How would you visualise your data?

COMP5048/4448 Assignment 2

Andrea Bosia
520600843

line 1: Given Name Surname
line 2: *Student ID number*

line 1: Given Name Surname
line 2: *Student ID number*

line 1: Given Name Surname
line 2: *Student ID number*

line 1: Given Name Surname
line 2: *Student ID number*

line 4: Your tutorial group number
line 5: Your tutor's name

I. Task 1

For the first task the objective was to identify groups of similar countries leveraging 12 attributes from the dataset provided. To achieve such goal, we firstly divided the time frame into three bins each encompassing 17 years. This rational seemed meaningful as countries undeniably evolve over time and such changes determine deep modification of their industrial backbone as well as of their socioeconomic structure. The latter are of course directly reflected in the GHG and CO2 emissions indexes, as well as in the socioeconomic ones, that are comprised in the dataset used. As countries change over time, their belonging group may change over time. By examining the situation within a narrower timeframe, it becomes possible to observe the fluctuations in the composition of these groups over time. Moreover, throughout the years, different heuristics were implemented to highlight groups of countries by experimenting with distinct visualizations, such as a parallel coordinate plot, scatter plot, and geographic map.

Analysis of countries from 1971 to 1987

A. Pre-Processing

For the analysis of the decade from 1971 to 1987, we tried to highlight clusters of countries based on 3 different heuristics:

- Countries were grouped together based on the amount of GHG emissions caused by the combustion of a fossil fuel (GHG_FC_Coal, GHG_FC_Oil, GHG_FC_Gas).
- Clusters of countries were also found analyzing the amount of Co2 emissions caused by maritime and aviation sector of a country (AVBUNK, MARBUNK).
- Lastly, countries were grouped considering the socioeconomic indexes provided (CO2_TES, CO2_GDP, CO2_GDP_PPP, CO2_POP).

For each of the previous set of variables the raw data required some pre-processing, so that they could be used in producing some effective visualizations capable of highlighting groups. Here's an overview of the dataset modifications:

- Missing values of each of the time series considered were imputed with the closest next available value.
- A scaled version (i.e. MinMaxScaler) of each index was added to the data frame.
- The mean value of each time series was computed obtaining the average of the index over the time.
- percentages were computed by combining fossil fuel-related variables, and similarly for AVBUNK and MARBUNK.

For the socioeconomic variables a further step was taken. Indeed, PCA was performed on the original four variables of this category in order to come up with one principal component that could be used to group countries based on socioeconomic factors. The component obtained successfully captured roughly 70% of the total variance of the dataset. Lastly, we crafted a binned version for each column, tailoring the number of levels to the distribution of the original data.

B. Evaluation

The type of graph utilized to showcase groups of countries is a world map (Figure 1) and is the result of an evaluation process based on firstly interacting with the visualization and then answering a series of open hand questions:

a. Which countries appear to be more similar based on socio economic indexes?
b. Identify groups of countries for which the aviation industry causes more than 75% of the emissions (over the totality of emissions caused by aviation and marine bunkers)
c. Are those countries where the aviation industry causes more than 75% of emission also among the countries with the highest emission values globally for the aviation sector?

Alternative visualizations like tree maps proved not appropriate for displaying data with small values and do not incorporate the geographic information. Also, bar charts were explored but it was impossible to plot a high number of countries at once without overloading the visualization.

C. Final Visualisations

The visualization produced is interactive in that it makes it possible to select a specific index based on which to populate the map. Countries are colored on the map based on their respective values. This results in countries with similar values sharing similar colors, effectively highlighting country groups. Originally, a continuous color scale was used, and the indexes were not binned into groups. This approach although proved not very effective as it was not immediate to grasp the number of clusters present and to which cluster a given country belong. However, once the indexes analyzed were binned into levels, and different colors were mapped to each bin, the clusters became easily visible. An added feature to enhance user interaction is the "hover-over" function. It displays vital information about a specific country when a user hovers over it. Moreover, also a dynamic description of a variable is added.

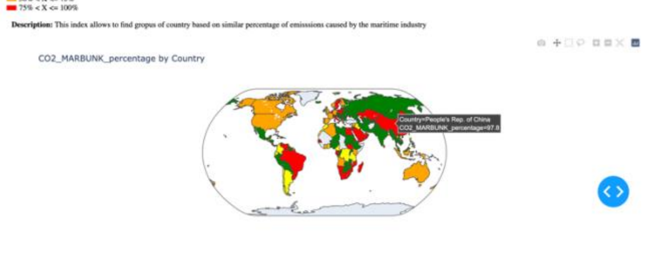

Figure 1 Interactive visualization for analysis of countries from 1971 to 1987 (World Map)

