

World Population: Data analysis and model for prediction.

This is a data science project done as a team in 2021 by the following members: Yusuf Tijani, Medina Lebari, Adetutu Isinkaiye, Aisha Mohammed.

Tools used: Python, SQL

Introduction/Problem Statement

There are about 7 billion people in the world currently, and this number is likely to keep rising as we add 227,000 people to the planet every day¹. As world population increases, the pressure on life sustaining resources (food, water, shelter, etc) also increases. There is no gainsaying that if not properly managed, the human population could outweigh resources available to sustain it. This is why it is important to study population data and devise a means to predict future population. Population projection can help governments to make decisions and plans about the future.

Thought process

We worked with the world population data set found on [Kaggle](#) for our project. The dataset contained the population of the world per year from 1951 to 2020. It also had the net change in population per year, density, and urban population from which we deduced the rural population. We explored the data of 70 rows and 7 columns, to see what we could learn about changes in population over the years, and also how we could use the data to predict the human population in coming years.

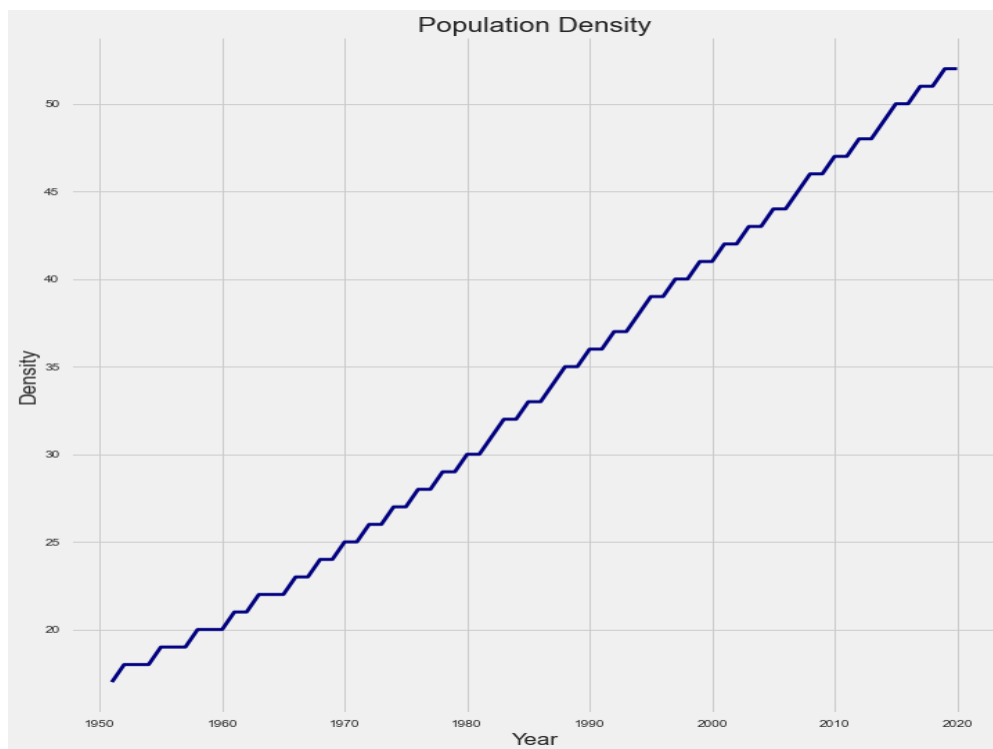
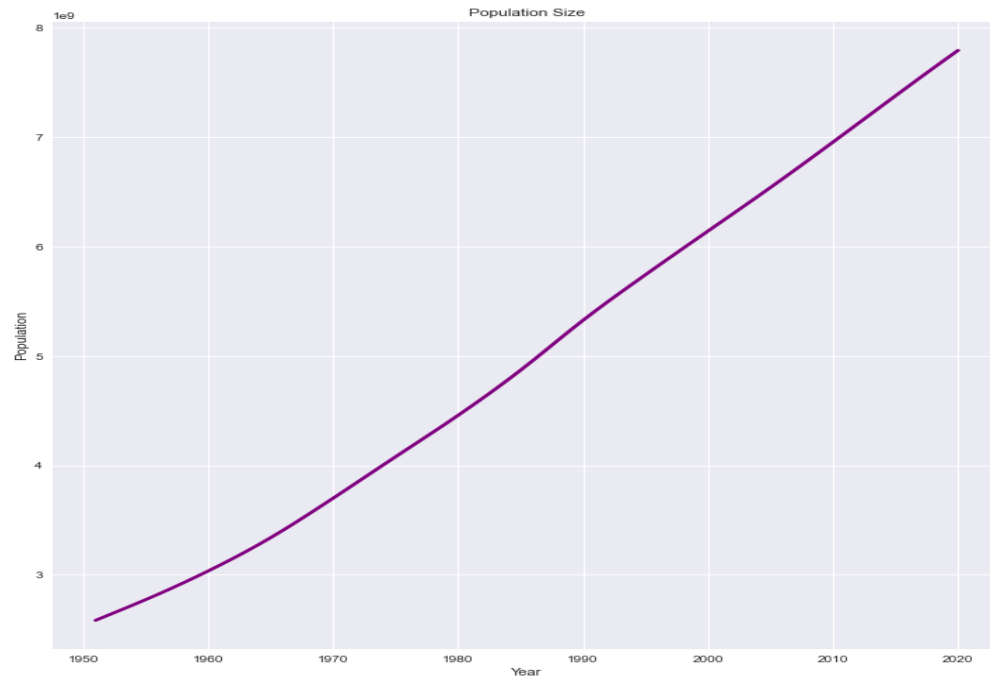
Model and techniques

