Renewable Energy Data Visualization: A study with Open Data

Gustavo Santos[1] Faculdade Nova Roma

Artur Lima[2] Faculdade Nova Roma

Lucas Pereira[3] Faculdade Nova Roma

Adrian Lauzid[4] Faculdade Nova Roma

Davi Maia[5] Centro de Informática CIn-UPE

Abstract

This study explores energy issues across various nations, focusing on sustainable energy availability and accessibility. Representatives from all continents were selected based on their HDI values. Data from Kaggle, spanning 2000-2020, was analyzed using Python to address questions on electricity access, renewable energy generation, and fossil fuel consumption. The research employed statistical and data visualization techniques to reveal trends and disparities. Findings underscore the importance of Python and Kaggle in data analysis. The study suggests expanding datasets and incorporating predictive modeling for future research to enhance understanding and decision-making in energy policies.

Index terms: Data Visualization, Open Data, Renewable Energy.

1 Introduction

Studying environmental data without any help can be seen as inefficient in some cases, as it requires prior knowledge. It is in this context that the visualization of environmental data plays an important role, making the perception of the state and trends of the environment understandable to a greater number of people. It makes it possible to reach concrete conclusions on topics such as air quality, water quality, biodiversity, climate, among others [1, 2].

Helping people with little access to information is a challenging task, and it is important to use simple and accessible resources. Open data proves to be extremely useful because it fits these requirements, is easily accessible and can be presented as a way of raising awareness. It can be used without restriction to raise awareness about public policies, the importance of sustainability and public health issues through maps and graphs [2].

This type of visualization is fundamental for identifying, analyzing and developing functional solutions to environmental problems. The visualization of environmental data is crucial for the creation of public policies, since in addition to helping to identify environmental problems where it makes it easier to locate areas that need more attention, it also ensures decision-making that takes important information into account.

Meanwhile, open data on renewable energy also proves to be extremely relevant, allowing the study of the impact of renewable energy installations and possible solutions, providing a solid basis for academic research, facilitating the monitoring of greenhouse gas emissions and influencing the creation of public policies, after all, it is transparent data that points out important aspects of public regions and can help predict the impact of technological advances and population growth.

Given these factors, the aim of this project is to create a visualization of open data collected over a 20-year period on renewable energy in some countries, covering topics such as renewable energy generation capacity, consumption and average access to electricity.

1 e-mail: gustavo.silva@novaroma.edu.br
2 e-mail: artur.silva@novaroma.edu.br
3 e-mail: lucas.souza@faculdadenovaroma.com.br
4 e-mail: adrian.lauzid@faculdadenovaroma.com.br
5 e-mail: djmm@cin.ufpe.br/davi.maia@faculdadenovaroma.com.br

Dealing with data is not as complex a task as you might think. Although it is extremely useful to everyone in general, data analysis and manipulation are relatively simple, so that many people are able to deal with data through intuitive and efficient manipulation techniques, thus generating useful and relevant work. These techniques can be applied in various areas, such as health, where data analysis can help determine the likelihood of a region suffering an epidemic.

In today's society, not everyone has free access to information or has extremely limited access, either through the internet or television. With open data, it is possible to bring information and awareness to a considerable part of the population, allowing them to use only the traditional resources in their homes to access and consult this data.

Data visualization helps us understand sustainability as something that is not exclusively affected by environmental causes, as is commonly thought, such as the poor distribution of recyclable waste and deforestation. It is something much broader that is totally linked to factors such as politics, public health, the economy, and the relationship a nation has with others around it. For example, a country that consciously consumes electricity will not be able to be sustainable if it has a turbulent economic relationship with its neighboring countries.

2 Related Literature

The article [2] presents a study demonstrating that data visualization tools, specifically Dashboards and Balanced Scorecards, can be used for creating strategic decisions about renewable development and showcasing the benefits of renewable energy for the energy industry. Although our project does not directly focus on or mention the commercial aspect, it aims to understand the availability of renewable energy, which can be used in developing strategies for the commercialization of renewable energies. Another similarity is the multidimensional approach used in both our project and the study in the article. In the article, the analysis is conducted considering various perspectives such as financial, sustainability, business processes, and innovation.

[3] is a study that seeks to present the impact of immersive visualization technologies, specifically virtual reality (VR) and augmented reality (AR), on the development and creation of renewable energy systems. Similar to the article, our project has an educational and informative character, as it aims to provide answers to questions about energy access and renewable energy generation. Presenting insights and information in the same way that immersive visualization tools can educate and inform about renewable energy systems. Another similarity is the focus on improving decision-making, using data, in the development of renewable energy.