D. Interactive graph of displayed results

The combined analysis of the percentages as well as the original indexes allowed to showcase interesting groupings. For GHG emissions from coal, the top contributors globally include the United States, China, Russia, followed by Poland and Germany. Yet, when examining the percentage of emissions, a different narrative emerges. Some nations, while not major contributors to global coal emissions, heavily rely on coal for their energy needs. For instance, more than 75% of fossil fuel emissions in South Africa, Zimbabwe, South Sudan, and the Czech Republic originate from coal. A similar trend appears when analyzing maritime bunker emissions. Russia significantly impacts global emissions, whereas China's contribution is comparatively lower. However, in terms of percentages, the roles are reversed. For Russia, maritime bunkers contribute to less than 25% of national emissions. In contrast, a staggering 91% of China's emissions are linked to maritime bunkers. Using the Principal Component Analysis (PCA) variable derived from socioeconomic indexes revealed more groupings. Countries like Australia, China, Russia, and those in Europe and North America predominantly fall within the top two categories. On the other hand, most South American, African, Middle Eastern, and Asian nations are clustered in the lower two groups. These visual patterns indeed echo the economic and developmental landscapes of these regions during the time frame.

Analysis of countries from 1988 to 2004

A. Interactive visualization of displayed information

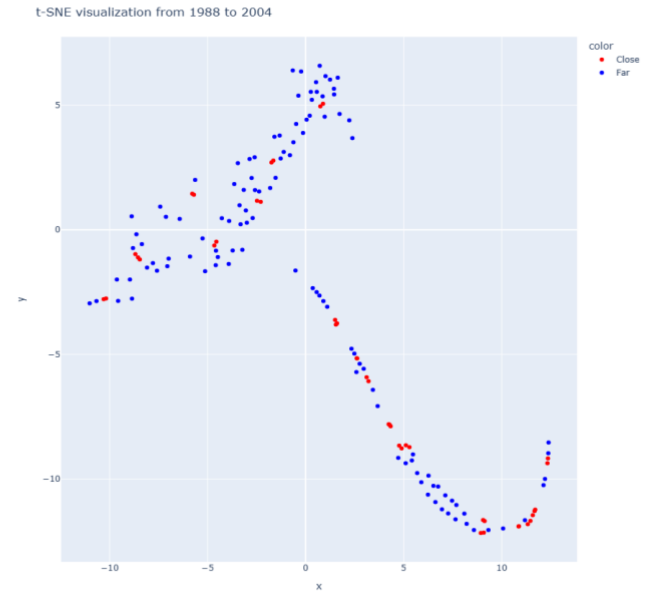

Figure 1 Interactive visualization for analysis of countries from 1988 to 2004 (t-SNE visualization from 1988 to 2004)

As shown in the Fig 1, the dataset is reduced into a two-dimensional visualization image using t-SNE technology. This can be easily obtained through the image title, which contains the 12 attributes data content from 1988 to 2004, and only analyzes its similarity for different countries. From the graph, it can be clearly seen that there are about 200 points representing each country, some of which are closer, and some are farther apart. From the color division on the right side of the image, the points closest to each other are labeled in red, while those farther away are labeled in blue. At the same time, most points are farther apart, and only a small number of points are labeled in red, indicating that only a small number of countries have similarities. The reason why it is possible to determine whether two countries are similar based on the distance between two points is because t-SNE technology can determine the similarity in high-latitude data and save the similarity in two-dimensional space. Thanks to this, we can intuitively observe countries with similarity through the visualization image above. To further understand what countries similar red dots represent, you can use the interactive function to move the mouse cursor over the red dots, which will automatically display the name of the country and its position on the X and Y axes in the two-dimensional image. For example, by moving the cursor over the two red dots on the far left of the graph, information can be obtained that these two countries are: Chinese Taipei and Republic of Türkiye, indicating a similarity between these two countries.

B. Evaluation and modification

After completing the first version of the visualization task, a survey was designed to gather feedback from classmates regarding the visualization content, reflect on the feedback, and make further changes. To demonstrate the design being tested, several closed-ended questions are raised, Also, open-ended questions are also raised to identify potential omissions in the design. The content of the survey is shown in Figure 2.

a. How easy is it to understand?
1) Very Easy 2) Easy 3) Average 4) Difficult 5) Very Difficult
b. Is it objective to notice the interaction button
1) Yes 2) No
c. Can you visually observe points and country names that are closer together?
1) Yes 2) No
d. If you want to change the color of the visualization graphic, which one would you Choose from below?
1) Red 2) Yellow

