



DEMOGRAPHIC ANALYSIS OF POPULATION OF UZBEKISTAN USING DATA SCIENCE TOOL PANDAS

Meliboev Azizjon Ikromjon o'g'li,

Department of Digital Technology and Mathematics, Kokand University.

Asqarov Elbek Erkinjon o'g'li,

Department of Digital Technology and Mathematics, Kokand University.

Azizbek Naraliyev Yakubjon o'g'li

MAQOLA HAQIDA	ANNOTATION
<p>Qabul qilindi: 24-mart 2025-yil</p> <p>Tasdiqlandi: 26-mart 2025-yil</p> <p>Jurnal soni: 14</p> <p>Maqola raqami: 62</p> <p>DOI: https://doi.org/10.54613/ku.v14i.1175</p> <p>KALIT SO'ZLAR/ КЛЮЧЕВЫЕ СЛОВА/ KEYWORDS</p> <p>Data science, Pandas tool, EDA, demographic analysis</p>	<p>The purpose of this study is to use the data science tool Pandas to perform a demographic analysis of the Uzbek population. We will investigate important demographic parameters as age distribution, gender ratio, population density, urban vs. rural population distribution, and population growth trends over time by utilizing Pandas' data manipulation and analysis capabilities. In order to produce insightful findings, the analysis will entail importing population databases, cleaning and preparing the data, and doing exploratory data analysis (EDA). The study will provide a thorough grasp of Uzbekistan's demographic structure by identifying patterns and trends within the country's population through statistical metrics and visualization.</p>

Introduction. An essential tool for comprehending a population's makeup, distribution, and trends is demographic analysis. Demographic information is essential for developing policy in sectors like urban planning, healthcare, education, and economic growth in the Central Asian nation of Uzbekistan, which has a rich history and diverse cultural legacy. Governments and organizations can make well-informed decisions to promote social well-being and sustainable growth by looking at the population structure.

The robust data science tool Pandas, a Python-based library, is used in this study to conduct a thorough demographic analysis of the people of Uzbekistan. Pandas is perfect for examining big datasets with demographic characteristics since it offers a strong foundation for data manipulation, cleaning, and analysis. Gaining a better understanding of Uzbekistan's population dynamics including age distribution, gender ratio, population density, urban versus rural populations, and historical demographic trends is the goal of this investigation.

Through the use of a data-driven methodology, this analysis will assist in highlighting changes over time, visualizing and comprehending important demographic patterns, and identifying possible future demographic concerns. Policymakers, scholars, and organizations looking to meet the needs of Uzbekistan's expanding and changing population may find the findings to be very helpful.

Exiting works

Social scientists, economists, and politicians have long been interested in demographic analysis because it offers important insights into population trends and structures. The population of Uzbekistan has been the subject of numerous studies, with an emphasis on a range of demographic factors such as age distributions, migration trends, and fertility rates. In Uzbekistan, demographic research has historically been carried out using statistical methods and census data. However, the subject of demographic analysis has evolved significantly in recent years because to the growing availability of big datasets and sophisticated data analysis tools, such Python's Pandas module [1].

Numerous studies have examined demographic statistics at the regional and global levels using data science approaches. For instance, Uakcaxova [2] used census and survey data to describe the state of demographic concerns in Uzbekistan. Similar to this, research on migratory patterns in South and Central Asia, such as that conducted by Narasimhan et al. [3], highlighted significant demographic shifts brought about by historical changes in the region. An old genetic continuity connecting groups linked to early hunter-gatherers from

Southeast Asia and Iran is the primary ancestral origin of present South Asians.

The regulations pertaining to birth, death, marriage, annulment, and the renewal of spouses and families as essential elements of these procedures are examined by Khasanova et al. [4]. They also examine the general rules of population renewal and the ways in which societal factors impact them.

Based on official figures, several papers released by the United Nations Development Programme (UNDP) [6] and the World Bank [5] have offered comprehensive demographic profiles of Uzbekistan. Important indicators including life expectancy, fertility rates, and urbanization have all been examined in these research. Although these reports give a broad overview of Uzbekistan's demographic composition, using data science tools like Pandas allows for a more thorough examination of these datasets, leading to more detailed insights on population trends over time.

