Populations Of Countries

Data Science Final Paper
Gianna Bagola

Introduction

Over the past few decades, the population of countries and Earth in general have grown largely and quickly. Just in 2022, the world population just broke 8 billion. In 1950, the world population was just a little over 2 billion! But something is happening where the populations are starting to slow down. While the population is still increasing, the growth rate is decreasing. What is causing this? Only time can really tell. People have predicted the future population growth but unexpecting things can happen. Population decrease and increase can both be good and bad things, but how do you know when it's going to be a certain case? There have been cases where the government gets involved in population numbers whether it's to increase them or decrease them because of the effects the population numbers have on the countries.

Problem Statement/Hypothesis

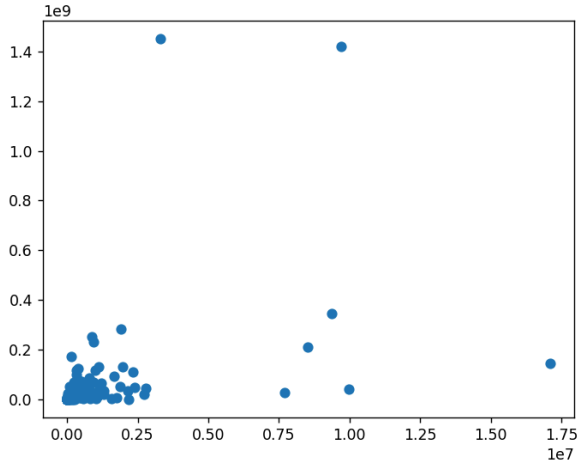
For the countries with large populations, do they have a larger area? For a large population, if there is not enough space for the people, will this affect their growth rate negatively? Is there a relation between population and area? And if so or not so, is there a relationship between population and growth rate? Also, would a country with a higher world percentage mean they have a higher net change?

It is being predicted that there will most likely not be a relationship between population and area because there are known countries with large land mass but low populations and vice versa. And for growth rate, there might be a similarity between each country and their growth rate since very similar things are happening in different countries around the world. World percentage and net change is tricky but it might be predicted that they are related in some way.

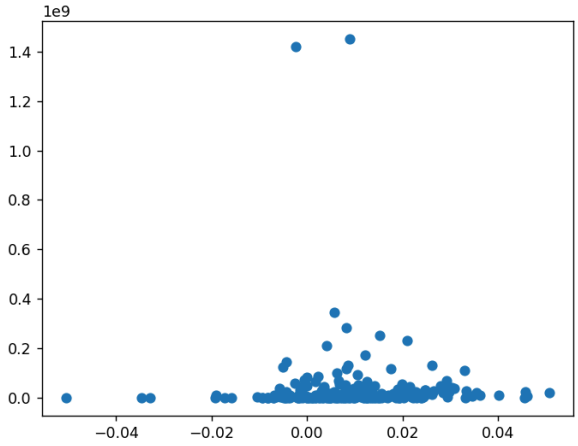
The Data

https://www.kaggle.com/datasets/arpitsinghaiml/world-population
The data includes the population of countries in 1980, 2000, 2010, 2023, and 2024. It also includes a predicted population for 2030 and 2050. It also includes the country of which it is populated, the area, net change between populations, growth rate, and world percentage which means the total world population that resides in the country.

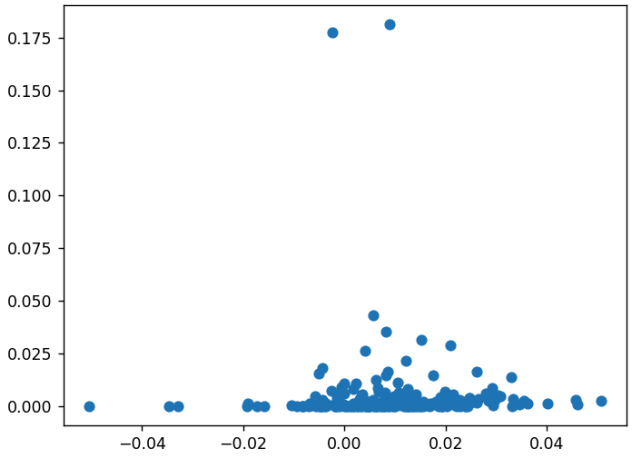
Testing



All Countries' Areas vs All Countries' Populations in 2024



All Countries' Growth Rates vs All Countries' Populations in 2024



All Countries' Growth Rates vs All Countries' World Percentages

```
ANOVA result between all populations in the different years: F = 0.6966314976087369, p = 0.5943233270390167
Fail to reject null hypothesis: all group means are the same

ANOVA result between area and population in 2024: F = 14.373868059827752, p = 0.0001695616325910518
Reject null hypothesis: at least one group mean is different

ANOVA result between growth rate and population in 2024: F = 14.867861835353017, p = 0.00013151466979476132
Reject null hypothesis: at least one group mean is different

ANOVA result between growth rate and world percentage: F = nan, p = nan
Fail to reject null hypothesis: all group means are the same
```

```
Spearmans correlation coefficient: 0.851
Samples are correlated (reject H0) p=0.000
```
-Area & Population

```
Spearmans correlation coefficient: 0.272
Samples are correlated (reject H0) p=0.000
```
-Growth Rate & Population

```
Spearmans correlation coefficient: nan
Samples are correlated (reject H0) p=nan
```
-Growth Rate & World Percentage

Results

To be honest, there must be something wrong with the code because the results do not seem likely. One result that I believe the most in is the ANOVA test done on all of the population years. The fact that the result failed to reject the null hypothesis shows that each country has a steady rate of growth and aren't changing differently than they have been. All the other results of the ANOVAs make a good amount of sense. For the correlation, each test rejects the null hypothesis. This means they were not correlated? Very interesting results, each of the data are all on their own and do not affect each other.

Conclusion

Because of the confusing results, it's difficult to conclude about the data. But the relationships between each of the data are all independent. The populations are still growing but not exactly at the most rapid rate. Perhaps if there are more years included in the data set, we can see another bigger picture and do more tests. But for now we know nothing is related.

Resources

Armstrong, M., & Richter, F. (2022, November 15). *Infographic: World population reaches 8 billion*. Statista Daily Data. https://www.statista.com/chart/28744/world-population-growth-timeline-and-forecast/