Brandenburg University of Technology Cottbus - Senftenberg

Data Exploration and System Management Using Artificial Intelligence / Machine Learning

Project Report

Name: Mirza Adnan Baig
Reg # 5006936
Email: baigmir1@b-tu.de

Table of Contents

Analyzing Demographic Trends and Scenario Simulations in Pakistan's Population

Mirza Adnan Baig, B.Sc. Computer Science, M.Sc. Artificial Intelligence Student

Abstract

This research investigates the demographic trends in Pakistan using data visualization and scenario simulations. The study analyzes datasets to reveal key insights about population distribution, growth trends, and demographic dynamics across various dimensions, such as religion, age, and geography. Advanced simulation techniques explore the impact of hypothetical scenarios, including GDP-driven growth and uniform growth models. The findings highlight critical insights into the evolving demographic landscape, offering valuable implications for policymakers and stakeholders.

Keywords: Gross domestic product.

Introduction

Understanding demographic dynamics is crucial for planning and resource allocation in any country. This study focuses on Pakistan, a nation with a diverse and rapidly growing population. Using Python-based analysis, we examine historical trends, current distributions, and potential future scenarios. By integrating data visualization and scenario modelling, this research provides a comprehensive understanding of Pakistan's demographic challenges and opportunities. The goal is to identify key trends and their implications for policy and development planning.

Pakistan faces unique demographic challenges, including a high population growth rate and significant regional disparities. This study seeks to provide actionable insights by leveraging scenario-based simulations and data-driven visualizations. The findings are expected to assist policymakers in crafting informed strategies for sustainable development.

Methods

Data Collection

The analysis used datasets covering aspects like religion-based population distribution, yearly demographic statistics, and province-level trends.

Data sources include national surveys and public demographic databases.

Tools and Libraries

Python was the primary programming language used for this analysis.

Key libraries include:

- Pandas for data manipulation and cleaning.
- Matplotlib and Seaborn for creating detailed visualizations.
- NumPy for numerical computations.
- Scipy for advanced calculations and projections.

Techniques and Algorithms

Data Manipulation

Pandas was used to clean and pre-process the data, ensuring consistency in column formatting and handling missing values if any existed

Scenario Simulations

Growth Rate Simulations:

Applied static growth rates for each religion (e.g., Muslims: 5%, Hindus: -2%).

Used iterative calculations to adjust populations based on these rates over a specified time period.

GDP-Driven Growth:

Dynamically modified growth rates based on an external variable (GDP), implemented using conditional logic.

The algorithm leveraged Scipy for efficient optimization of growth rates over multiple iterations.

Uniform Growth:

Assumed a fixed 2% annual growth rate across all religions, maintaining their relative proportions over time.

Visualization Techniques

Bar charts and line plots were employed to display population trends by religion, age group, and geographic region.

Forecasting Models

Projections were generated using dynamic models to forecast population trends over a 10-year horizon.

The simulations incorporated economic factors to refine growth rates and align them with real-world possibilities.

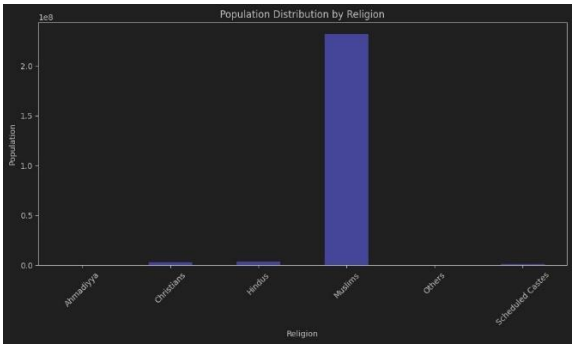
Results

Current Demographic Insights

The majority population is Muslim, followed by Hindus and Christians.

Visualizations revealed stark regional and age-based variations in population distribution.

Figure 1:



Scenario Analysis

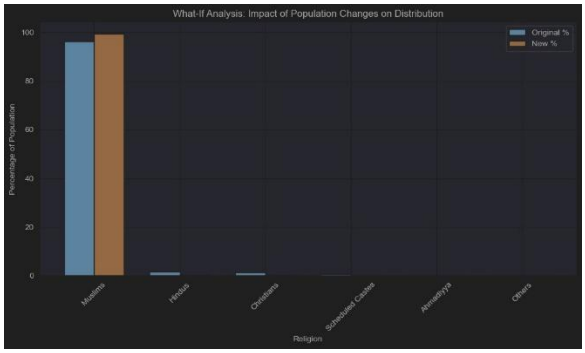
Growth Rate Variations

Applied religion-specific growth rates to simulate population changes.

Muslims experienced a significant increase, maintaining dominance across all scenarios.

Hindus and Scheduled Castes showed a decline due to negative growth rates.