This work intends to advance research by using Pandas to provide a more effective and scalable method of demographic analysis. Data science techniques allow for more flexible data manipulation, processing of big datasets, and the use of dynamic visualizations for a better understanding of patterns, in contrast to traditional methods that frequently rely on predetermined statistical models. By utilizing Pandas' capabilities, this project processes and analyzes Uzbekistan's demographic data, allowing for a more thorough examination of trends including age distribution, gender ratios, and urban-rural population dynamics. As a result, it adds to the expanding corpus of work on the application of data science to population research.

Analysis method

The data is a vital tool for analysis and gaining a variety of insights. We made use of statistical data gathered by the President of the Republic of Uzbekistan's Statistics Agency [8]. The aforementioned dataset has demographic statistics from several places across a number of years, most likely from 2010 to 2023. In the dataset, each row represents a particular year, while the columns stand for different districts or areas. Each region's population numbers for a specific year are represented by the numerical values. The graphic shows the population trends in each region over a given time frame, most likely 2010–2023. The graph provides a comparative examination of the changes in the population sizes of different regions throughout time, with each line representing the population trend for a particular region.

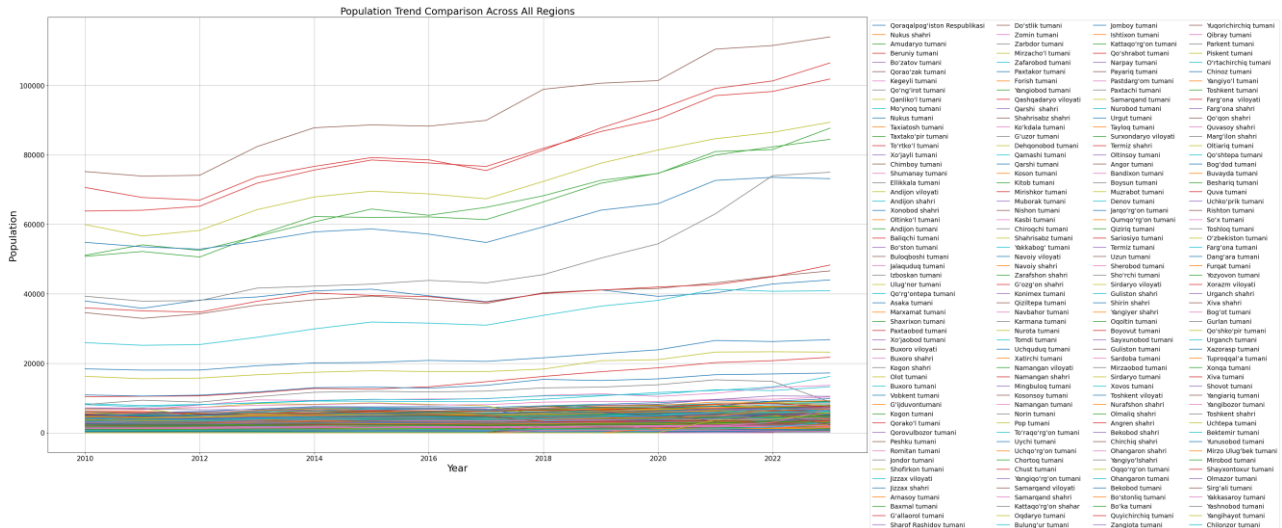


Figure 1. All regions representation.

Years stands for the time period that the data covers, which begins about 2010 and ends around 2023. Every point on the x-axis represents a distinct year. The population of each region at different times is shown by values in the population shows, which range from zero to over 100,000. Leading Population Areas Numerous places, especially those in the upper portion of the graph, exhibit noticeably larger populations. These places may be densely inhabited or significant urban hubs. For instance, considering their high population trends, areas like "Samarqand viloyati," "Farg'ona viloyati," and "Qashqadaryo viloyati" are probably well-known.

Comparative Analysis by Region With some areas exhibiting consistent growth, others exhibiting fluctuations, and a few remaining relatively stable, there are notable disparities in the population

patterns. Rapid urbanization or expansion is indicated by the sharp population increases in some areas over time. Every region included in the dataset is listed in the lengthy legend on the right-hand side, with each location being assigned a unique color that matches the plot's lines. The legend aids in recognizing the patterns of particular regions of interest due to the vast number of locations.

Several of the regions near the graph's lower end, known as Bottom Population Regions, have smaller populations. These areas most likely reflect rural or less inhabited places. This chart is helpful for policy research, urban planning, and demographic studies since it successfully offers a thorough picture of population dynamics across several locations over time for scholarly discussion.

