained from China Statistical Yearbook and provincial statistical yearbook. In addition, this paper also collected the birth rate and death rate of the entire population since 2001 as a supplement for research.

2.2. Data Processing

Because there are some missing years and indicators in the statistical yearbooks of each province, this paper fills in these missing values by linear interpolation method. If we know the point,, interpolated at, the value formula is as follows,

$$y = y_0 + \frac{y_1 - y_0}{x_1 - x_0}(x - x_0)$$

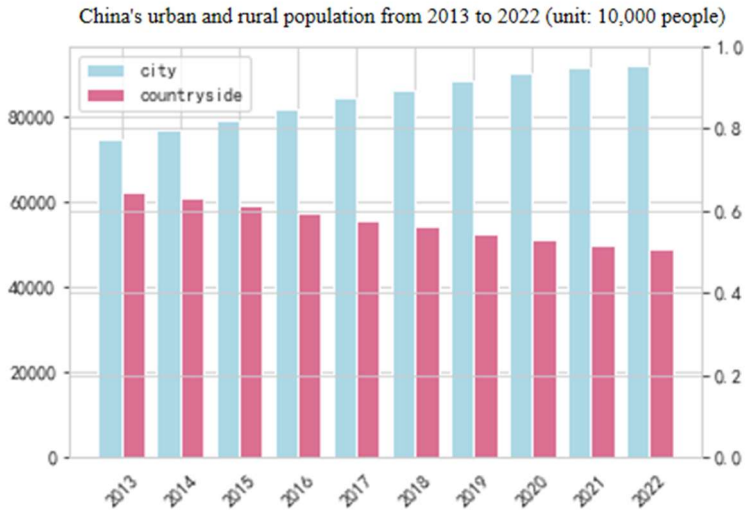
In order to ensure the unity of index attributes, this paper uses the range normalization method to normalize the data in order to eliminate the dimensional influence. The specific formula is as follows:

$$x_i^{'} = \frac{x_i - \overline{x}_i}{S_i}$$

Where, x_i represents the actual value of the i indicator, $\overline{x}_i$ is the mean value of the i indicators, S_i is the variance of the i indicators, and $x_i^{'}$ is the value after the standard of the i indicator.

