## Module Name: Applied Data Science 1

**Assignment 2:** Statistic and trends

**Student Name:** Ali Hamza Yasin

**Student ID:** 21090741

**GitHub Repository link:** <https://github.com/AliHamzaYasin/Applied_Data_Science_Assignment2.git>

**Dataset Link: https://data.worldbank.org/topic/climate-change**

Shape

Description automatically generated with medium confidence

**Abstract**In this project, we have presented the statistical analysis and trend observation. As part of this assignment, we have looked at climate change-related indicators across nations using World Bank data (available at https://data.worldbank.org/topic/climate-change). Access to energy, agricultural activity, urban population, etc. is only a few of the many variables that are significant to climate change.

Chart, line chart

Description automatically generatedThe above line graph showing the greenhouse gas emissions (in kt of CO2 equivalent) for six countries (United States, China, India, Japan, South Korea, Indonesia) from 1990 to 2020 in five-year increments. The graph shows that the United States has consistently been the largest greenhouse gas emitter among these six countries since 1990, with emissions peaking in the mid-2000s and declining slightly thereafter. China's greenhouse gas emissions have steadily increased over the years, and in the early 2000s surpassed those of the United States, making it the largest emitter globally. India's greenhouse gas emissions have also been steadily increasing, albeit at a lower rate than China's. The other three countries (Japan, South Korea,

and Indonesia) have lower greenhouse gas emissions compared to the top three emitters but show increasing trends over time.

Moreover, the graph illustrates the differences in greenhouse gas emissions trends and magnitudes among the six countries, highlighting the need for continued efforts to reduce emissions globally.

Chart, line chart

Description automatically generated

In the above graph the x-axis shows the year, the y-axis shows the energy consumption per capita, and the hue (color) of the line represents each country. The graph shows how energy consumption per capita has changed over time for each of the selected countries. Graph shows that the Qatar is high energy consumer per capital than other provided countries. The title of the graph is "Energy Consumption by Country (1990-2020)" and the x-axis is labelled "Year" while the y-axis is labelled "Kg of oil equivalent per capita".

Chart, line chart

Description automatically generated

The above line graph shows the population trends of five countries over a 40-year period (1990-2020). It contains the results of five countries: the United States, China, India, Japan, and Pakistan. The x-axis of the graph represents the years, and the y-axis represents the population. Each country's population is plotted as a line on the graph, with a different color and a label for each country. The graph shows the population trends of each country over time, allowing for comparisons and analysis of changes in population growth rates. Graph shows that the China has greater population growth rate than other countries from the provided list and Japan has least.

Chart, bar chart

Description automatically generated

This bar graph showing the comparison between urban population and total population for four countries: United States, China, Pakistan, and Russian Federation. The x-axis shows the years from 1960 to 2020 in 20-year increments, and the y-axis shows the population size in millions.

Moreover, we can see that the total population of China has been growing steadily since 1960, while the total population of the other countries has remained relatively stable. The urban population of China has grown rapidly since 1980, while the urban populations of the other countries have also grown, but at a slower pace. The United States has the largest urban population and the second-largest total population after China. Pakistan has the smallest urban and total population among the four countries. The Russian Federation has a relatively small urban population compared to the other countries, but its total population has remained stable.

Chart, treemap chart

Description automatically generated

The above shows three correlation heatmaps between forest area and another factor for four countries (United States, China, India, and Japan) over a range of years.

The first heatmap shows the correlation coefficients between forest area and another factor for each country over the entire range of years. The correlation coefficients range from -1 to 1, where -1 represents a perfect negative correlation, 0 represents no correlation, and 1 represents a perfect positive correlation. The heatmap uses a color scale to show the correlation coefficients, with cool colors (blues) representing negative correlations and warm colors (reds) representing positive correlations. The annotations on the heatmap show the actual correlation coefficients.

Chart, treemap chart

Description automatically generated

The second heatmap shows the correlation between forest area trends and the other factor for each country over the subset of years. The heatmap shows only the correlation coefficient for forest area, and the annotations show the actual correlation coefficients.

Chart, treemap chart

Description automatically generated

The third heatmap shows the correlation between the other factor trends and forest area for each country over the subset of years. The heatmap shows only the correlation coefficient for the other factor, and the annotations show the actual correlation coefficients. The heatmap has labels on the x and y axes indicating the variables being compared (forest area and total population). The heatmap uses the same color scale as the first heatmap.

Chart, waterfall chart

Description automatically generated

The above heatmap shows the seven indicators of interest related to urban and rural populations, land use, economic growth, greenhouse gas emissions, and energy use in Pakistan. The resulting heatmap shows the correlation between the selected indicators for Pakistan. The plot title indicates that the heatmap is for Pakistan specifically. The annotations in the cells display the correlation coefficients. The heatmap can be used to identify patterns and relationships between the indicators, which can provide insights into the country's social, economic, and environmental dynamics.