KU LEUVEN

Visual reality

Data Visualization in Data Science

May 27 2021

Rana Cansu Kebabcı	r0772826
Ömer Yiğit	r0767950
Sybren De Boever	r0583966
Maxime Keutgen	r0828809

Project: Energy across the world

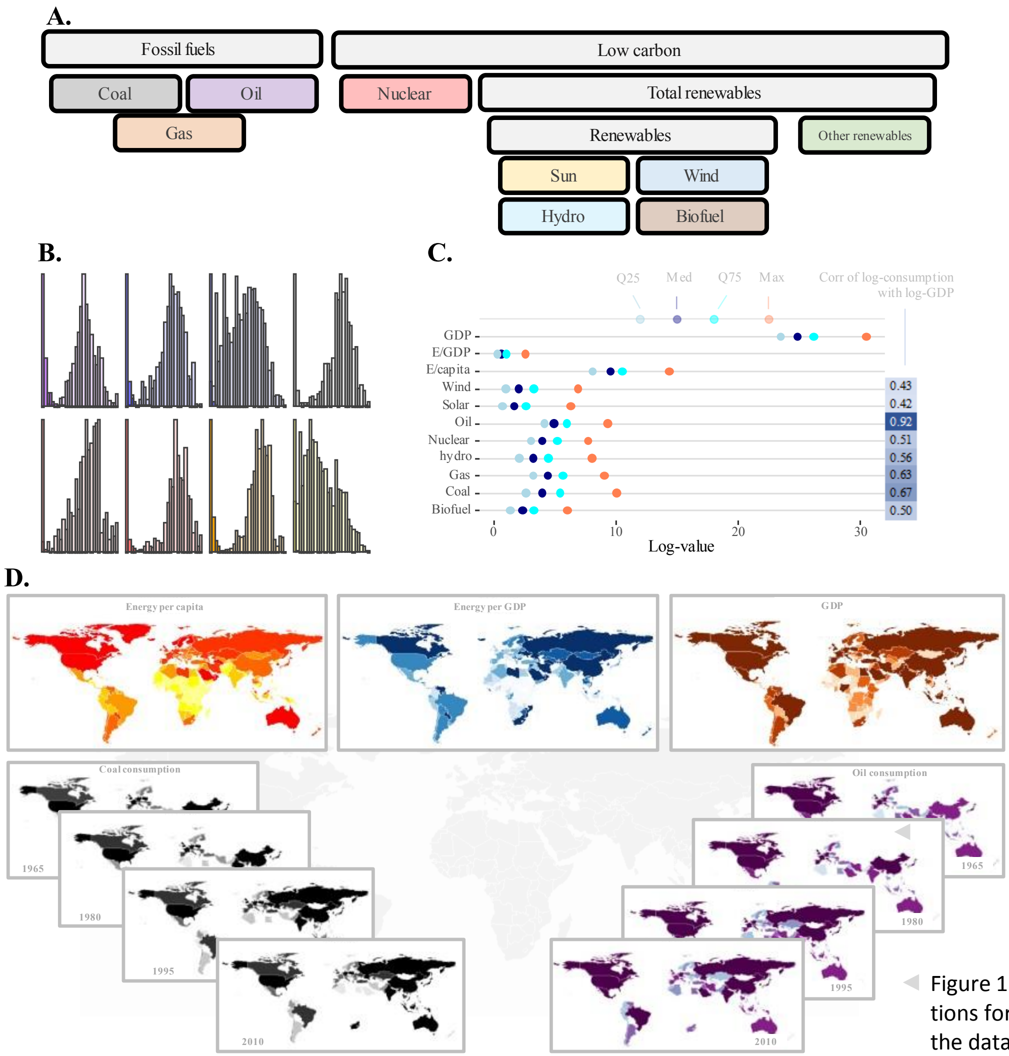
Video Link: https://www.youtube.com/watch?v=aESSk4CHe8c

1. Introduction

Energy source	% electricity	% growth	% sub energy	Consumption (TWh)	Consumption (kWh/ capita)	Consumption change (TWh)	Electricity generation (TWh)	electricity (kWh/capita)	production	production/ capita	Anual change in production (%)	(kWh)
	Collected data associated with the source											
Biofuels		•	•	•	•	•						
Coal	•	•	•	•	•	•	•	•	•	•	•	•
Fossil fuel	•	•	•	•	•	•	•	•				
Gas	•	•	•	•	•	•	•	•	•	•	•	•
Hydro	•	•	•	•	•	•	•	•				
Nuclear	•	•	•	•	•	•	•	•				
Oil	•	•	•	•	•	•	•	•	•	•	•	•
Solar	•	•	•	•	•	•	•	•				
wind	•	•	•	•	•	•	•	•				
Renewables	•	•	•	•	•	•	•	•				
Other renewables	•	•	•	•	•	•	•	•				
Low-carbon	•	•	•	•	•	•		•				

Background variables
Year
Country
Energy/GDP
Energy/capita
Electricity/capita
Population
primary energy consumption
GDP
HDI
Continent

Table 1. Variables included in the dataset. Dots in the large table indicate the presence of the corresponding energy measure for the energy resource. The small table presents additional non-resource related variables.