3. NATIONAL AND PROVINCIAL POPULATION STRUCTURE VISUALIZATION ANALYSIS

3.1. Visualization of Population Urban-rural Structure



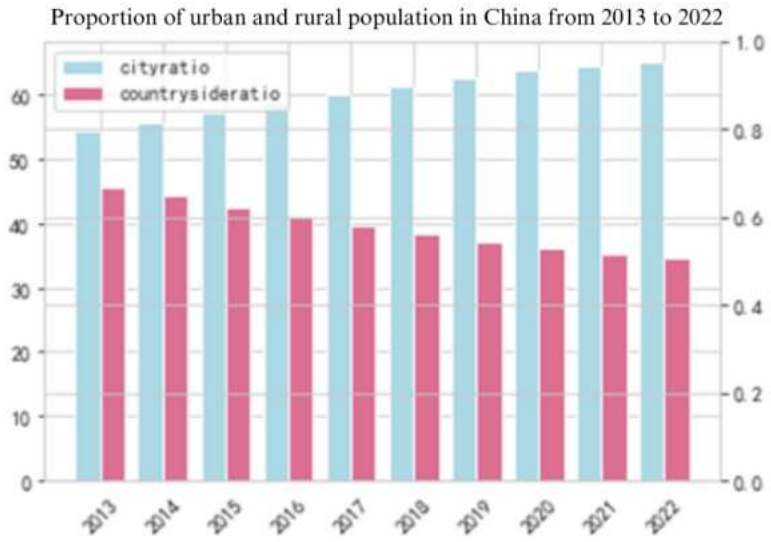


Figure 1. Visualization of population urban-rural structure

3.2. Visualization of Population Gender Structure

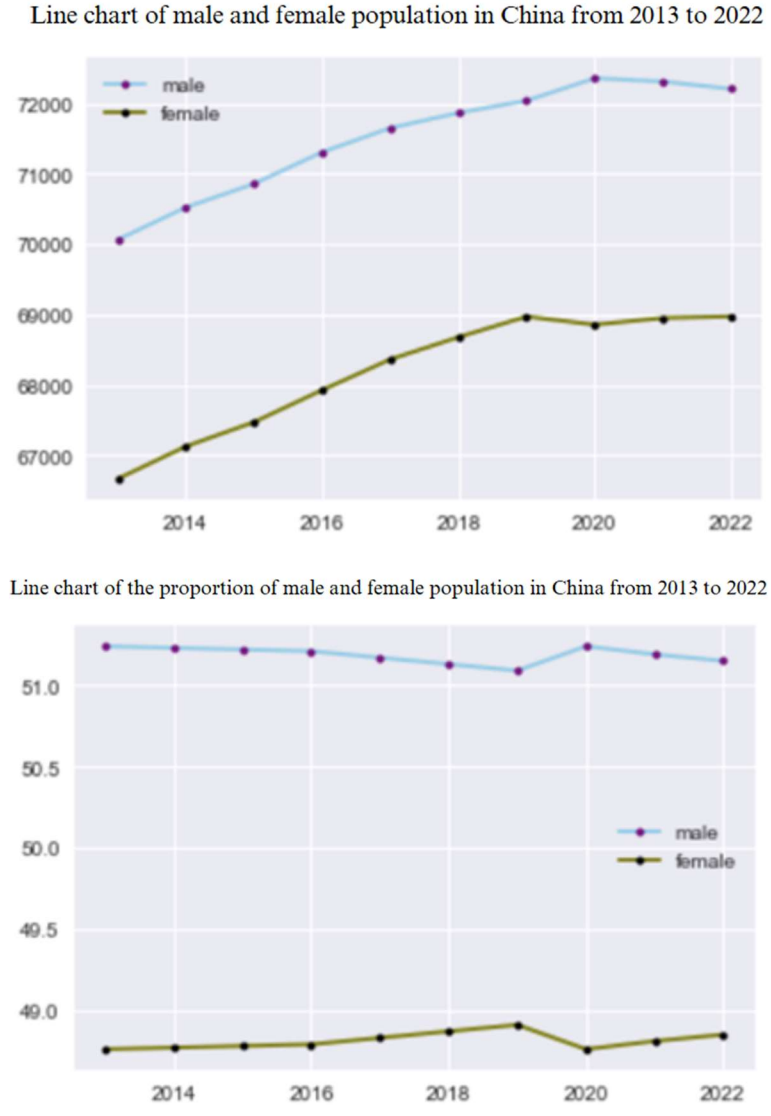


Figure 2. Visualization of gender structure of population

3.3. Visualization of Population Age Structure

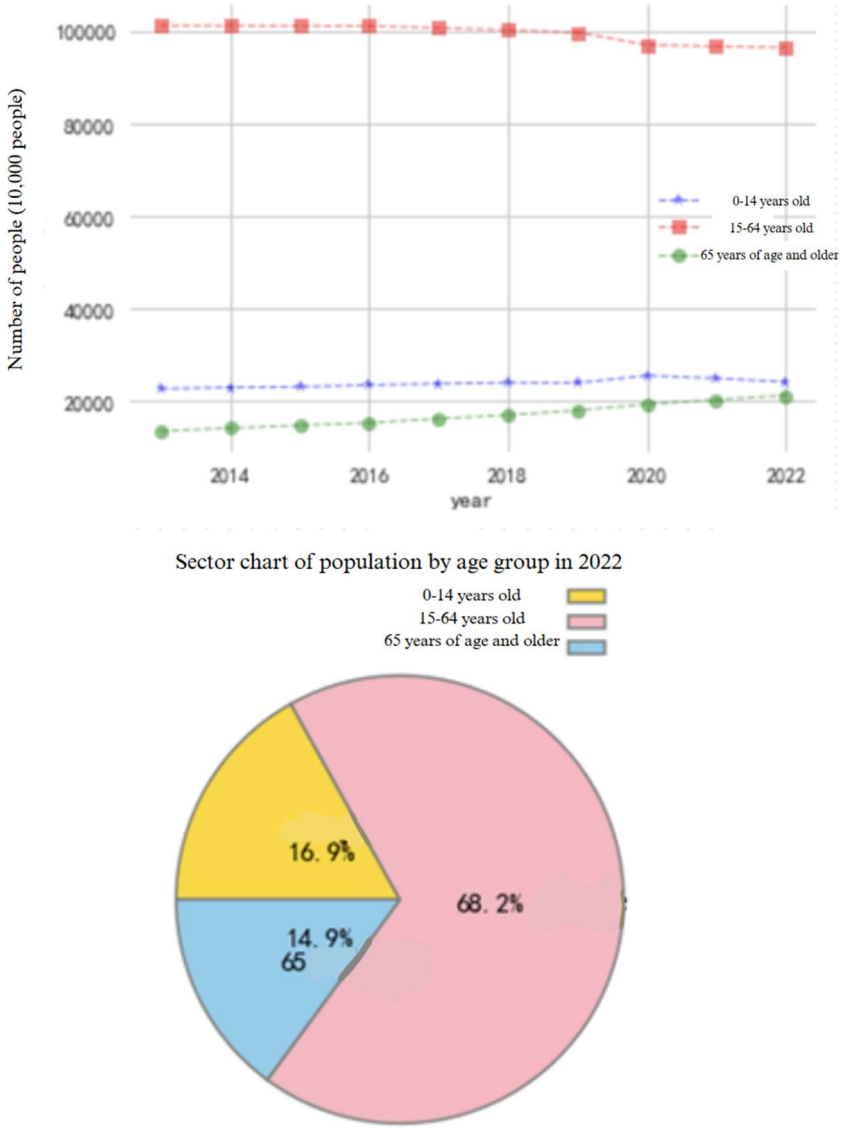


Figure 3. Visualization of population age structure

From 2001 to 2009, the national birth rate in China decreased gradually, and fluctuated and rose from 2009 to 2016, while from 2016 to 2022, the birth rate decreased rapidly, and even the death rate exceeded the birth rate in 2022. At the same time, from 2001 to 2022, the national mortality rate changed steadily, showing a slow upward trend. Observing the changes of urban and rural population structure, from 2013 to 2022, the urban population gradually increased, while the rural population gradually decreased, and the gap between them continued to widen. In the past two years, the urban and rural population tended to be stable, and the ratio of urban and rural population was close to 2:1. The total population of the country and the population of both men and women have gradually increased, with men always outnumbering women by about 2%. From 2013 to 2022, the number of people aged 15-64 decreased slowly, and the number of elderly people over 65 increased continuously, which will be the same as the number of people aged 0-14. The population aged 0-14 is relatively stable. During 2016-2020, due to the policy of having two children and three children, it increased slightly, but it decreased recently. The age structure of the population in 2022 is olive-shaped, with the population aged 15-64 accounting for the largest proportion, and the number of elderly people over 65 is expected to exceed the number of people aged 0-14.

4. BASED ON K-MEANS CLUSTER ANALYSIS OF PROVINCIAL POPULATION STRUCTURE

In this paper, data from 28 provinces in 2017 and 2021 are selected for clustering respectively to observe which provinces have the same type of population structure, and to observe the changes in the population structure classification of each province from 2017 to 2021.

There are three methods to select K clustering centers. The first is the elbow method: the distance between each data point and the cluster center is measured by the SSE error square sum, and the elbow point is found to determine the optimal number of clusters. The second is the elbow method: Determine the optimal number of clusters by calculating the average deviation of each cluster. The third is the contour coefficient method: the effect is evaluated by combining the cohesion and separation degree of clustering, the range is [-1,1], and the larger the value, the better the effect. These methods help improve the effectiveness of K-means clustering.

In this paper, the above three methods are comprehensively used to analyze the data of 2017, and the result K value should be 3. The analysis of data in 2021 shows that K is 4. The result is shown below. (Where the purple dot is the cluster center).