Fig. 1. Survey for t-SNE visualization

In response to the final feedback results, I have made the following improvements to the visualization content: first, I will consider improving the interaction content: (1) adding interactions that can be used to add or remove countries, so that similar countries can be judged based on the overall visualization graphics of all countries, or interested countries can be individually selected to study their distances in 2D images in detail. (2) Add scaling function, as there are many countries, the connections between different points in the overall visualization content may appear very noisy. Therefore, it is possible to further indent and select the detailed observation points of interest to observe the connections between them. Another consideration is the improvement in visual effects: (1) The original visual content of the country name is directly displayed in the image, which will result in a chaotic effect where two country names overlap. Therefore, it was removed and replaced with a cursor interaction, where the country name is viewed by floating the cursor above the point. (2) The original image was all blue, and one could only rely on intuition to feel whether the two points were close enough. Then, different colors were introduced, and the Euclidean distance between the two points was calculated to determine whether they were close enough. For two points that were close enough, they turned red.

C. *Interactive graph of displayed results*

There are the following explanations for the final visualization effect displayed: (1) There is no specific meaning or representation of measurement units for the use of the X and Y axes in the final displayed graphics. This is because when using the t-SNE method, the final generated two-dimensional visualization graph is only to maintain its similarity in high-dimensional situations. Therefore, the X-axis and Y-axis only represent a direction to display the positions of different points, without direct physical or data significance. Therefore, the positions of each point are obtained through optimization algorithms, so that the closest points in the original high-dimensional space are also close in the two-dimensional space, and points that are far away are also far away in two-dimensional space. (2) The use of two different colors, red and blue, in the graphics is also to more clearly reflect the unique characteristics of similar countries, which can help the audience visually and clearly observe similar points and determine what similar countries are. (3) For different interactions, it has also been discussed earlier, including the cursor floating above the points to display detailed content such as countries, Scaling can be used to adjust the overall distribution of observations or explore detailed parts, as well as to add and subtract different countries for personalized selection of countries of interest.

When implementing interactive functions, the dynamic features of the Dash application allow users to select or deselect countries, which updates the t-SNE visualization. This provides a direct manipulation approach to studying individual or group patterns and provides insights on how certain countries affect overall distribution.

In such visualization, zoom, pan, and hover functions are achieved. By zooming in and out, two similar points can be observed up close to prevent visual effects from overlapping and blurring caused by text overlap. Pan can search for adjacent points on a large image. Hovering over a country may display more information about the aggregated data of that specific country, enhancing the audience's understanding.

Analysis of countries from 2005 to 2022

Parallel coordinate plot is used to display the multivariate data, where the representation of each item in the parallel coordinate plot is in the form of polylines, where the intersection of the polylines with the axes encodes the attribute values [1]. This is considered as the one of the suitable approaches to visualize the group of countries because parallel coordinates allow users to visualize multiple numeric attributes in one single graph. Since this dataset contains 155 countries, there are going to be 155 lines within the parallel coordinates to represent different countries. However, it is very hard to visualize it by just using a single parallel coordinate plot as lines are often clustered together and it is impossible to see the individual values clearly. A choropleth map is also suitable for comparing statistics between regions immediately. In this case, this map is a suitable medium for visually representing country groupings and effectively conveying the precise values of each attribute. Using this interactive visualization, the user can select the parcoords on the parallel coordinates plot for all 12 attributes and show the specific details of the selected attributes within the country on the map.

A. *Evaluations and modifications*

An interview method is used to evaluate the effectiveness and usability of the graph. In this case, only the open-ended questions are asked. The participants are asked to interact with the interactive parallel coordinate plots with the choropleth map and a task is given to them to find a cluster, such as 5 to 10 countries that are similar.

1. What was your initial approach using the selection and screening tools in the parallel coordinates?

2. Which specific variable did you focus on during your analysis and did these characteristics help you find any clear groups of countries?

3. As you hovered over the map, how did the information presented match your expectations or needs in understanding the clusters of countries? Is there anything that was unexpected?

4. How did the color scheme of the parallel coordinates plot make it harder or easier for you to understand and make sense of the data?

5. Does the color scheme of the choropleth map enhance or detract from your understanding of the data?

6.What features or types of activities do you think could be added to the visualization to make it more useful?

Fig. 2. Interview Questions for Parallel coordinate plot and choropleth map.

Based on the results of the conducted interviews, the interactive parallel coordinates and choropleth map require some improvements. First, the choropleth map lacks information regarding the selected features, making it difficult for users to determine the value of each selected feature for each country. Therefore, a tooltip displaying the value of selected attributes is added to the map, so that the user can more intuitively understand the measures for each country. In addition, users are distracted by the color that used in the visualization. The excessive brightness of the color cause distraction on differentiating the countries and lines in the map and parallel coordinates plot respectively. In this case, a color scheme is implemented using HSL model to generate 155 different colors, while a different blue is used for the parallel coordinates plot. This ensures that the chosen hue is not too vibrant, bright, or dark for the map and use a consistent color can improve visual clarity. Therefore, uses the parallel coordinates plot for the selection of CO2 emission and the choropleth map to visualize the clusters of countries based on the selection.

