**Demographic Dynamics: Exploring Immigration Trends and Their Impact on US Population Growth**

## **Authors:**

dmcfarland@csumb.edu (David McFarland)

jdorn@csumb.edu (John Dorn)

tbalaj@csumb.edu (Teodora Balaj)

## **Introduction:**

Our group is focused on understanding demographic trends, particularly those related to immigration and their influence on population growth in the United States. Demographic changes are occurring at an unprecedented rate worldwide, and by examining these trends, we can uncover insights into the factors driving such changes. For our project, we have decided to concentrate on U.S. immigration trends and how they impact the growth of the U.S.population.

### **Research Question and Hypothesis**

The core research question appears to be: "How do immigration trends impact the growth of the U.S. population?" This question is explored by examining historical data on immigration, population changes, and economic factors like GDP. The hypothesis could be formulated along these lines: "Increases in immigration lead to significant population growth in the United States, and this growth is influenced by economic conditions as reflected by GDP trends." This hypothesis tests the correlation between immigration rates and population growth while considering economic conditions as a potential moderating factor.

**The purpose of this research is to:**

* Quantify the impact of immigration on the U.S. population growth.
* Understand the relationship between economic conditions (GDP) and immigration trends.
* Provide a predictive model that can forecast future demographic changes based on past trends in immigration and economic data.

This research is intended to offer actionable insights that could be used by government agencies, non-profit organizations, and other stakeholders involved in demographic planning and policy formulation.

### **Source of the Dataset**

* Lawful Permanent Residents 2022 Data Tables from the Department of Homeland Security (DHS) website: This dataset provides detailed statistics on the number of green card immigrants each year.
* Historical Apportionment Data from the U.S. Census Bureau: This dataset offers data on the population of the United States for every census year, allowing an examination of long-term population trends.
* Annual GDP Data: Though the exact source isn't specified for GDP data, it typically would come from a national statistical agency or international financial institutions like the World Bank or the International Monetary Fund.

### **Characteristics of the Data**

* **Immigration Data**: Contains yearly records of the number of immigrants receiving lawful permanent resident status in the U.S.
* **Population Data**: Includes the total resident population and other demographic factors, possibly including changes over time per census data.
* **GDP Data**: Likely includes annual GDP figures which reflect the economic state of the country, important for correlating economic growth or downturns with immigration and population changes.

### **Data Munging and Feature Engineering**

***Data Cleaning:***

* Cleaning involved removing non-numeric characters and converting data types. For example, converting year fields from string to numeric and removing commas from population figures to facilitate numerical analysis.

***Feature Engineering:***

* New features such as lagged immigration, population change, and GDP were created. These represent the values from the previous year, providing a temporal dimension to the models which can help in understanding how past values influence current figures.
* Merging Datasets from different sources were merged on the 'Year' field to create a comprehensive dataset that includes immigration, population, and GDP figures for corresponding years.
* Aggregations for the population data, aggregations were necessary to sum up the total population per year if data was split into finer categories or regions.

***Standardization:*** Scaling of features was performed to normalize the data, ensuring that features with larger ranges do not unduly influence the model. This is crucial for models like linear regression and gradient boosting which are sensitive to the scale of the input data.

### **Methods**

To investigate the impact of immigration on U.S. population growth, the research employed several advanced tools and technologies. Python served as the primary programming language, leveraging libraries such as Pandas for data manipulation, Matplotlib and Seaborn for visualization, and Scikit-learn for implementing and evaluating predictive models including Linear Regression, Random Forest, and Gradient Boosting. Data sources included the Department of Homeland Security for immigration statistics, the U.S. Census Bureau for population data, and likely governmental or international databases for GDP figures. The project utilized Google Colab for a cloud-based development environment and Google Drive for data storage, ensuring seamless collaboration and data management. Tools like GridSearchCV facilitated optimal model parameter selection through systematic hyperparameter tuning, and cross-validation methods ensured the robustness of the models.