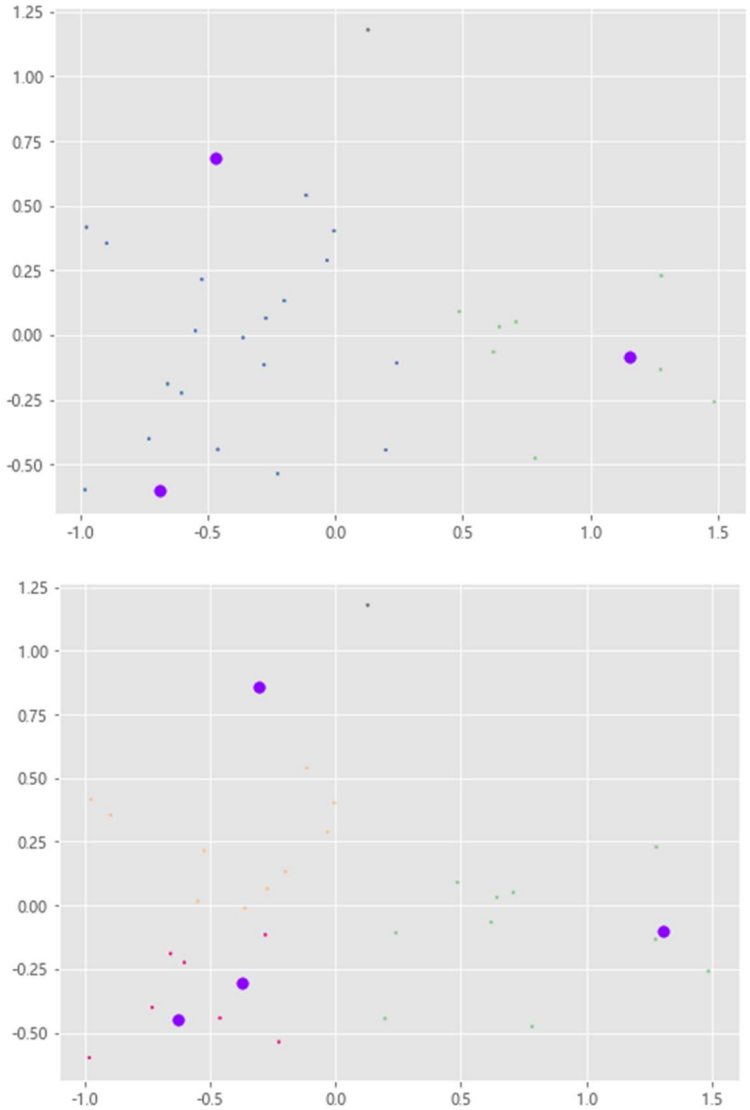


Figure 4. Clustering results (left: 2017, right: 2021)

This paper takes 2017 as an example to select one province from each category and carry out visual analysis respectively, which are Shaanxi Province, Anhui Province and Tianjin City, as shown in the figure below.