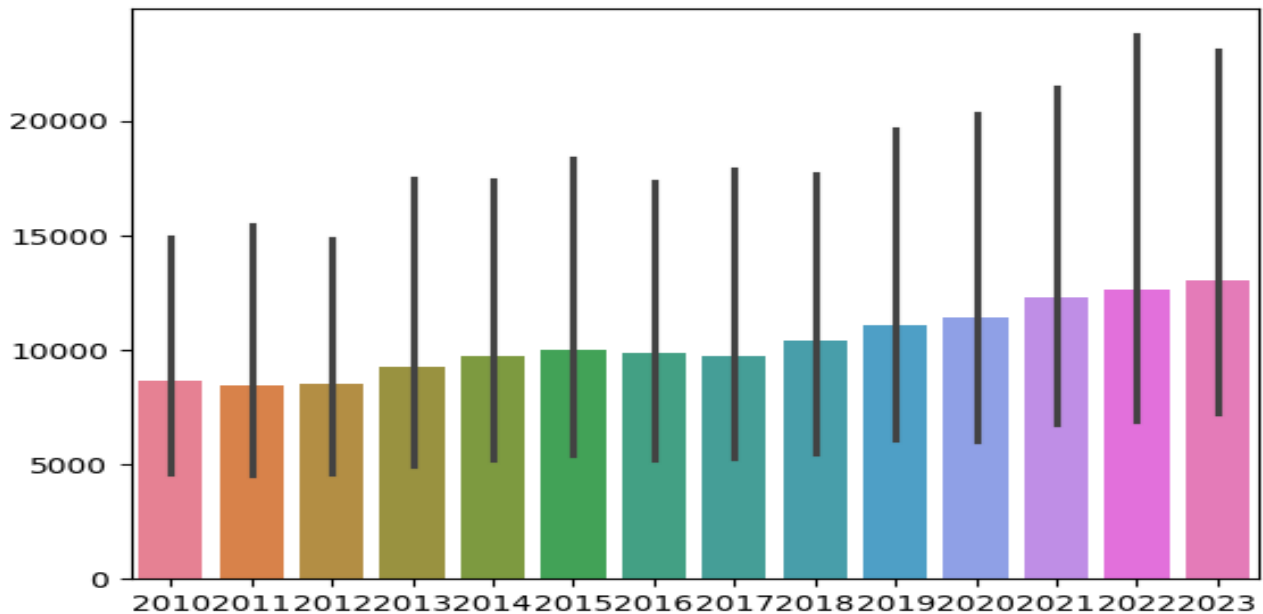


Figure 2. Bar plot across years.

It seems like a bar plot of population values over time (perhaps from 2010 to 2023) in Figure 2. The height of the bars reflects the amount of the population, with each bar representing the population for a particular year. From 2010 to 2023, the years are shown chronologically on the x-axis. Every year is symbolized by a bar. The population size is represented by the y-axis, which has values between 0 and just over 20,000. This shows the annual population for a certain area or region.

Growth of the Population Consistent population growth points to potential causes including migration, urbanization, or natural population expansion. There may have been differences in the accuracy of the population data or other demographic shifts over those years, as evidenced by some years (such between 2016 and 2020) having somewhat higher error bars or growth dips.