Description of the project

Our project aims to investigate (1) the evolution in the use of different energy sources throughout the world and (2) whether the energy characteristics of countries are associated with their wealth. We obtained time series data on twelve energy sources (from here) including over 10.000 rows, over 100 variables, and over 200 countries/regions. In addition, we also obtained geographical data (from here) to link each country to their respective continent. Finally, we also included data on the life quality measure HDI (human development index, from here) that, together with the provided GDP (gross domestic product), provides a general view on countries' wealth. We created visualizations in Vega-lite, Vega, Rstudio, and D3.

With our visualizations, we hope to contribute to sever shortcomings we noticed throughout the energy literature: (1) Worldwide data are frequently used in line graphs, bar plots, pie charts or tables, without providing country specific measures (Cao 2016, Ma 2020, Balat 2006, Kan 2019), (2) A selection of years is sometimes presented in time evolutions rather than all available time points (Cao 2016),(3) Different energy sources are presented individually without a clear visual comparison between them or the countries (Balat 2006), (4) Shifts in energy sources have only been visualized for a few countries over a limited time range (Kan 2019, Stern 2015), (5) No articles were found that provide an extensive investigation of the evolution in energy sources across countries while including their evolution in GDP(Ramanathan 2006).

Description of the data

The retrieved dataset was a large multidimensional dataset, covering 10.134 rows and 119 variables (Table 1, Figure 1D). Twelve energy sources were reported in the dataset, which consisted of both base resources and aggregated resources (Figure 1A). The variables in the dataset provide numeric measures related to the production and consumption of the different energy sources, as well as data on country population sizes and wealth measures. Measures were provided as time series from 1965 to 2019. Missing data was present on the variable level and the measurement level. Not all variables were reported for each energy source (Table 1), and missing values occurred in the time series of resources. Such missingness could be the result of unemployment of the corresponding resource or from lost records.

Some modifications were performed to the original measures. First, we retained data on single countries by excluding aggregated regions that were also reported, such as 'middle Africa', 'Asia Pacific' or 'Other Middle East', retaining 221 countries in total. This allowed us to investigate energy evolutions and transitions on both a country and world level without including countries more than once. Second, we focused on the consumption measures as these provided data for all the energy sources. For these, log-transformations were considered because of their right skewness, indicating the presence of outlying countries with extremely high energy characteristics. This transformed the extreme distributions into more bell-shaped distributions (Figure 1B).

Several summary statistics are displayed in Figure 1C to further explore the energy data across the world. Gas, coal and oil seem to be the most employed resources, which together constitute the branch of fossil fuels (Figure 1A). The least employed resources are wind, solar and biofuels, which are part of renewable energy sources. The distribution for the measure 'energy per GDP (E/GDP)' shows a small spread, indicating a more consistent energy consumption per GDP compared to energy per capita. Figure 1D shows the GDP distribution over the world and their corresponding energy consumption, which seems to be increased for large countries. Finally, correlations between GDP and the energy sources indicate that countries with a large GDP show a clear increase in oil consumption, but less in the consumption of renewable energy.

Figure 1. A. Representation and compartmentalization of the reported energy sources. B. Original and log-transformed distributions for the consumption of oil, hydro, wind, coal (upper row), biofuel, nuclear, gas, solar (lower row). C. Example measures in the dataset.