Figure 2

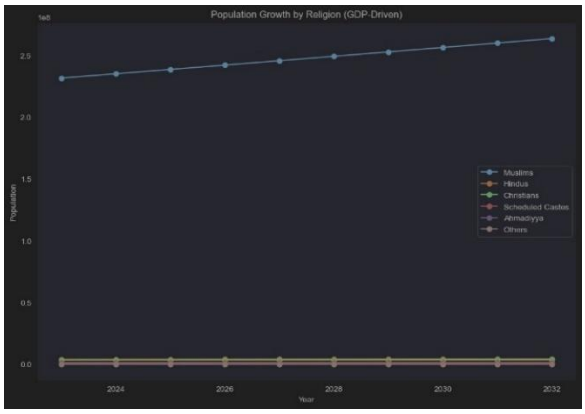


GDP-Driven Growth

Population growth rates were reduced as GDP increased, reflecting improved living standards.

The dominance of Muslims persisted, with minimal changes to relative proportions across other groups.

Figure 3

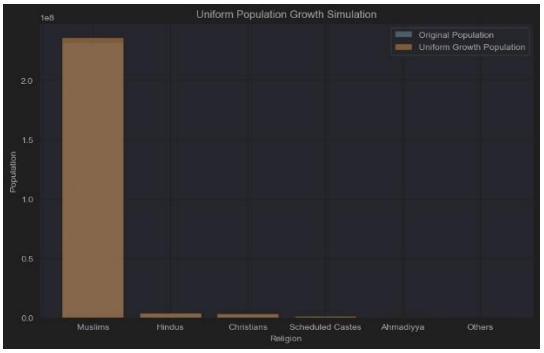


Uniform Growth

A consistent 2% growth rate was applied to all religions, resulting in proportional population increases.

This scenario demonstrated the effect of equal growth, maintaining existing population distributions.

Figure 4



Discussion

Implications of Findings

Religion-specific growth rates highlight disparities that could influence social dynamics.

GDP-driven simulations suggest that economic improvements can lead to population stabilization, emphasizing the importance of sustainable development policies.

Uniform growth scenarios provide a baseline for understanding how proportional distributions can remain stable under equal growth conditions.

Methodological Strengths

The use of Python libraries allowed for efficient data processing and high-quality visualizations.

Scenario simulations provided a flexible framework for exploring multiple hypotheses about demographic trends.

Limitations

The analysis assumes static growth rates for simplicity, which may not fully capture real-world complexities.

Datasets primarily focus on high-level trends; deeper granularity (e.g., district-level data) could refine insights.

Recommendations

Policymakers should consider economic and educational investments to moderate population growth sustainably.

Further studies should incorporate additional variables, such as migration and fertility rates.

Conclusions

This study offers a comprehensive analysis of Pakistan's demographic trends and the potential outcomes of various growth scenarios. Visualizations and scenario simulations provide actionable insights, emphasizing the interplay between population dynamics and socioeconomic factors. The integration of Python-based analysis techniques highlights the value of computational tools in demographic research. Future research should aim to expand the scope and granularity of the analysis to support more targeted policymaking.

About Dataset:

This dataset contains demographic and population data for Pakistan that spans multiple years. It includes important statistics such as:

Population Age Distribution: Insights into the age groups within the population, highlighting the proportion of individuals aged 0 to 14, 15 to 64, and 65 and older.

Gender Distribution: Data on the male and female populations, along with the overall sex ratio, providing a clear picture of gender balance in the country.

City and Province Trends: Population growth rates and changes in various cities and provinces from 1998 to 2023, offering a view of urbanization and development over time.

District Fertility Rates: Information on total fertility rates across different districts, along with the margin of error and survey year, to understand family size trends in Pakistan.

Annual Demographic Statistics: Key indicators like annual live births, deaths, natural increase, crude birth and death rates, infant mortality rates, total fertility rates, and life expectancy to analyze the overall health and growth of the population.

Religious Population Distribution: Data reflecting the population divided by religion, giving insights into Pakistan's diverse cultural and religious landscape.

This dataset is useful for researchers, policymakers, and anyone interested in understanding the demographic dynamics of Pakistan. It can help make informed decisions and develop effective strategies for the country's future.

References

- Python Libraries: Pandas, Matplotlib, Seaborn, NumPy, Scipy. URL: Popular Python Libraries - NumPy, Pandas, Seaborn, Sklearn
- Data Sources: National demographic surveys, public databases. Data Sources for Demographic Research | Pew Research Center
- Kaggle URL: https://www.kaggle.com/datasets/kirahhayatdata/pakistan-demographic-and-population-data
- Demographics of Pakistan. URL: https://en.wikipedia.org/wiki/Demographics_of_Pakistan
- Worldometer URL: https://www.worldometers.info/demographics/pakistan-demographics/
- World Bank Group. URL: https://data.worldbank.org/country/pakistan
- Department of Economic and Social Affairs Population Division. URL: https://population.un.org/wpp/
- Trading Economics. URL: https://tradingeconomics.com/pakistan/indicators