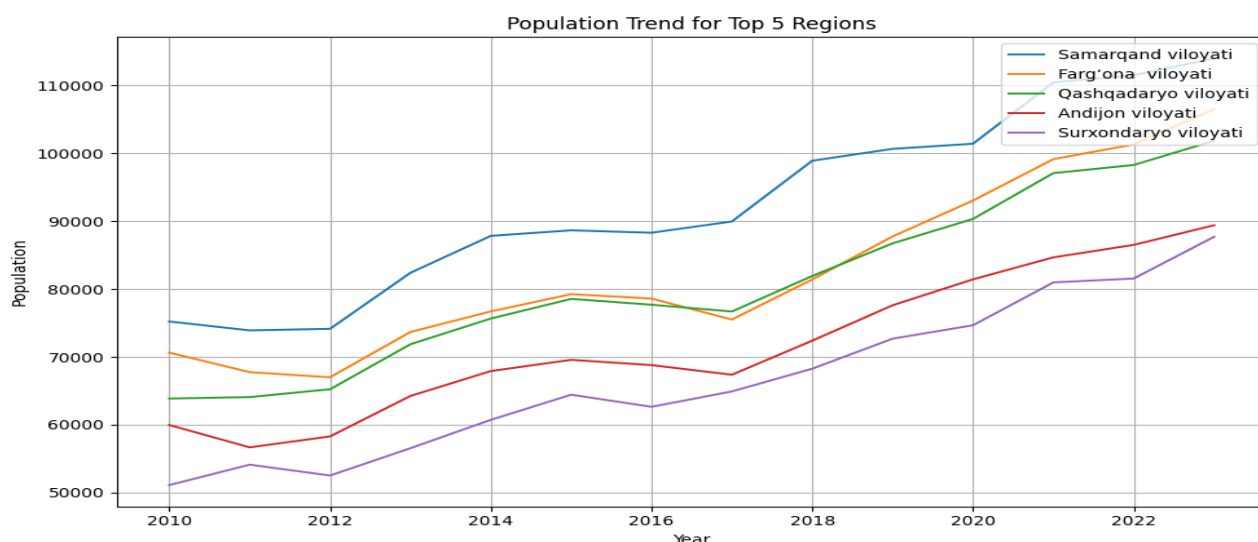


Figure 3. Line plot of to 5 regions.

The population trends for Uzbekistan's top five regions from 2010 to 2023 are depicted in Figure 3. Samarqand's substantial lead that figure emphasizes the region's steady population expansion, and the causes of this, such as migration, economic issues, or birth rates, might be investigated further. Regions such as "Farg'ona" and "Qashqadaryo" exhibit consistent growth, which is probably due to regional

development, legislation, or demographic changes. Some areas show minor variations, especially between 2020 and 2022. These might be brought on by outside variables like migration or shifts in the birth and death rates, which may have been impacted by the COVID-19 pandemic.

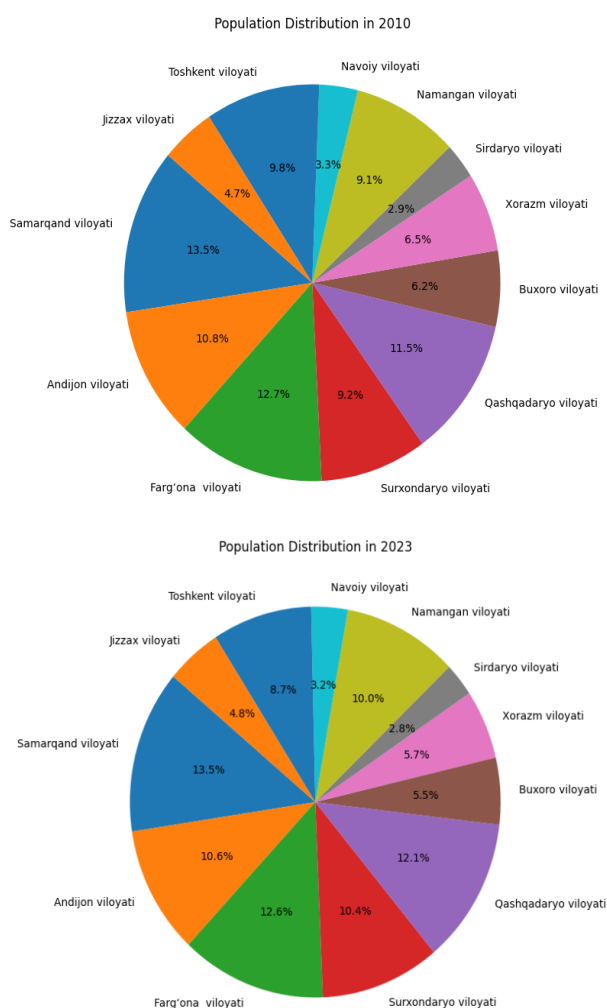


Figure 4. Population distribution across regions.

The population distribution in Uzbekistan in 2010 is depicted in Figure 4 as a pie chart, which shows how the population was dispersed over different areas (viloyats) at that time. A glimpse of Uzbekistan's demographic situation at the time may be found in the 2010 population distribution, which emphasizes the importance of key areas like "Samarqand" and "Farg'ona." This distribution is consistent with the historical, economic, and geographic elements affecting the nation's population patterns. These patterns have not changed much over time, although migration or development has caused demographic changes in some areas.