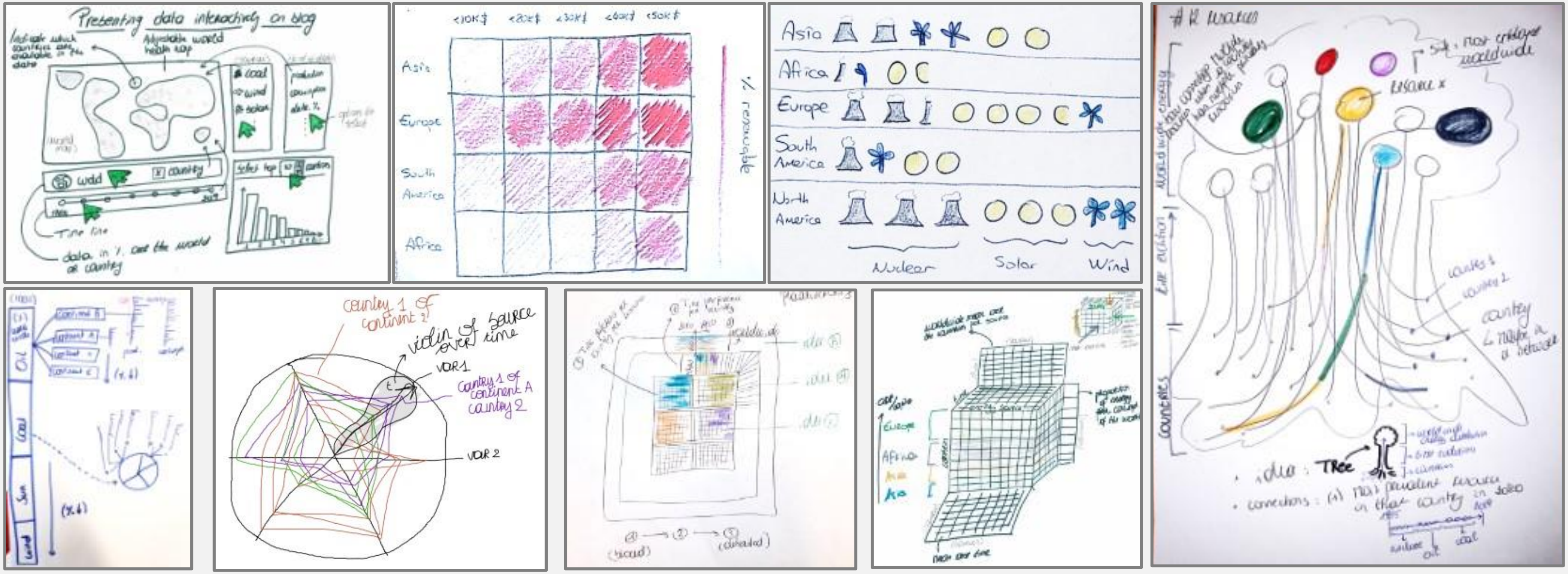


Figure 2. Examples of hand-drawn designs. Some of them are not implemented because of infeasibility.

2. Visualization development

We searched the design space for various visualizations to represent our data as accurately as possible. First, we came up with simple designs like area chart or plots with symbols to represent a specific part of the data (Figure 2). Then, we created designs that can encompass as much data as possible, like a world map dashboard or a heatmaps. All our initial sketches can be accessed at the Miro board.

Research questions

Our research questions are phrased broadly to have a larger design space to explore and to be able to come up with numerous ideas for visualizing the different aspects of the patterns in the data.

1. "How does the global energy consumption evolve over time and how do countries and continents differ from each other?"

2. "What is the quantitative and qualitative relationship between wealth and energy?"

Visualizations were created in:

- Vega-lite
- Vega
- D3
- Rstudio
- Tableau

2.1 General visuals

An interactive dashboard is created on Tableau to see the energy patterns over time for different countries. In general, USA, and recently China, are the two biggest energy consumers in the world, mainly because of their high population and industrial development. In energy production, Russia joins them. In clean energy share, different countries lead for different energy sources, e.g. Brazil for biofuel (8%), Denmark for wind (21%), Japan for solar (4%) in 2019. Also, some small countries, e.g. Oman and Trinidad & Tobago, consume only fossil fuels. In the next step, more specific plots are created in Vega and Vega-lite to see how we can answer our research questions. We also had inspirations from the visuals made by others for the similar subjects.

Figure 3. Interactive world map dashboard (accessible here). Yearly visualization for different energy sources and measures in the world. Top countries for the selection are shown below. Detailed insights can be accessed here.

Figure 4A illustrates well the dominance of fossil fuels on every continent. However, proportions do not give insights about absolute values. As an example, Figure 4B illustrates the electricity consumption per energy source of each continent. Each emoji symbolizes a 1000 GWh (gigawatt hour) of energy consumed. The consumption of Africa is marginal compared to other continents and it fully relies on fossil fuels to generate electricity. Worldwide electricity consumption remains largely dominated by fossil fuels, with Asia being by far the leader. Europe is still consuming three times as much oil and gas than renewables or nuclear. Europe is also the only continent to rely on a substantial amount of nuclear energy for its electricity generation.

