Name: Nethravathi Shivachar Gurumallesh

Student Number: 22089836

Course: Applied Data Science

Module Leader: Ralf Napiwotzki

GitHUB Report Link: https://github.com/nethra791/ADS1-Visualization

Data Links: https://data.worldbank.org/indicator/EG.USE.ELEC.KH.PC for bar graph.

https://data.worldbank.org/indicator/AG.LND.ARBL.ZS for line plot

https://data.worldbank.org/indicator/SP.POP.TOTL for pie chart

Contents

Increasing Insights through Visual Data Analysis: A Co2 Emission Trends Study Using Python's Visualisation Libraries

Abstract

This study examines the crucial function of data visualisation in data analysis, emphasizing its importance in understanding and extracting useful insights from complex data (Gbadago, Moon and Hwang, 2023). The study focuses on CO2 emission data from the World Bank and highlights the use of Python libraries, especially Pandas and Matplotlib, to process and represent data in various visual ways. Findings show that information presented visually, such as in line graphs, bar graphs, and pie charts, is more easily understood. The reasons for selecting two other graph methods are: a bar graph displays information in the form of rectangular bars whose lengths are proportionate to the values that the bars represent and a pie graph is a way to show data visually. It is a circle chart with slices in it, and each slice shows a part or amount of the whole.

Developments in purchasing land, differences in energy use among countries, and shifts in population distribution can all be better understood with the help of these visualisations. This research contributes to the growth and understanding of data visualisation in modern data analysis by highlighting the importance of visualisations in highlighting how excellent graphical representations aid in grasping complex data, leading to faster and more accurate conclusions.

Introduction

The human brain dedicates over half of its resources to processing and feeding back visual information, and the human eye is capable of proficient pattern recognition. Information contained in data can be discovered for the first time using data visualisation graphics because a human brain is more intuitive & specific in its recognition of aspects like sizes, graphics, & colours (Li, 2022). Bar charts, column charts, histograms, & other statistical charts are all examples of data visualisation tools that can be used to effectively and efficiently convey numerical and graphical data and image information. Python is a good choice for data visualisation due to the abundance of helpful third-party modules, vibrant open-source communities, & up-to-date documentation.

Data visualisation in this study is achieved using the pandas and matplotlib libraries in Python(Cao *et al.*, 2021).

1.0 Dataset Description

In this task, a dataset on CO2 emissions has been obtained from the website of the World Bank. Total rows in the dataset equal to 20216, and there are 66 columns in total.

2.0 Data Visualisation Results

This section provides a detailed summary of the data visualisations created from the World Bank's CO2 emission dataset, focusing on three major chart types.

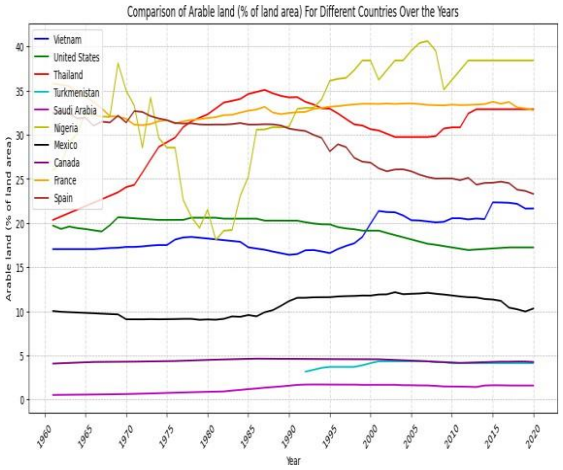

Figure 1: Line chart

Figure 1 depicts how 10 distinct countries take up land on an annual basis. The graph indicates that the quantity of land purchased is increasing annually. And from 2005 to 2010, "Nigeria Country" purchased the greatest amount of land.

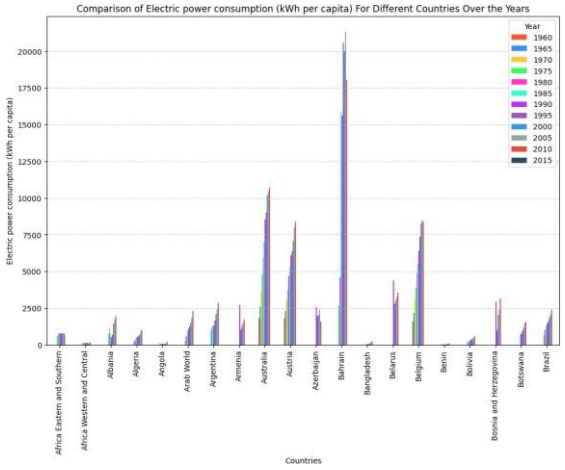

Figure 2: Bar chart

The amount of energy consumed by the 20 greatest countries is depicted in the bar chart that follows Figure 2. The amount of energy consumed varies from year to year. According to this bar chart, Bahrain consumes the most electricity, while Benin consumes the least.