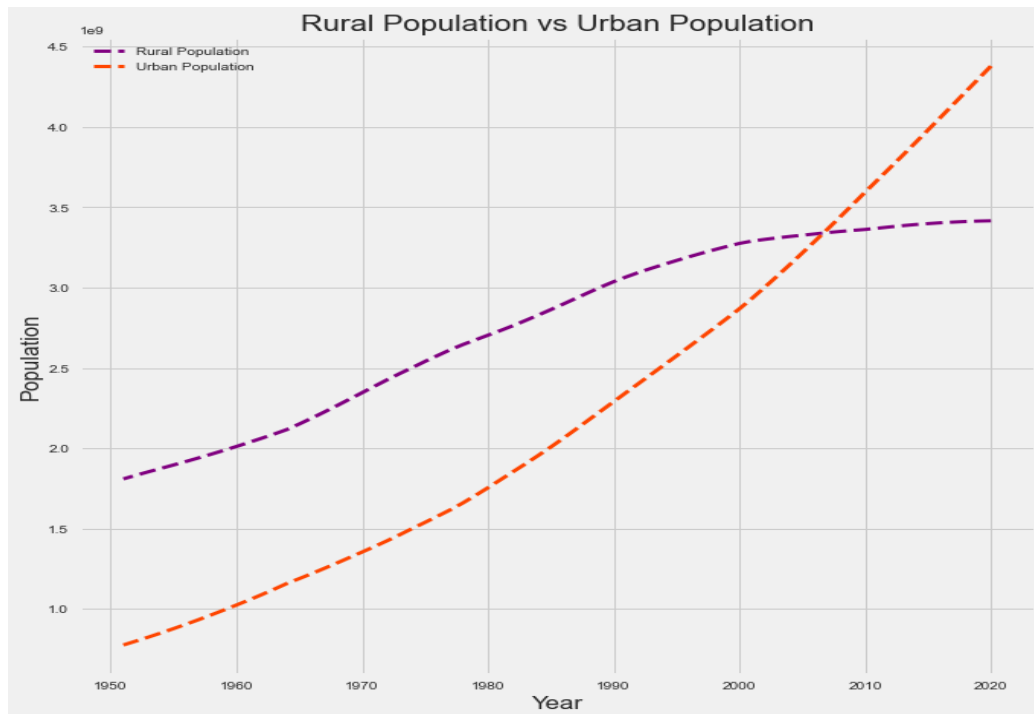
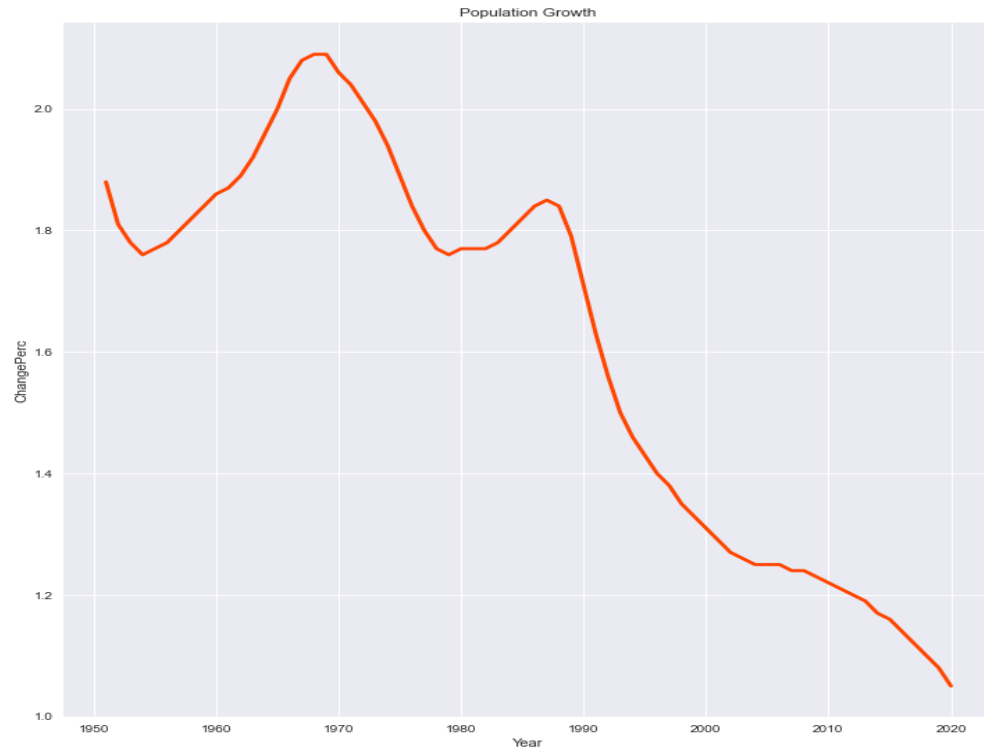
We analysed the data column-wise and created charts to observe the changes in population size, growth rate, density, and changes in urban versus rural population (charts below).