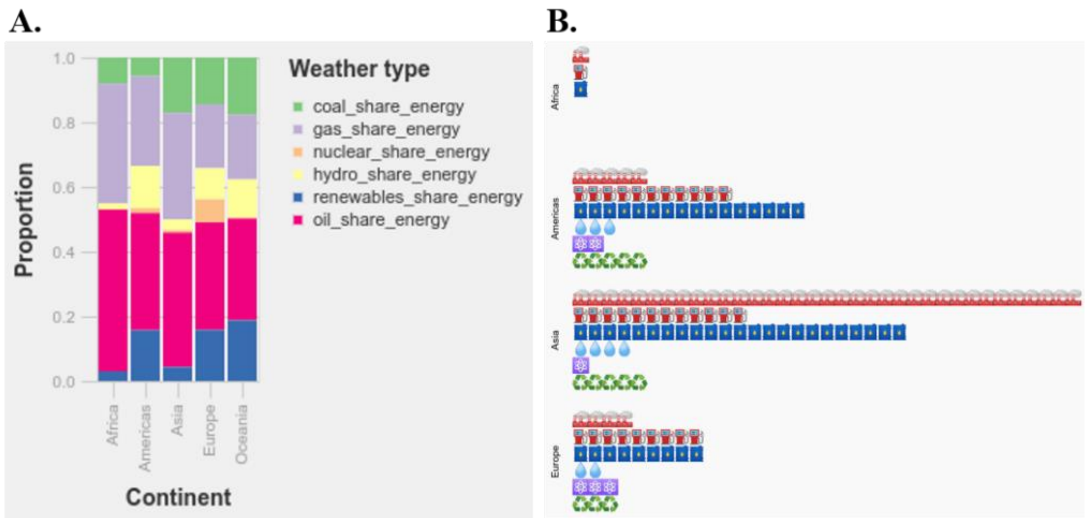


Figure 4. A. Relative amount of energy per resource for continents. B. Emoji plot showing the energy consumption per continent.

2.2 Visual development of research question 1: country, continental and global energy evolutions

First, an animated bar chart (Figure 5A) is created in D3 to visualize the worldwide (log-transformed) energy consumption per capita for different energy sources over time. If the same increasing pattern can be extrapolated in the future, clean energy sources may dominate in the future. Second, an interactive line plot (Figure 5B) is created for different measures on all energy types. Low-carbon and fossil energy sources create two clusters of lines recently. Then, two stacked area charts (Figure 5C) are created for consumption per capita and for energy share. The aim for these plots is to observe if different patterns occur indifferent time periods in different continents.

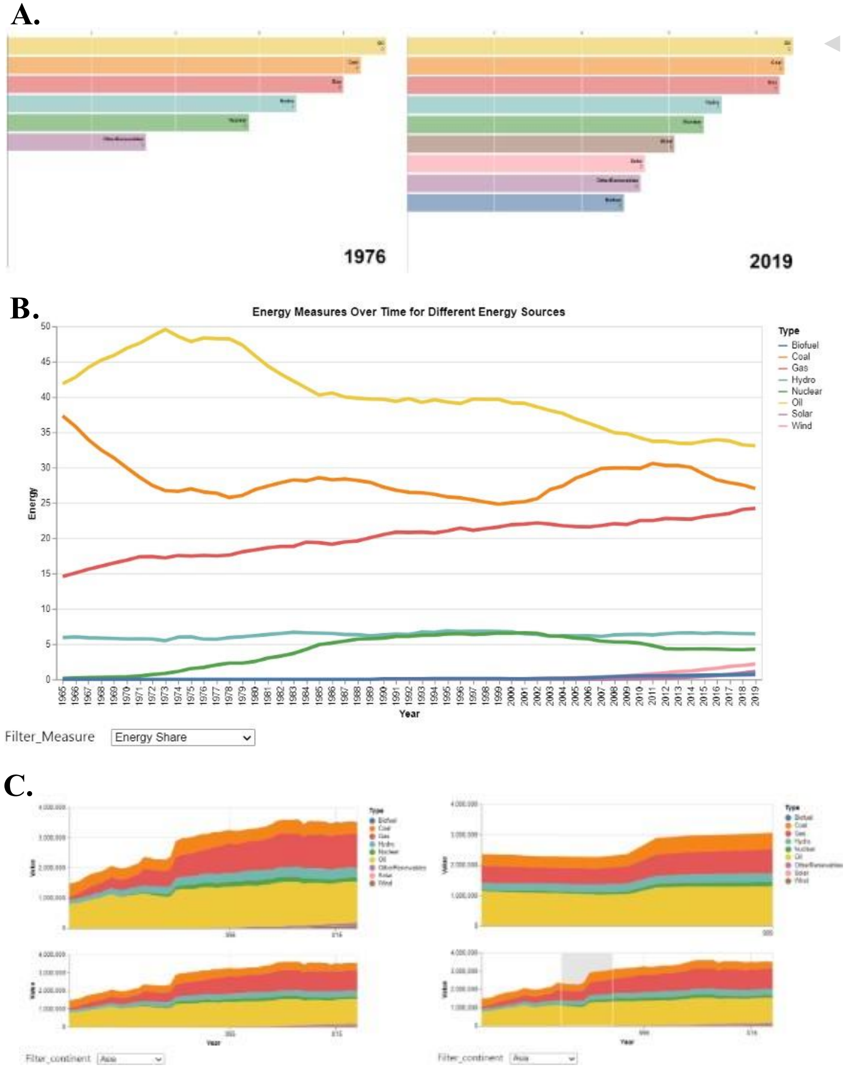
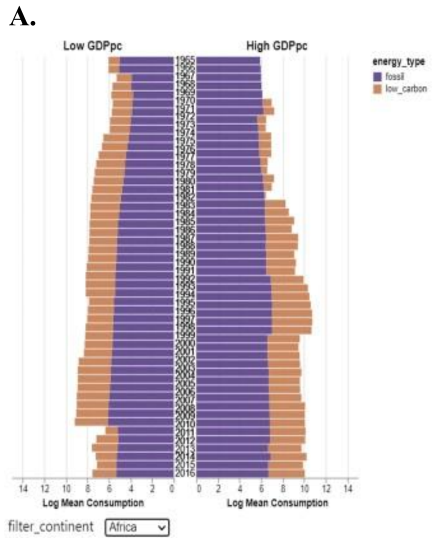
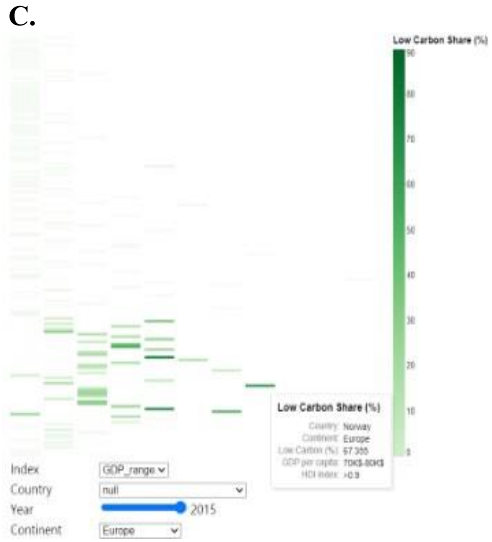