4.1. Shaanxi Province

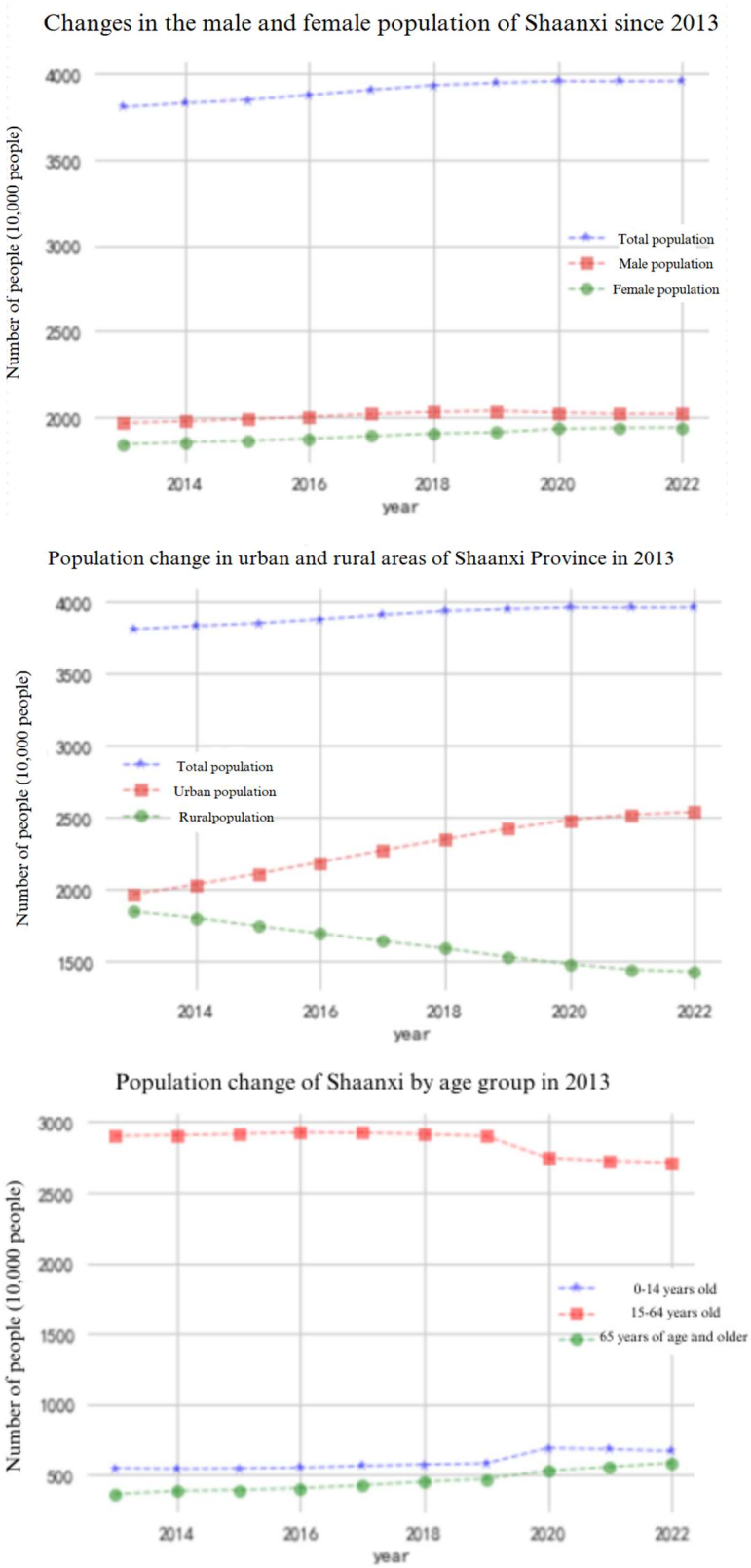


Figure 5. Visual analysis of Category 1 taking Shaanxi Province as an example

Over the past decade, the total population has witnessed a steady growth, with the male population consistently outnumbering the female population. This demographic trend is accompanied by a rise in urban population and a decline in rural population. Notably, both the population aged 0-14 and those over 65 have increased, although the growth rate among the latter group has surpassed that of the former. Additionally, the population aged 15-64 has experienced a continuous decline since 2019.