We also built a simple linear regression model to predict the world population for any year that was entered into the algorithm. We used 70% of the data as our training dataset and 30% was used for testing the model. Our model has a score of 99.68% and an R2 score of 0.9965. Our R2 score being close to 1 indicates a good fit, and tells how reliable the predictions are. Here is the link to our complete project on [Github](#).

Results and Discussion

The data showed that the world population has grown rapidly over the last seventy years. We went from having about 2.6 billion people in the world in 1951 to about 7.8 billion in 2020, a threefold increase. This increase in size could be explained by medical advancements in recent centuries and associated decline in death rate. Using our model, we predicted that the population in 10 years would be 8.5 billion.





As world population numbers have continued to go up, while the size of our planet remains the same, it is not surprising that we found the population density to be increasing also. In 1951, density was 17 people per square kilometer, while in 2020, it had gone up to 52 people per square kilometer. Increasing density can pose a challenge in managing outbreaks of communicable diseases and pandemics like COVID-19 where distancing is a protective measure.

While the world population has tripled, it is noteworthy that the rate of this growth is actually falling. We saw the highest growth in population between 1960 and 1970, and a steady decline from around 1990. This could be good because smaller populations have been associated with less crowding, maintaining green areas, and improved quality of life², while overpopulation is associated with increased stress on available resources, resulting in diminished quality of life³.

In terms of location, there were initially more people in rural areas in the early 50s, compared to urban areas. From the year 2000, the rural population seemed to plateau, while the urban population continued to increase thereby reversing the rural-urban numbers. This could be explained by possible migration of people to urban areas for economic reasons, industrialization, etc.

Limitations

We would have loved to have the population per country or continent in our dataset in order to compare population growth/changes across countries and continents. Also, to build a better and more effective model, we would have liked to have a bigger data set.

Conclusion

We saw how much the population of the world has increased and predictions show that future increase is to be expected, albeit at a slower rate than what was observed 50 years ago. With such predictions, governments can plan ahead, e.g. in terms of estimating basic human needs like food, water, power, transport, housing, etc. Population planning methods, as is already in place in some of the most populated countries in the world, can be employed to control the growth rate of the population.

References

1. https://www.biologicaldiversity.org/programs/population_and_sustainability/population/
2. <https://www.eurekalert.org/news-releases/619604>
3. https://en.wikipedia.org/wiki/Human_overpopulation#Proposed_impacts