Figure 5. Preliminary interactive plots for energy evolutions.





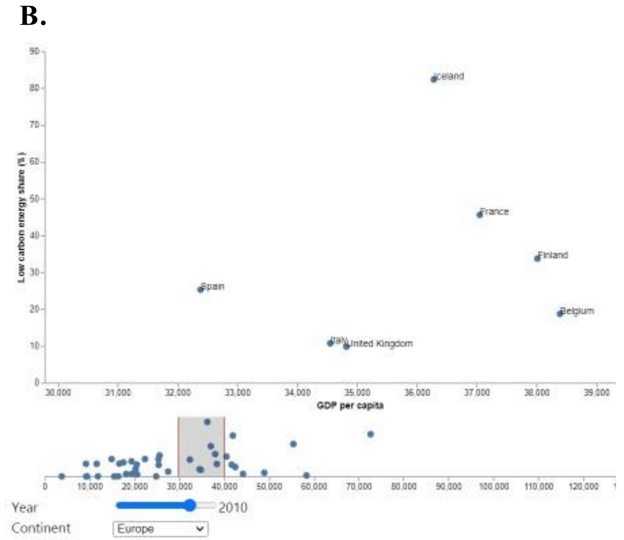


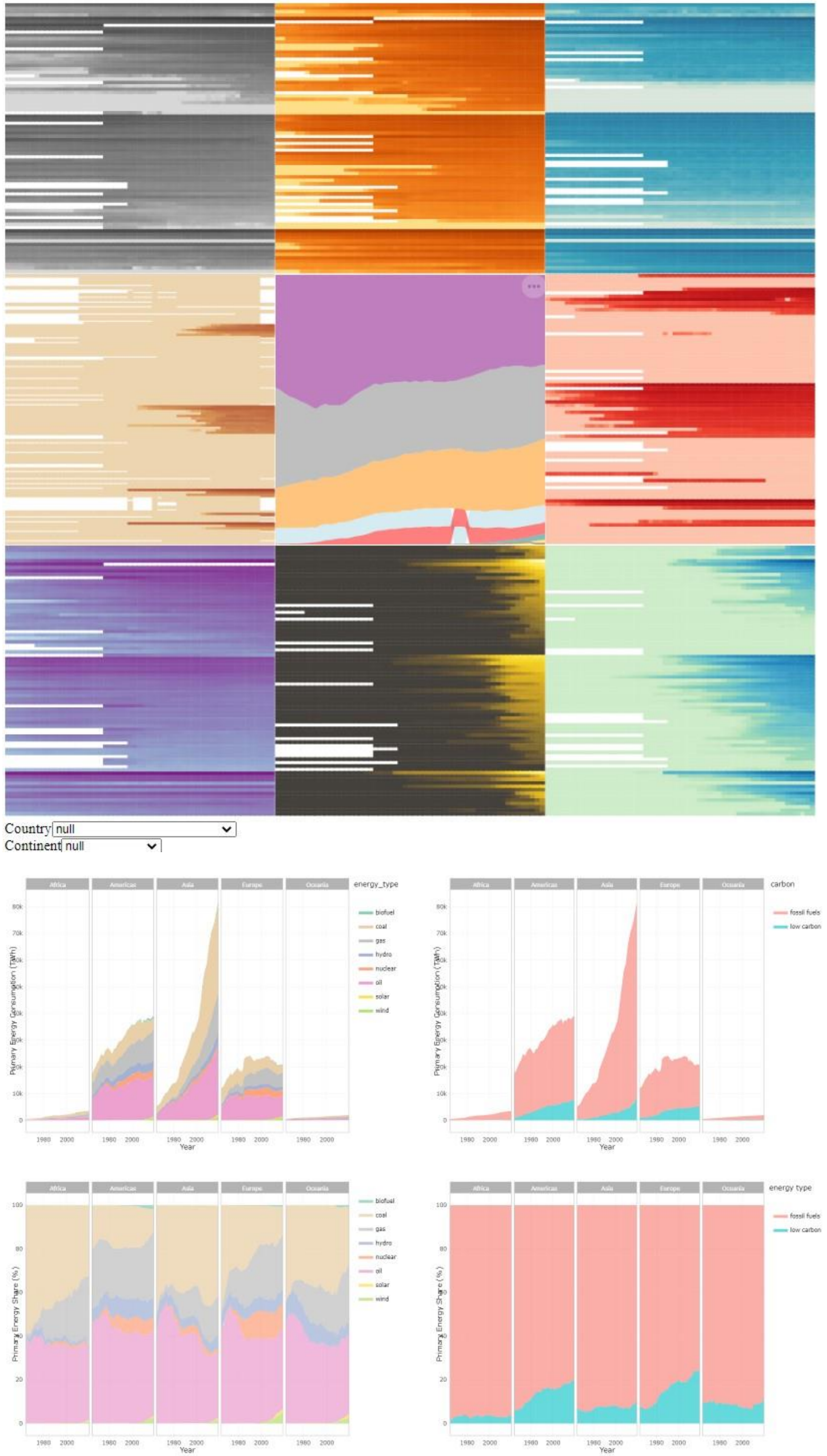
Figure 6. Preliminary plots for wealth-energy relationship.

