

Applied Data Science 1

Assignment#2: Statistics and Trends

Name: Umair

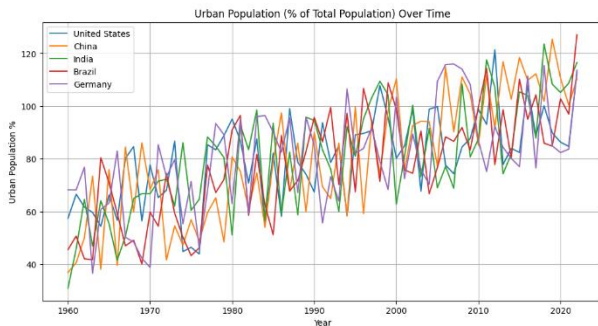
Roll# 22098012

Github: <https://github.com/umair556677/ADS1-assignment2>

Title: "Unlocking Global Insights: A World bank Data Exploration"

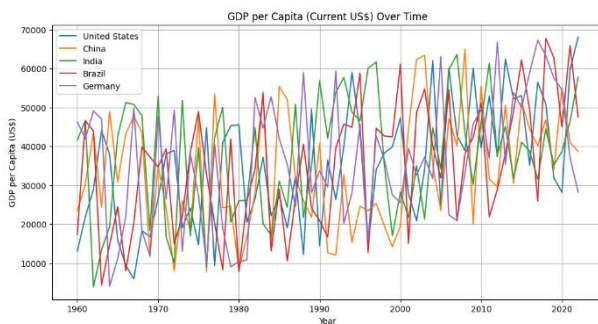
Abstract: This study shows a comprehensive exploration of Worldbank data, which aims the uncovering profound insights into global trends. First starting by developing a robust data manipulation function that changes the data into two structured dataframes. One with years as columns while the other with countries as columns—after data cleaning. Our analysis includes on select indicators, enabling comparisons between countries and aggregated regions. We have done statistical methods which includes the .describe() method, and also other tools to gain a highly deeper understanding of these indicators, the statistical properties, and their evolution of over time.

The analysis we will do of some urbanization trends overtime. The change in urban population of the 5 countries, US, China, India , Brazil and Germany.



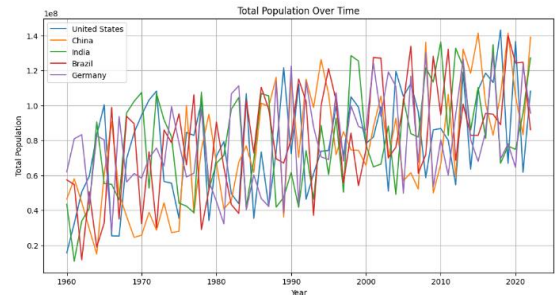
The above time series graph we can see the urban population over time in 5 countries. The overall trend we can see that the urban population increases in time.

Now we will analyze the GDP per capita over time of these 5 countries.



From the above time series chart we can see that the x axis is upto 70000 which is in US dollar and in time domain it is upto 2020. So from the above trend we can analyze that how the GDP of the countries are going.

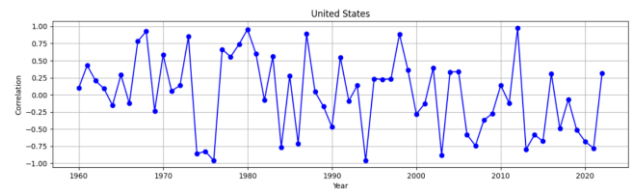
In this we will see the overall population of time.



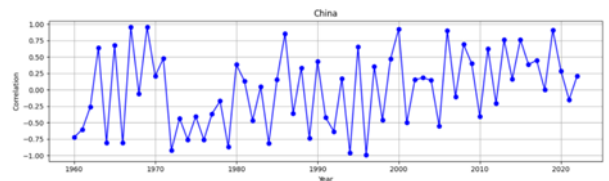
The overall population increases with increase in time of the countries. From 1960 to 2020 the trend goes increasing.