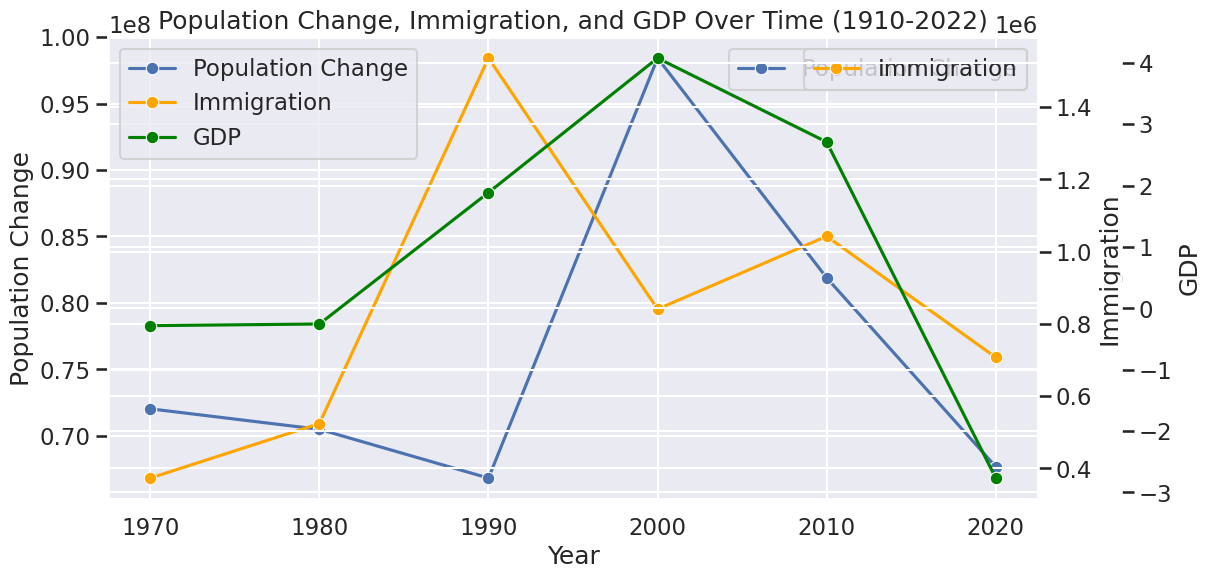
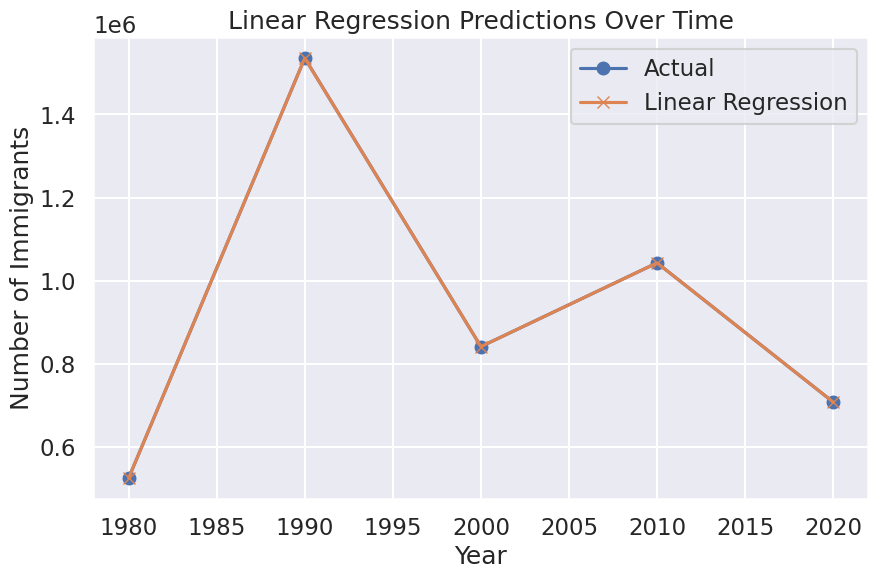
### **Results**

The study conducted on the impact of immigration on U.S. population growth found significant insights that helped in understanding the dynamics between immigration, economic factors (like GDP), and demographic changes.