2.3 Visual development of research question 2: energy and wealth

First, we plotted an interactive pyramid plot with discretized wealth measure (log-GDP per capita) based on its median (Figure 6A). It is a compact graph in which the evolution of the clean energy share per continent can be tracked. Then, a zoomable scatterplot is created as an improvement to capture more detail (Figure 6B). It enables to focus on different ranges in wealth. Additionally, a detailed heatmap is created to compare the progresses of countries and continents over time (Figure 6C). Overall, there does not always seem to be a correlation with GDP per capita and low-carbon energy use, yet the correlation mainly depends on region and time. These graphs were further developed for the final visualizations.

A.

B.



3. Implementation and annotation of final designs

3.1 Research question 1: country, continental and global energy evolutions

Aim of the visuals

With our final implementation, we aimed to improve the shortcomings in literature where often a restricted part of the data is shown, i.e. only aggregated world or continent data, only a few countries, or discretized time series. With our heatmap approach created with a combination of Vega and HTML (Figure 7A), we were able to combine (1) complete time series (2) for all individual countries (3) for all base resources (4) with aggregated world data. In addition to country specific data and worldwide data, we also present continent aggregated data (Figure 7B), which did not fit the heatmap visual, for the base energy sources (left) and grouped resources (low-carbon vs. fossil fuels). The latter was created using plotly in R.

Understanding the visuals

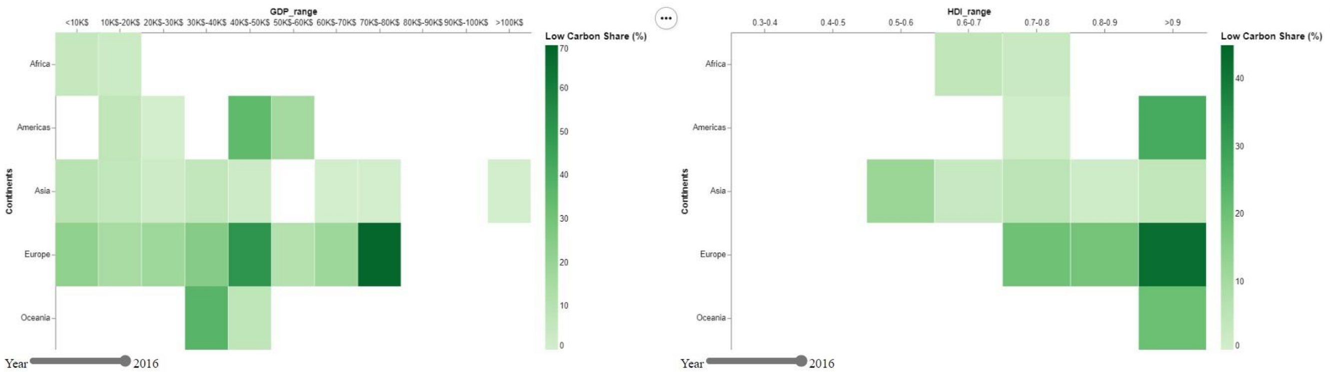
Figure 7A displays the total consumption for coal (grey), gas (orange), hydro (blue), biofuels (brown), nuclear (red), oil (purple), solar (yellow) and wind (light blue) for every country (rows, grouped per continent) per year between 1965-2019 (columns). The central visual shows the sorted world aggregated measures for the respective energy sources. The philosophy of this plot is that the outer layer, i.e. the individual country data, constitutes the central global energy consumption. Three levels of interactivity are added to improve visibility and comparability: mouse hovering highlighting individual countries, which can be combined with selection boxes for both countries and continents. In addition, Figure 7B shows absolute and relative continental stacked plots for the same energy sources over time. These are also implemented with interactivity, allowing individual selection of specific resources.