Now we will see the yearly correlation analysis of each country.

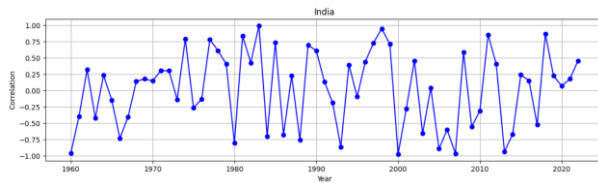
United states:



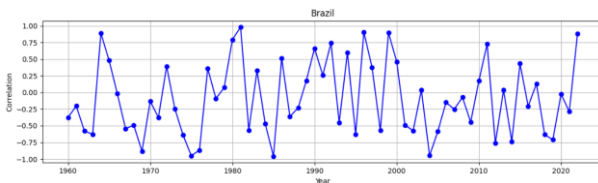
China:



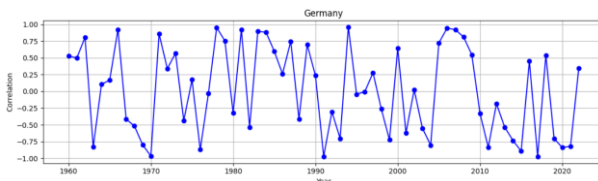
India:



Brazil:

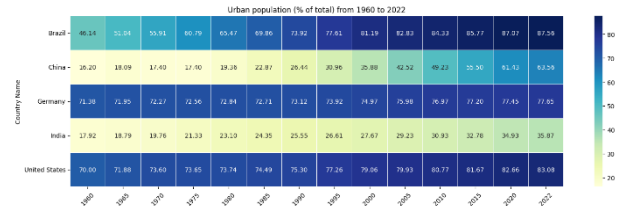


Germany:



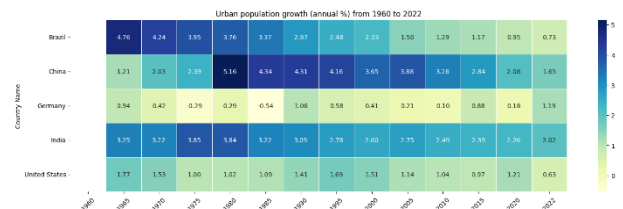
These above correlation graphs shows the annual relationships between population growth and energy that is used in various countries. Each graph represents a specific nation's correlation trend over time. Some countries have a stable positive correlation, showing that as their populations grow, energy consumption increases. Also, others display a consistent negative correlation, showing that as population size expands, energy use tends to decrease. Some countries, show erratic and unpredictable correlations, showing that factors beyond population growth significantly influence their energy consumption patterns. The visualizations shows the diverse and complex nature of the interactions between the population dynamics and energy usage across different nations.

Heatmap:



This is the heatmap for the indicator 'SP.URB.TOTL.IN.ZS', which shows the urban population as a percentage of the total population, from 1960 to 2022 for Brazil, China, Germany, India, and the United States.

The color intensity represents the normalized value of the indicator for every country and year, with darker shades showing higher percentages of urban population.



This heatmap shows the 'SP.URB.GROW' indicator, which displays the urban population growth (annual %) from 1960 to 2022 for the selected countries.

The color intensity shows the normalized value of the annual growth rate, with darker shades indicating higher growth rates.

The heatmap's pattern shows the following insights:

- 1. Temporal Trends:** Over the years, we can see the variations in data values for each country. In some years exhibit larger values (warmer colors), while others show lower values (cooler colors). These fluctuation present changes in the indicator that is being analyzed, such as economic growth, environmental factors, or any other relevant metric.
- 2. Between-Country Comparisons:** Comparing the columns (countries) in the heatmap shows for an assessment of how different countries' indicator values evolve over time.