4.2. Anhui Province

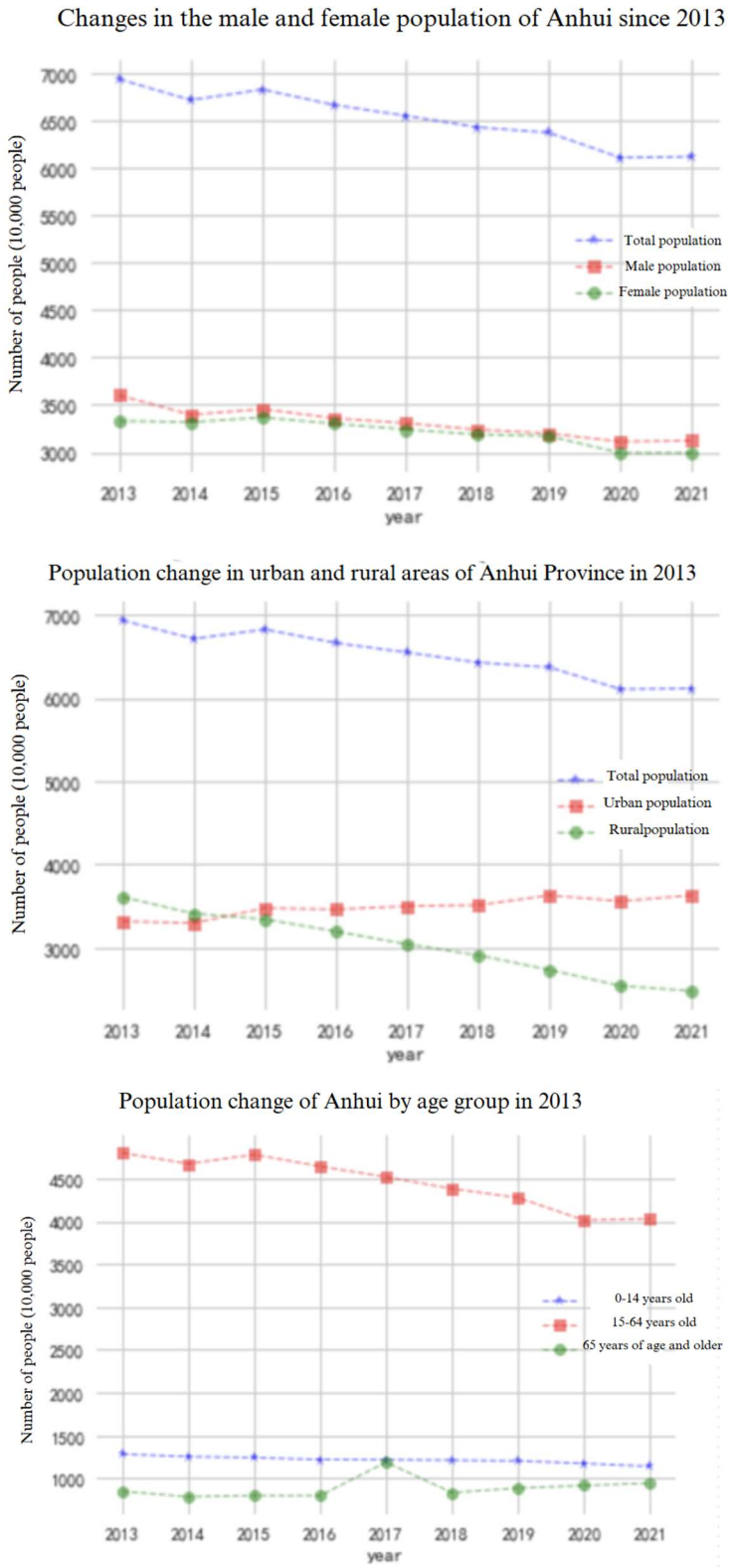


Figure 6. Visual analysis of Category 2 taking Anhui Province as an example

Over the last ten years, there has been ongoing fluctuation in Anhui province's total population with an overall downward trend. Gender distribution remains relatively balanced while urban areas experience growth and rural areas witness decline. Notably, urban migration rates are slower compared to those in rural regions—suggesting inter-provincial movement within Anhui's populace. Moreover, there's a notable increase in individuals aged over 65 alongside a steady decrease in those aged 0-14; meanwhile, individuals between ages 15-64 have experienced almost a decade-long decline mirroring that seen in overall demographics.