Insights obtained from the visuals

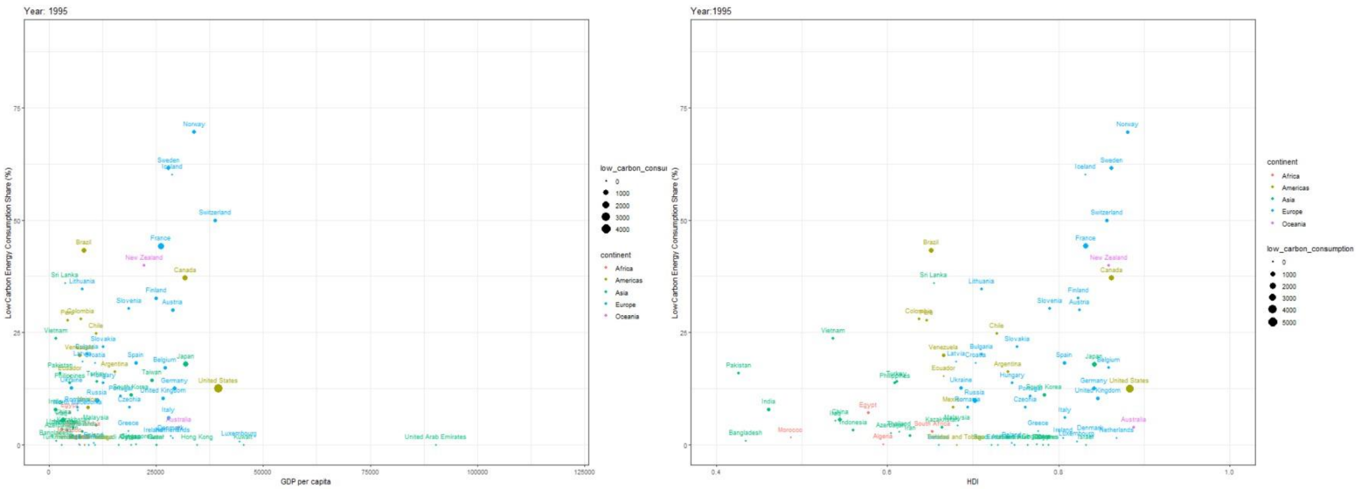
1. On a global level, fossil fuels (coal, oil and gas) are still dominating, but the use of renewable resources is slowly increasing.
2. Most of the countries use fossil fuels at a seemingly steady rate.
3. Many countries have started employing renewable resources, of which hydroelectric is by far the biggest in volume. The use of biofuels, wind and solar is only apparent in a limited number of countries. The use of nuclear energy hardly occurs in a country's timeline when it had not been employed in the very past (1970's).
4. Different continents show different energy characteristics: Africa reports little energy consumption compared to Europe, America and Asia, despite that its population size amounts twice that of Europe.
5. Different energy transitions occurred throughout history. For instance, the relative use of coal decreased in Africa and Europe, associated with an increase in gas, while Asia and Oceania decreased their oil usage and Europe increased its nuclear energy consumption.
6. The relative usage of low-carbon resources increased most for America and Europe. This is the result of a substantial use in nuclear energy for Europe, and nuclear and hydro energy for America.

Figure 7. A. Interactive heatmaps, accessible here where you can also find a description of the philosophy and implementation of the visual. B. Upper row: evolution in raw energy consumption across continents. Interactive version here. Lower row: Relative energy consumption across continents from 1965 to 2015, interactive version here.

A.



B.



3.2 Research question 2: wealth and low-carbon energy consumption relation

Our assumption is that clean energy resources are favored by more wealthy and developed countries. We think that those countries are more conscious of climate change and have the financial and technological resources to invest more in clean energy. To investigate this relation, energy types are grouped into two major groups such as fossil and low-carbon energy, which are considered polluting and clean resources respectively. The reason for the grouping is that the consumption of each clean energy resource namely, the solar, wind, hydro, and nuclear are almost neglectable when they are considered individually. Thus, in the analysis clean energy consumption is represented by low-carbon energy share in percentages. The wealth is represented by GDP per capita and HDI.

Considering the pros and cons of the previous preliminary visualizations, two types of plots are here produced which are the continent-level heatmaps and country-level animated scatter plots. The heatmaps were plotted in Vega-lite and the animations were implemented in the R software using the gganimate library. The visualizations were restricted to the time period between 1990 and 2016 because the data for HDI was only available for that time period.