In the article [4] the issue of remapping territories around the world is addressed, assessing each one's renewable energy potential. This initiative arose from concerns about climate change and the availability of fossil fuels, with a particular emphasis on the RERMI project in Vietnam. Our project also aims to analyze the political and socioeconomic impact by examining the variation in energy access, the average access to electricity, and the capacity for renewable energy generation, issues that directly influence the development of the analyzed countries. Just as the study examines the incentives promoted by the World Bank and other institutions for remapping in countries of the Global South, highlighting the social and political dynamics of these

countries, we have conducted a specific analysis of the access to and development of sustainable energy in countries in this part of the globe, such as Madagascar, Pakistan, Papua New Guinea, and Haiti. Another similarity is that both projects aim to use the results of the analyses conducted on sustainable energy in different countries to influence policies and investments regarding these issues.

[5] Explains basic methods and instruments to source, retrieve, and interpret analytical open data using APIs and Python. It highlights the practical uses of Python in the collection and exploration of sparse data where there are multiple sources and emphasizes the relevance of open data in enhancing transparency, innovation, and decision-making processes. Open data in addition to the flexibility of python makes it easier for researchers, policymakers, and developers to extract data and or solve problems faster to utilize the data and address global problems including renewable energy development. Ours, like many other projects, can use Python to conduct a simple analysis, insight, and visualization of the data. This is so that policymakers can make better debates, discussions, and decisions in seeking energy access and energy development across different parts of the world.

3 Research Methodology and First Results

This study aims to examine energy issues in different countries, with emphasis on the availability and accessibility of sustainable energies. The study focuses on all continent, of which there will be two representatives per continent based on their HDI (Human Development Index): Africa-Madagascar and Egypt, Asia-Pakistan and China, Oceania - Papua New Guinea and Australia, America-Haiti and Brazil and Europe-Greece and Germany. The data for this study were obtained from publicly available datasets on Kaggle. As shown in Figure 1.

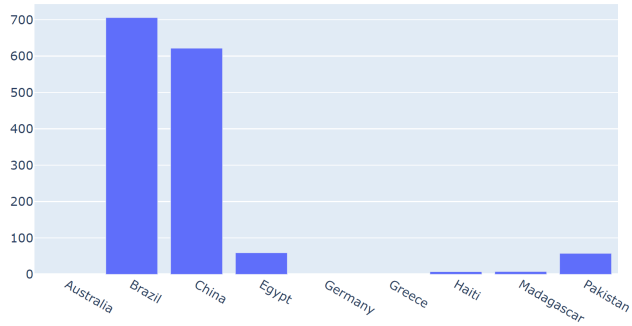

Figure 1: Per capita renewable energy generation capacity by country

Several research questions guide the study. One of these is to understand how levels of access to electricity differ among these nations by assessing its scope as well as distribution. Another question it seeks to answer is the average number of people who have access to electricity between the year 2000 and 2020 arrived at by summing up all the years values then dividing by the number of years. Next, it reveals which countries had renewable energy per person, thus dividing the total amount of renewable energy generation by population size. It also measures fossil fuel electricity generation in terawatt-hours (TWh) from 2000 to 2020 by combining yearly information together.

Finally, we will plot a normal curve for Brazil's renewable energy generation that includes calculating its mean score as well as standard deviation before. As shown in Figure 2.

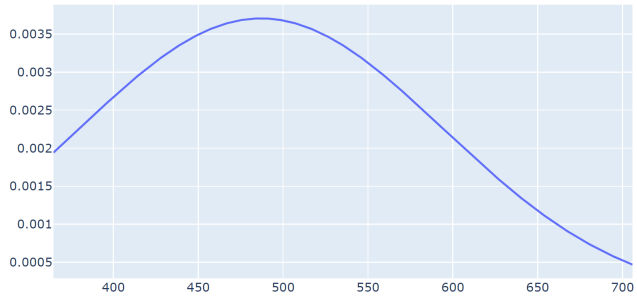

Figure 2: Renewable energy generation capacity per capita in Brazil

The complete source code utilized for this study is available at this link: https://github.com/GustavoSS01/Dash-APS.git

Many statistical techniques were used to explain the provided data, such as average percentage calculation, fossil fuel electricity generation measurement, per capita renewable energy generation, normal curve analysis and variation space analysis. For all data analysis and visualization tasks Python was employed using its strong libraries such as Pandas, NumPy and Matplotlib [5]. These methods gave a broad understanding of energy access and industrialization, which helped evaluate future prospects for sustainable energy cultivation in the countries selected [2].