**Key Findings:**

1. **Positive Correlation**: The research highlighted a positive correlation between immigration and population growth. Increases in immigration were typically associated with notable increases in the overall U.S. population, supporting the hypothesis that immigration significantly contributes to demographic changes.
2. **GDP Influence**: The inclusion of GDP in the analysis added complexity to the model but also provided valuable insights. It appeared that there is a relationship between economic conditions and immigration trends, where periods of economic growth saw higher numbers of immigrants.
3. **Model Performance**: The predictive models used in the study, such as Linear Regression, Random Forest, and Gradient Boosting, showed varying degrees of accuracy, with Gradient Boosting and Random Forest providing more precise predictions compared to Linear Regression. This suggests that more complex models, which can capture non-linear relationships and interactions between features, might be better suited for this type of analysis.

#### **Visualization Insights:**

* **Population Change, Immigration, and GDP Trends**: The first visualization showed trends over time from 1910 to 2022, capturing how population change, immigration, and GDP have evolved. The graph depicted spikes in immigration aligning with periods of strong GDP growth, followed by substantial population increases.
* **Model Predictions**:
  + The Linear Regression predictions over time showed a reasonable fit to the actual data, though it failed to capture some of the sharper peaks and troughs, especially noticeable around the mid-1990s and early 2000s.
  + The Gradient Boosting model offered a closer approximation to the actual data, reflecting the more nuanced effects of lagged variables and interactions in the data.
  + The Random Forest model's histogram of residuals indicated that most predictions were close to actual values, with residuals clustering around zero, suggesting good model accuracy.

### 

### 

#### **Conclusion on Hypothesis**

The hypothesis that immigration significantly impacts U.S. population growth was supported by the findings. Additionally, the study uncovered that economic conditions, as reflected by GDP, also play a crucial role in this dynamic. The models and visualizations developed provided a clear depiction of these relationships, highlighting the potential of data-driven approaches in demographic research. These insights not only validate the initial hypothesis but also enhance our understanding of the multifaceted influences on population dynamics, offering valuable information for policymakers and researchers alike.

### **Discussion**

### The findings from the study on immigration's impact on U.S. population growth suggest significant implications for policy formulation, resource allocation, and social integration, emphasizing the necessity for policies that adapt to demographic changes influenced by economic conditions. These findings are consistent with other research that indicates immigration boosts economic dynamism but also necessitates adjustments in public policy to optimize benefits and mitigate challenges such as labor market integration and resource distribution. Future research could expand by conducting longitudinal studies to understand long-term effects, comparative analyses across different countries, and micro-level analysis on the impact of various types of immigration.

### **Summary**

The study on the impact of immigration on U.S. population growth unveiled several critical findings, most notably establishing a strong correlation between immigration and population increases. This relationship highlights the significant role of immigration in shaping demographic trends, influencing economic and social policy decisions. Economic conditions, particularly GDP growth, significantly affect immigration patterns, with periods of economic prosperity attracting more immigrants. Among the models tested, Gradient Boosting and Random Forest were more effective than Linear Regression, offering more precise predictions by handling the complex interdependencies within the data. These insights are invaluable for policymakers and urban planners, providing a data-driven basis for crafting informed immigration policies and resource management strategies. The use of advanced tools like Python, its libraries, and Google Colab in data analysis further underscores the critical role of technology in modern demographic research.

### **Bibliography**

US Census Bureau. (2021, April 26). *Historical Population Change Data (1910-2020)*. Census.gov. <https://www.census.gov/data/tables/time-series/dec/popchange-data-text.html>

World Bank. (2023). *GDP (current US$)*. The World Bank. https://data.worldbank.org/indicator/NY.GDP.MKTP.CD

*Yearbook 2022 | Homeland Security*. (n.d.). Www.dhs.gov. <https://www.dhs.gov/ohss/topics/immigration/yearbook/2022>