Figure 8. A. Wealth vs low-carbon energy share heatmaps, accessible here. B. Wealth vs low-carbon energy consumption animated scatter plots, accessible here and here. In addition to the low-carbon consumption, the raw consump-tions are also shown by the size of the scatter points. The coloring is based on the continent and the wealth measures are used on continuous scales. Countries with complete observations between 1990-2016 are included. Year increases as the animation proceeds.

Insights for GDP heatmap (Figure 8A, left):

1. Overall, GDP per capita increases over the years and the low-carbon consumption remains stable.
2. Different relationships for different continents: none for Asia and Africa, positive for Europe and America (between 1998-2003 and 2010-2016), and reverse relationship for Oceania.

Insights for HDI heatmap (Figure 8A, right):

1. Overall, HDI increases over the years.
2. Different relationships: positive relationship overall years for Europe, Americas and Oceania, but none for Asia and Africa, which also show overall small low-carbon share.

Remarks:

1. Because of the observed differences, HDI might be a better predictor of low-carbon share or it may be the result of the scale difference and discretization.
2. The reverse relation of Oceania is not generalizable because there is not enough data to be grouped into more than two wealth index groups.

Insights for GDP scatter plot (Figure 8B, left):

1. GDP increases over time.
2. Low-carbon energy share is overall low (0-25%).
3. Small low-carbon share + high GDP/capita for Kuwait, Qatar, United Arab Emirates, Singapore, and Luxembourg.
4. Large low-carbon share + high GDP/capita for Iceland, Sweden, and Norway.
5. Positive relationship, especially after 2000, for Europe, and no relationship for the other continents.

Insights for HDI scatter plot (Figure 8B, right):

1. Overall improvement in HDI with countries condensing in the range of 0.6-1 in later years.
2. Large low-carbon share + high HDI for Iceland, Sweden and Norway.
3. Mild positive relationship for Europe, and no relationship for the other continents.

Remarks

1. Both the GDP and HDI scatter plot show that the relation of wealth and low-carbon share is only detectable for Europe.
2. However, Americas, Oceania and Africa are underrepresented, which might impede the detection of a relationship.
3. In later years, the relation is hard to be detected for HDI because of condensing.

4. Conclusion

Research question 1:
country, continental and global energy evolutions

We were able to visualize energy evolutions for individual countries and each energy source, in combination with continental and global evolutions. On a global scale, oil is the largest energy source, followed by coal. Although the share of low carbon energy is increasing rapidly, its rate of growth is not fast enough to offset the growth of fossil fuels consumption. We also saw that continents are characterized with different energy transitions : Europe reduced its coal share at the benefits of gas and nuclear. O

Research question 2:
relationship between wealth and energy

With our visualizations, we were able to see that wealth only had a mild effect on improving the clean energy share of countries and continents. However, because of representability issue, this could only be concluded for Americas and Europe. We compared the measures GDP and HDI and saw that these gave parallel insights.

All final implementations can be accessed on the collective HTML page.

We would like to thank the assistant's team for the time they were willing to invest in inspecting our visuals, trying to understand them, and providing useful suggestions. We would like to especially thank Danai Kafetzaki for helping us several times with Vega and Vega-lite implementation and her fast replies!

Individual contributions

Cansu: Obtaining HDI dataset – design sketching – data cleaning – dealing with missing values – visualizations in gganimate (R), Vega-lite – presentation script – blog and report writing.

Ömer: Design sketching – data cleaning – dealing with missing values – visualizations in Vega, Vega-lite, D3, Tableau – presentation script – blog and report writing.

Sybren: Obtaining energy dataset – design sketching – visualizations in Vega and Vega-lite – creating HTML pages – logo and report layout – blog and report writing.

Maxime: Obtaining energy dataset – design sketching – data cleaning – visualizations in R, plotly and vega – blog and report writing – video recording.

"Yes we made our own logo"

- Visual Reality, 2021

References

Balat M. Electricity from worldwide energy sources. Energy Sources, Part B: Economics, Planning, and Policy (2006). 1(4):395-412.

Cao X, Dai X, Liu J. Building energy-consumption status worldwide and the state-of-the-art technologies for zero-enery buildings during the past decade. Energy and Buildings (2016). 128:198-213.

Kan S, Chen B, Chen G. Worldwide energy use across global supply chains: decoupling from economic growth? Applied Energy (2019). 250:1235-1245.

Ma X, Fu Q. The influence of financial development on energy consumption: worldwide evidence. Int. J. Environ. Res. Public Health (2020). 17(4):1428.

Ramanathan R. A multi-factor efficiency perspective to the relationship among world GDP, energy consumption and carbon dioxide emissions. Technological Forecasting and Social Change (2006). 73(5):483-494.

Stern DI. Energy-GDP relationship. The New Palgrave Dictionary of Economics (2015).