Module Code: 7PAM2000

Module Title: Applied Data Science 1

Assignment 2: *Statistics and Trends*

Link of GitHubRepository: https://github.com/Machindrareddy/ADS1-Assignment-2

Short Abstract

This analysis explores climate change indicators using World Bank data, focusing on Afghanistan and later on Nigeria, Ireland, and Aruba. The dataset covers 1960-2019, addressing health and socioeconomic factors, with identified missing values. Python, pandas, and seaborn are employed to filter and visualize data. Specific attention is given to 'CO2 emissions from gaseous fuel use (kt),' revealing emission differences among nations through a boxplot. Line plots depict trends in CO2 emissions and urban population percentages. A correlation heatmap suggests a significant. relationship between forest area and arable land. Time series analysis highlights environmental and demographic changes in Aruba, Nigeria, and Ireland. External World Bank data on climate change enriches the analysis.

STUDENT NAME: T MACHINDRA REDDY

STUDENT ID: 22076527

Climate Change Data Analysis Based on World Bank Data

Figure 1: Basic Descriptive

The given data frame seems to include data on several Afghanistan-related statistics arranged annually. The indicators address several issues, including the treatment of ARI for children under five, access to clean fuels, and power availability for both urban and rural communities. The dataset includes values for each indicator for the years 1960 through 2019. Nevertheless, several cells have missing values (NaN), which suggests that the data for certain indicators or years is not full. Over the given period, the dataset enables the investigation of trends and patterns in Afghanistan's socioeconomic and health-related
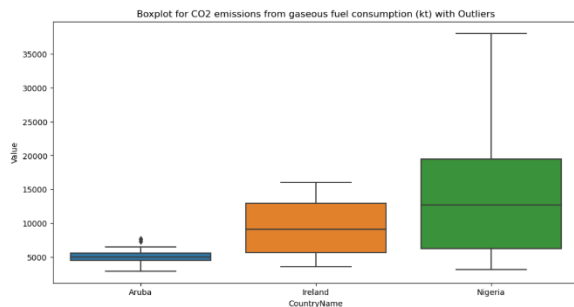
metrics.

Figure 2: Boxplot for CO2 emissions from gaseous fuel consumption (kt) with outliers

The World Development Indicators data is filtered and altered using pandas and seaborn in the accompanying Python code. It focuses on certain metrics for Nigeria, Ireland, and Aruba.

The 'CO2 emissions from gaseous fuel use (kt)' in each nation are shown in an outlier scenario whereby Nigeria has the highest emissions and Aruba has the lowest. The difference in emissions between the chosen nations is then shown by the code via the use of a boxplot to visualize these emissions.
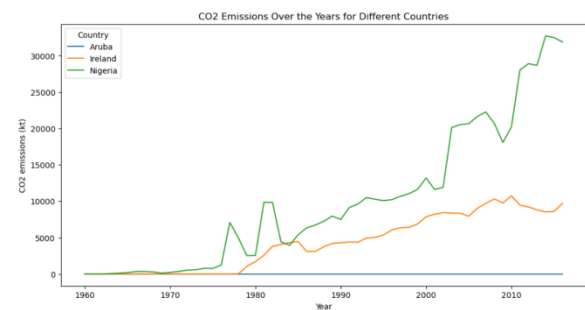

Figure 3: CO2 Emissions Over the Years for Different Countries

Focusing on the designated indicators and filtered data, the given Python code creates a line plot showing the trend of 'CO2 emissions from gaseous fuel use (kt)' throughout the years for various nations. The resultant map shows how CO2 emissions have changed over time, with different lines representing Nigeria, Ireland, and Aruba.
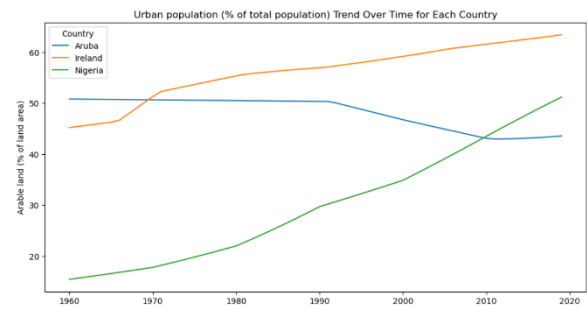

Figure 4: Urban population (% of the total population) Trend Over Time for Each Country

Figure 5: Correlation Heatmap for Countries Based on Indicators

The claim that "Forest area and arable land have the proper correlation relationship" implies that, in this particular situation, there is a significant and predictable association between the two. This might mean that, in line with ecological or land use trends, one of these variables tends to rise while the other tends to decline, or vice versa.
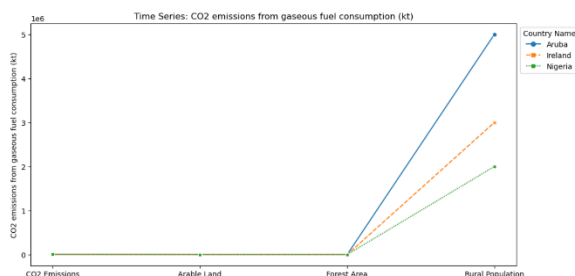

Figure 6: Time Series: CO2 emissions from gaseous fuel consumption (kt)

The code represents and visualizes time series data for CO2 emissions, arable land, forest area, and rural population in Aruba, Nigeria, and Ireland using a pandas data frame. A line plot showing the changes over time is made possible by the transposed data frame, giving a clear comparison of key environmental and demographic characteristics across the three nations.
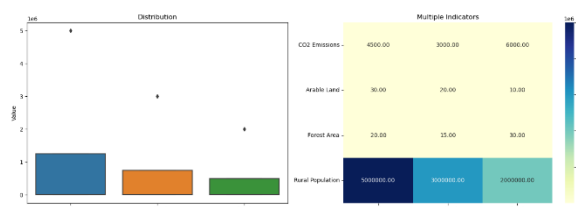

Figure 7: Multiple distribution indicator

https://data.worldbank.org/topic/climate-change