The Chinese University of Hong Kong, Shenzhen
Final Project for MAT3300

November 15, 2024

Final Project 1

Weitang LI & Paolo PIERSANTI

Instructions

1. You should choose only one between Project 1 and Project 2 and work on one of the projects. If solutions to both projects are submitted, the Instructors will decide which one to grade.

2. Submit your work in a form of a `.zip` file onto Black Board. The `.zip` file you submit must contain a short report in `.pdf` format, the source code (either MATLAB or Python files) for the scripts you coded to solve the proposed problems, as well as the files for the figures you are required to generate.

3. The report you are required to prepare should thoroughly address the questions that are raised in this Project. The report can contain figures and tables.

4. The report cannot exceed **15 pages**. Please use LATEX or the provided `.docx` template.

5. This is an individual project. You can collaborate with your classmates, but the work you submit must be yours and must be phrased in your own words.

6. Comment the code in the proper manner to clearly illustrate your solution strategy.

7. The submission deadline for the Final Project is **December 15, 2024**, at **11:59PM**.

8. Late submissions are **not** accepted. If you fail to meet the deadline, your score for the Final Project will be set to zero.

9. If you need clarification on this Final Project 1 statement, please email Prof. Weitang Li (`liwt31@gmail.com`). Please note that the Course Instructors are not supposed to provide assistance in bringing the objectives laid out in the Final Projects to fruition.

Global Populations 2050: Trends and Insights

Background

Global population growth and density are key factors shaping economic development, environmental sustainability, and social stability. Understanding and predicting population growth trends is critical for planning in areas such as infrastructure, resource allocation, healthcare, and policy-making. However, population growth is highly variable across countries, influenced by factors like land area, density, and economic development. Accurate forecasting is essential for both rapidly growing countries and those facing population stagnation or decline.

Task

Using a dataset of over 200 countries, which includes the historical population from 1970 to 2022, population density, growth rates, and land area, you are asked to develop a model to project each country's population up to the year 2050. The model should account for different growth patterns, geographic limitations, and any significant trends observed in the data.

Dataset Description

The dataset has 234 rows and each row represents a country/territory. The columns in the dataset are

- Rank: Rank by Population.
- CCA3: 3 Digit Country/Territories Code.
- Country/Territories: Name of the Country/Territories.
- Capital: Name of the Capital.
- Continent: Name of the Continent.
- 2022 Population: Population of the Country/Territories in the year 2022.
- 2020 Population: Population of the Country/Territories in the year 2020.
- 2015 Population: Population of the Country/Territories in the year 2015.
- 2010 Population: Population of the Country/Territories in the year 2010.
- 2000 Population: Population of the Country/Territories in the year 2000.
- 1990 Population: Population of the Country/Territories in the year 1990.
- 1980 Population: Population of the Country/Territories in the year 1980.
- 1970 Population: Population of the Country/Territories in the year 1970.
- Area (km²): Area size of the Country/Territories in square kilometer.
- Density (per km²): Population Density per square kilometer (as of the year 2022).
- Growth Rate: Annual Population Growth Rate by Country/Territories of the year 2022.
- World Population Percentage: The population percentage by each Country/Territories.

The dataset is downloaded from Kaggle and more information on the dataset can be found following the link.

Requirements

1. Data Analysis and Trend Identification (20 points)

Analyze the historical data to identify growth patterns and classify countries into categories (e.g., rapid growth, stable, declining) based on trends. Determine which factors (population density, land area, continent, etc.) most significantly influence growth trends.

2. Model Development (30 points)

Develop a mathematical model or set of models to project population changes from 2022 to 2050 for each country. Consider creating different models for different categories of countries, if needed. If possible, incorporate uncertainty factors such as migration, economic changes, pandemic outbreaks, and potential impacts of global challenges (e.g., climate change).

3. Regional Comparisons (20 points)

Aggregate, analyze, and visualize population projections at a regional or continental level. Predict the total world population from 2022 to 2050, and analyze the implications of regional differences on global trends.

4. Visualization (30 points)

Create clear visualizations (e.g., graphs, maps) to illustrate key findings, such as projected population changes and regional growth comparison.

Deliverables

You should submit a zip file containing the following components:

- Project Report. The report should include a detailed explanation of your model(s) and assumptions, the method you used to develop your model, and visualizations that clearly convey your results and conclusions. The report should begin with an executive summary with no more than 1 page and the total length should not exceed 15 pages.
- Code. Include all code and necessary environment setup to allow for reproducibility of your findings. Ensure that the code is well-documented for clarity.
- Data. If you have collected any self-collected data, include this in your submission. Document your data collection process, methodology, and any preprocessing steps taken.

Evaluation Criteria

Your submission will be evaluated by the following criteria:

- Accuracy. How well does your model capture historical trends, and how effectively does it project future growth?
- Insight. Does your analysis provide insights into global population trends of the 21st century?
- Presentation. Regardless of the complexity of the model, can a reader easily understand your reasoning and conclusion?