More people live in areas like "Samarqand" and "Farg'ona," most likely as a result of a mix of reasons like improved infrastructure, more chances for employment, and faster rates of urbanization. The higher percentages in areas like "Samarqand," "Andijon," and "Farg'ona" may be related to their significance as both cultural and industrial centers. Smaller population shares in areas like "Navoi" and "Xorazm" might be a sign of less urbanization, less economic activity, or more difficult geographic circumstances, all of which could be causing slower population growth or lower population density.

The population distribution of Uzbekistan by region in 2023 is depicted in this figure. While smaller locations may still be more rural or less developed in comparison, the concentration of people in some localities, such as "Samarqand" and "Farg'ona," suggests that these areas are important centers of activity in terms of commerce, infrastructure, and development. More people live in areas like Samarqand and "Farg'ona," most likely as a result of a mix of reasons like improved infrastructure, more chances for employment, and faster rates of urbanization.

The higher percentages in areas like "Samarqand," "Andijon," and "Farg'ona" may be related to their significance as both cultural and industrial centers. Furthermore, areas like "Navoi" and "Xorazm" with lesser population shares might be less urbanized, have less economic activity, or have more difficult geographic circumstances, all of which could be factors in slower population growth or lower population density.

Conclusion

Key insights into the regional distribution, population trends, and growth patterns are provided by the demographic analysis of Uzbekistan's population, which was carried out using data from 2010 to 2023. Some areas, including "Samarqand," "Farg'ona," and "Qashqadaryo," have continued to have a dominant population share over time, which is indicative of their historical, cultural, and economic importance. These areas have shown steady population

expansion, most likely due to a mix of industrialization, urbanization, and increased agricultural output.

Smaller areas such as "Navoi," "Xorazm," and "Sirdaryo," on the other hand, have retained lower population shares, which may be a reflection of their more rural character and possibly less developed infrastructure. Although regional differences in growth rates and population densities are still noticeable, the country's consistent population growth over the past ten years demonstrates Uzbekistan's overall progress and development.

These demographic trends have been greatly influenced by important variables like migration, birth rates, and local economic

prospects. Although the majority of the population still congregates in the main regions, some smaller regions have gradually seen population growth, indicating possible shifts brought on by economic or policy changes.

Policymakers, urban planners, and researchers can use this data as a basis to better comprehend Uzbekistan's demographic environment and make well-informed decisions about future planning, infrastructure development, and resource allocation. In the upcoming years, keeping an eye on these demographic patterns will be essential to tackling issues with population density, regional inequities, and sustainable development.

Reference

1. The pandas development team. (2020, February). *pandas-dev/pandas: Pandas* (latest version) [Software]. Zenodo. <https://doi.org/10.5281/zenodo.3509134>
2. Максакова, Л. П. (2020). DEMOGRAPHIC SECURITY IN MODERN UZBEKISTAN. *TIME DESCRIPTION OF ECONOMIC REFORMS*, (2), 22-30.
3. Narasimhan, V. M., Patterson, N., Moorjani, P., Rohland, N., Bernardos, R., Mallick, S., ... & Reich, D. (2019). The formation of human populations in South and Central Asia. *Science*, 365(6457), eaat7487.
4. Khasanova Shoirakhon Abdujabborovna. (2023). Demographic Indicators in The Republic of Uzbekistan and Their Importance in Medicine. *Texas Journal of Medical Science*, 27, 117–119. <https://doi.org/10.62480/tjms.2023.vol27.pp117-119>
5. Statistical data. Online available: <https://stat.uz/en/>
6. World bank: <https://www.worldbank.org/en/home>
7. United Nations Development Programme (UNDP): <https://www.undp.org/>
8. Atatistics Agency under the President of the Republic of Uzbekistan: <https://stat.uz/uz/>
9. Meliboev, A. (2024). IOT NETWORK INTRUSION DETECTION SYSTEM USING MACHINE LEARNING TECHNIQUES. *QO'QON UNIVERSITETI XABARNOMASI*, 11(11), 112–115. <https://doi.org/10.54613/ku.v11i11.972>
10. Meliboev, A. (2024). MACHINE LEARNING METHODS FOR BREAST CANCER CLASSIFICATION BY USING DATA SCIENCE TECHNIQUES. *QO'QON UNIVERSITETI XABARNOMASI*, 11(11), 101–104. <https://doi.org/10.54613/ku.v11i11.969>