Python is an ideal choice for this study because it is easy to use, offers diverse online resources, such as courses, tutorials, and guides in multiple languages, most of those without cost attached to it. This wide availability allows people from all backgrounds to access and learn Python, making it a versatile tool for data analysis [5].

A clear and comprehensive presentation of complex data is made possible by the use of data visualization techniques, which help in making patterns and trends easy to follow. It enables us to transmit results to a wider audience including policy-makers, stakeholders among others who would be able to use the information for decisions regarding energy policies as well as investments [2].

The solution developed for this study was deployed on the follow link: https://dash-gunicorn-app-server.onrender.com

This paper carries out a comparison of energy accessibility and sustainability in various regions, revealing gaps and possible areas for further assessment [4]. By examining industrialized and developing nations alike, this report provides useful information on the overall trend of sustainable power world over, which can help guide equal access plans as well as facilitate the move to renewable sources worldwide [3].

4 Conclusion

In an understandable and easy to understand way, this study improves data visualization by making complex data appear simpler. In this way, the discovery of patterns and trends is simplified for energy policymakers and investors who are interested in making informed decisions. Consequently, this not only makes the data more comprehensible but also simplifies the process of sharing results with a wider audience, thereby generating actionable insights [2].

On one hand, Kaggle has played a big role in supporting data visualization as it offers high-quality publicly available datasets that have been used as the basis of analysis. Therefore, reliability and coverage of Kaggle's datasets enhance the validity of this work by giving us accurate and extensive visualized information. Also, Kaggle's platform creates room for interaction where data scientists can share ideas and experiences, thus promoting innovation in analytics [5].

Authors such as [6] also claim Kaggle as an important tool for learning the fundamental concepts of data science, especially for its GUI-based integrated Python IDE. However, according to [6], Kaggle also has limitations, such as less support for open data standards, which means that several organizations can simply share their data with the public without further considerations towards the platform's support for open data standards. Furthermore, [6] claims that both the supply and quality are completely dependent upon the data providers.

Python is indispensable because it is easy to use, with online resources and libraries such as Pandas, NumPy, and Matplotlib. The project utilized these features to undertake complex data manipulations and produce detailed visualizations. We have improved the application of Python in this study outcome by exploiting its versatility in handling various data analysis tasks efficiently that guarantees robust and meaningful visual representations [5].

In future research, we intend to do an enlargement of the dataset to include more nations and longer periods of time to provide a more holistic understanding. The incorporation of machine learning algorithms could help in predicting the future energy access trends and sustainability. Moreover, bettering interactive visualization tools would go a long way into engaging users more while allowing them to do real-time exploration of data, which supports strategic decision-making for policy development too [3].

References

[1] Awari Ltda., "Análise de Dados Ambientais: Descubra as Tendências e Impactos no Meio Ambiente." https://awari.com.br/analise-de-dados-ambientais-descubra-as-tendencias-e-impactos-no-meio-ambiente/ (accessed Jun. 1, 2024).

[2] Lea, B. R., Yu, W. B. V., and Min, H. "DATA VISUALIZATION: A CASE STUDY IN RENEWABLE ENERGY." Issues in Information Systems, , 2017, 18.

[3] Rozmi, M. D. A. B., Thirunavukkarasu, G. S., Jamei, E., Seyedmahmoudian, M., Mekhilef, S., Stojcevski, A., & Horan, B. "Role of immersive visualization tools in renewable energy system development." Renewable and Sustainable Energy Reviews, 2019, 115, 109363.

[4] McCarthy, J., & Thatcher, J. "Visualizing new political ecologies: A critical data studies analysis of the World Bank's renewable energy resource mapping initiative." Geoforum, 2019, 102, 242-254.

[5] Anjaneya Reddy N M "Open Data and APIs-data extraction and exploration using python" 4th LIS Academy Conference on 'Open Scholarship and Libraries, 2021.

[6] Mohsan Ali "A comprehensive review of open data platforms, prevalent technologies, and functionalities". 15th International Conference on Theory and Practice of Electronic Governance. 2022