4.3. Tianjin City

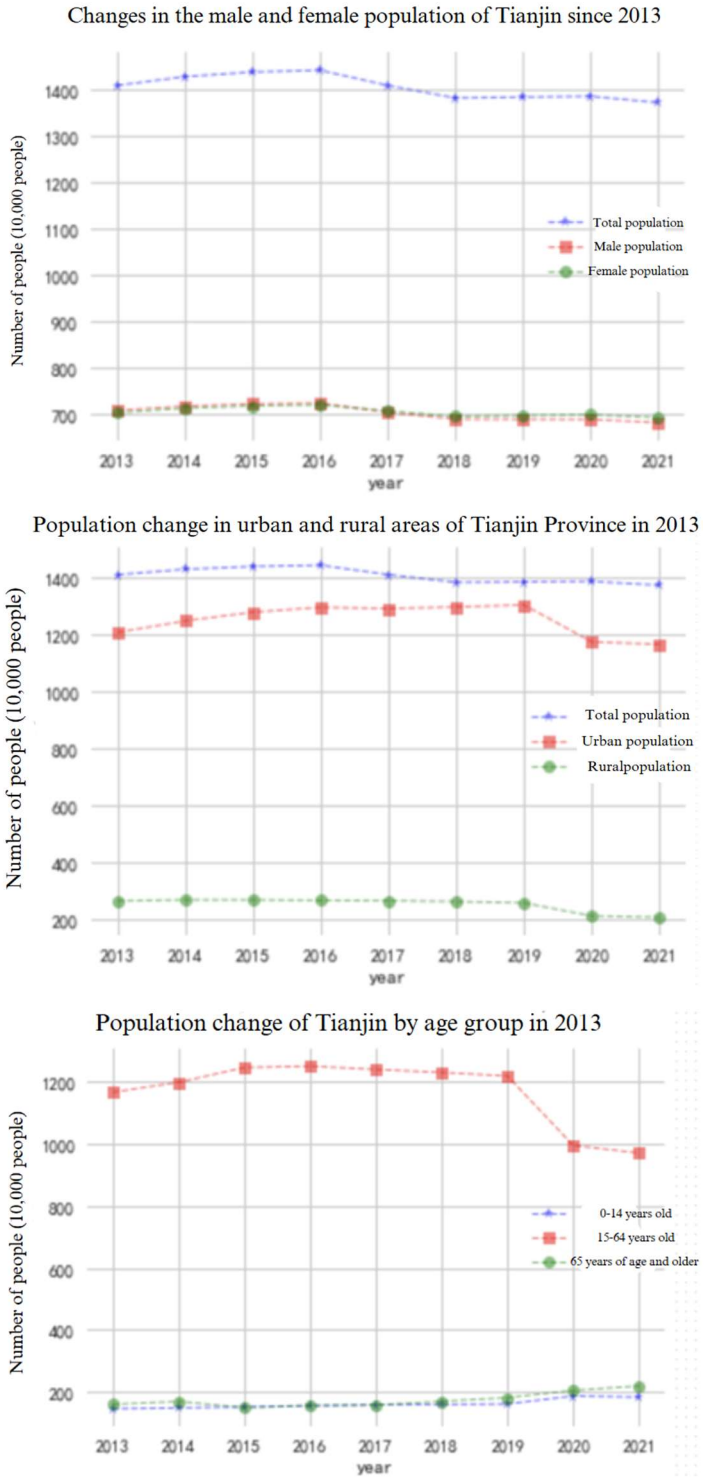


Figure 7. Category 3 Visual analysis taking Tianjin as an example

The overall population has experienced fluctuations over the past decade, yet it has predominantly remained stable, with a near-equal distribution between male and female populations. The urban population consistently outweighs the rural population, and from 2013 to 2019, the urban population exhibited gradual growth before experiencing a sudden decline in 2019 possibly attributed to an epidemic outbreak; subsequently stabilizing. Conversely, the rural population saw minimal change during this period. It was only in 2019-20 that a decline occurred. The age group of 15-64 witnessed slow growth from 2013 to 2019 but sharply declined in 2019, while both the age groups of 0-14 and over 65 showed similar rising trends at approximately equal rates.

5. CONCLUSION AND SUGGESTIONS

The ongoing demographic transformation in China is leading to a significant shift within the socio-economic framework. This pivotal transition is characterized by a slowing population growth, a decrease in average family size, an acceleration of the aging population, rapid urbanization, improved educational standards, optimized familial structures, and gender balance. These phenomena go beyond mere demographic changes and represent significant milestones in societal advancement and economic evolution.

Confronted with this monumental transition, the government has implemented a range of forward-thinking measures aimed at addressing emerging challenges. However, ensuring that population development aligns sustainably with economic and social progress requires intensified efforts across multiple domains. In terms of gender demographics, it is crucial to continue advocating for gender parity and expedite efforts to bridge the gender gap through educational initiatives, informational campaigns, and legislative interventions. Simultaneously, attention must be paid to gender selection dynamics while endorsing policies that support a balance between professional and domestic spheres to strengthen familial policies' effectiveness.

In the context of age demographics, it is essential for the government to conduct careful assessments and adjust fertility policies promptly to moderate the increase in the aging population while refining labor force composition. Furthermore, increasing investment in geriatric services and medical reforms are imperative as well as cultivating industries catering to the elderly while enhancing social security provisions for seniors due to escalating demands from this growing demographic.

Regarding urban-rural demographic disparities necessitate fostering institutional reforms and policy endorsements that facilitate robust urbanization's orderly evolution. The government could expedite overhauls of household registration systems & land management systems; elevate urban public service quality; integrate urban & rural areas thereby reducing developmental disparities between these regions. Concurrently reinforcing rural reforms including modernizing agricultural practices are vital for elevating income levels among agrarian populations.

Policymakers also need to consider how demographic changes impact various societal sectors such as education employment healthcare & elderly care devising corresponding policies & measures ensuring comprehensive harmonious societal progression For instance within education sector there's need fortify lifelong learning elevate pedagogical standards accommodate novel exigencies engendered by demographic shifts Employment sector should optimize labor market offer abundant occupational opportunities vocational training surmount nascent challenges Healthcare sector needs reinforce public health infrastructure enhance medical services address medical exigencies precipitated by aging populace.