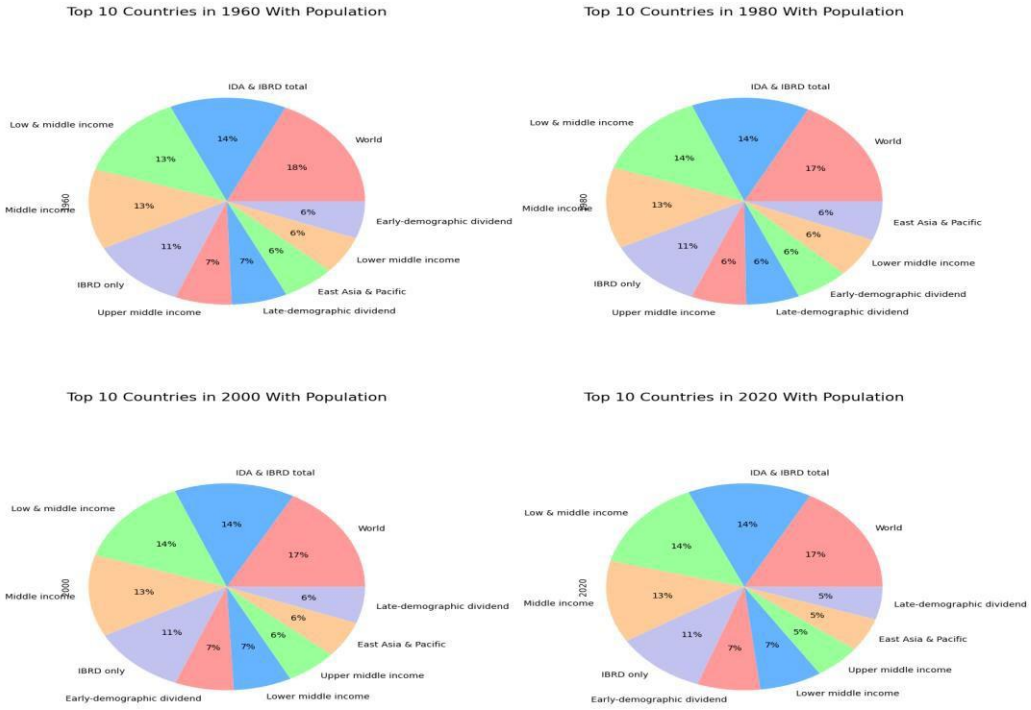

Figure 3: Pie Chart

From 1960 to 2020, the population distribution among the top ten countries is depicted in a pie chart that reveals intriguing trends across years. As of 1960, the country classified as "world" possessed the largest population share, approaching 18%. In contrast, the countries classified as "lower middle income," "early-demographic dividend," and "east Asia & Pacific" each displayed the smallest population representation, approximately 6%. As of 1980, the 'world' country maintained its preeminent position with a population share approaching 17%. East Asia & Pacific, Upper Middle Income, Late-Demographic Dividend, Lower Middle Income, and Early-Demographic Dividend all had population shares of Approximately 6%. As the year 2000 approached, the 'world' country maintained its greatest population share at around 17%, whereas the 'late-demographic dividend,' 'east Asia & Pacific,' and 'upper middle income' countries maintained the lowest population shares at around 6%. In conclusion, the 'world' country maintained a substantial population share of approximately 17% in 2020, while the 'late-demographic dividend,' 'east Asia & Pacific,' and 'upper middle income' regions possessed the smallest population shares, each approaching 5%.

References

Cao, Shengjia *et al.* (2021) 'Research on Python Data Visualization Technology', in *Journal of Physics: Conference Series*. Available at: https://doi.org/10.1088/1742-6596/1757/1/012122.

Gbadago, D.Q., Moon, J. and Hwang, S. (2023) 'Exploring advanced process equipment visualization as a step towards digital twins development in the chemical industry: A CFD-DNN approach', *Korean Journal of Chemical Engineering* [Preprint]. Available at: https://doi.org/10.1007/s11814-022-1273-2.

Li, F. (2022) 'Research on Data Visualization Technology Based on Python', *International Journal of Multidisciplinary Research and Analysis* [Preprint]. Available at: https://doi.org/10.47191/ijmra/v5-i5-03.