In conclusion China's demographic transformation isn't just socio-economic mutation but represents formidable test nation's governance acumen Government must persist proactive stance adopt more efficacious array policy measures ensure development synchronized with economic social development actualizing aspiration sustainable development This endeavor transcends purview government alone demanding collective wisdom fortitude engagement entire societal fabric.

REFERENCES

[1] Cai Fang. Labor shortage: should we save for a rainy day[J]. China Population Science,2005,(06):1116+95.

[2] Chen Wei,Du Yang,Hou Dongmin. Is it a demographic dividend? Or is it a demographic problem? [J]. Population Research,2007,(02):41-48.

[3] Zhang Xiaoqing. Study on the Impact of Population Age Structure on Regional Economic Growth[J]. China Population-Resources and Environment,2009,19(5):4.

[4] ZHANG Xiuwu,LIU Chengkun,ZHAO Xindong. Whether population age structure affects economic growth through human capital - A test based on the mediating effect[J]. China Soft Science,2018,(07):149-158.

[5] Chen Daiyun,Chen Xi. Research on the problem of population age structure change and its effect--Based on the information of population census in Shandong Province[J]. Shandong Social Science,2020,(11):103-110.

[6] TIAN Xueyuan,HU Weiliu,YANG Yongchao. Population Aging and Economic and Technological Progress in Japan: A Report on a Study Tour to Japan and Academic Exchange[J]. China Population Science,1990,(06):59-62.

[7] Li Xing,M.W.LukeChan.A study on the relationship between demographic change and economic growth in China based on SYS-GMM[J]. Statistical Research,2012(4):5.

[8] Zhang Peng. Study on the impact of population age structure transition on economic growth in China[D]. Nankai University,2013.

[9] CHENG Wanjing,WU Kangping,TIAN Yajun. The impact of labor force age structure on China's economic development[J]. Technical Economy,2019,38(01):104-111.

[10] Liu HY. Research on the relationship between human capital "homogenization" and the quality of China's economic growth[J]. Management World,2004(11):7.

[11] XIAO Yiping,YANG Yanlin. Research on the impact of population age structure change on economic growth[J]. Population Research,2017,41(4):13.

[12] SUN Aijun,LIU Shenglong. Analysis of economic growth effect of demographic change[J]. Population and Economy,2014(1):10.

[13] WANG Weiguo,LIU Feng,HU Chunlong. Fertility policy, population age structure optimization and economic growth[J]. Economic Research,2019,54(01):116-131.

[14] WANG Wei,LIU Yufei,XU Yan. The age structure of the working population and the dynamic evolution of labor productivity in China[J]. Academic Monthly,2019,51(08):48-64.

[15] Wang Jinying. Accelerating population urbanization to achieve sustainable economic growth in China[J]. Journal of Hebei University (Philosophy and Social Science Edition),2004,(02):50-54.

[16] Shen Kunrong,Jiang Rui. An empirical study on the impact mechanism of urbanization on economic growth in China[J]. Statistical Research,2007,24(6):7.

[17] Sun Wenkai. Analysis of the relationship between urbanization and economic growth--An analysis of Chinese characteristics[J]. Economic Theory and Economic Management,2011,000(004):33-40.

[18] Huang Xiangmei,Xia Haiyong. The dynamic relationship between population urbanization and economic growth and industrial structure--Taking Jiangsu Province as an example[J]. Urban Issues,2012(5):6.

[19] Hong Yeying. An empirical study on the relationship between population urbanization, economic growth and industrial structure[J]. Business Times,2013,(08):16-17.

[20] Kong Xiaoni,Deng Feng. Empirical study on the mechanism of population urbanization-driven economic growth[J]. Population and Economy,2015(6):11.

[21] WANG Yongjun,ZHANG Donghui. The impact of new urbanization on economic growth from the perspective of urban-rural integrated development[J]. Gansu Social Science,2020,(03):177-184.

[22] Liu Qi. Study on the impact of rural population structure change on local economic development in Tumet Right Banner of Baotou City[D]. Inner Mongolia Agricultural University,2012.

[23] Shandan. Population structure analysis and economic development in Shenzhen[J]. Northwest Population,2002(4):3.

[24] Yu Xuejun. China's demographic transition and the "period of strategic opportunity"[J]. China Population Science.