B. *Final Visualization*

To understand the attributes of different countries between the year 2005 and 2022, the cumulative sum of each attribute across this specified timeframe is calculated. Also a min-max scalar is used to normalize the variable in each axes.

In the CO2 emission parallel coordinate plot, the selection of axes refers to the choices of variables representing greenhouse gas and carbon dioxide (GHG/CO2) emissions and various socioeconomic measures from 2005 to 2022. Each axis represents a distinct type of greenhouse emission, and all the axes are arranged in a parallel manner. The similar attributes, such as MARBUNK and AVBUNK, are placed adjacent to each other to help to better identify the countries with similar CO2 emission measures. The lines connecting across the axes represent different countries, where the position of each line on the axes indicates the corresponding attribute value for that country. Location is commonly used as a visual variable to leverage the position of an element to comprehend the value associated with a specific dimension [3]. The parallel coordinate plot incorporates transparency as a visual variable to enhance the viewer's attention towards the chosen data, while maintaining the information provided by the unselected data. In this case, the selected range of attributes are fully opaque, and a bright color is given to emphasize, and unselected ranges are more transparent and given a grayer color to de-emphasize.

Choropleth map is in a network arrangement, where the axes are neither linear nor circular. The color used as a visual variable in the choropleth map is HSL values, with a saturation of 60% and a lightness of 50% that ensures the color is relatively pure but not the most vibrant version of the hue and is neither too dark nor too light to represent the different countries. Therefore, a line in the parallel coordinate represents a block of a country in the choropleth map.

The interaction is a large part of this visualization. The combination of two visualizations, namely the map and parallel coordinates plots, allows for seamless integration and facilitates dynamic updates, as graphs can update and display the user selected ranges. The brushing data selection feature allows the user to select a range of points of an attribute or multiple attributes on the parallel coordinates plot by defining an attribute range on one of the axes through clicking-and-dragging the mouse. Thus, the polylines that intersect the axes are highlighted, which provide a focused view of the chosen content. At the same time, this selection updated and the corresponding countries are highlighted on choropleth map, showing the detailed values of the attributes in the tooltips.

To improve the visualization from the evaluation, a few interaction features are added. Once the user can hover on the highlighted countries on the choropleth map, a tooltip is displayed to show the name of the country, selected attributes and its value that are the user performed in the parallel coordinate plot. Users are also able to select the countries to hide on the country legend by clicking the color square displayed on the map.

C. *Interactive graph of displayed clusterred countries*

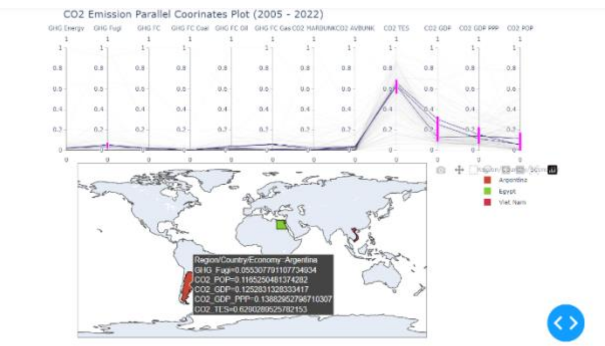

Fig. 3. The selected range of attributes in parallel coordinates plot and displayed in the choropleth map (CO2 Emission Parallel Coordinates Plot)

To illustrate the country clusters, the user has the option to initially specify a range of values for one or multiple axes on the parallel coordinate plot. The selection is done by clicking and dragging on the axes to create brushes to define the range of interest. In this case, GHG Fuji, GHG POP, CO2 GDP, CO2 GDP PPP, CO2 TES are selected with user's desired data range. Then, the parallel coordinates plot only highlights those lines that fall within the selected range to provide the immediate visual feedback to the subset of the data. Then, the choropleth map is updated to include 3 distinct countries, including Argentina, Egypt, and Vietnam, each of which has been allocated a unique color scheme in accordance with the HSL color model. When the user navigates to a specific country by hovering, the tooltip will display the precise value of each attribute. This visual clustering of countries facilitates the quick detection of geographical patterns and groupings associated with carbon dioxide (CO2) emissions.

II. Task 2

A. *Evaluation*

A questionnaire is designed to evaluate the user satisfaction for using the heatmaps. The questionnaire contains both closed-ended and open-ended questions to gather user feedback on the understanding, usability, interactivity and overall user experience (Figure 1).

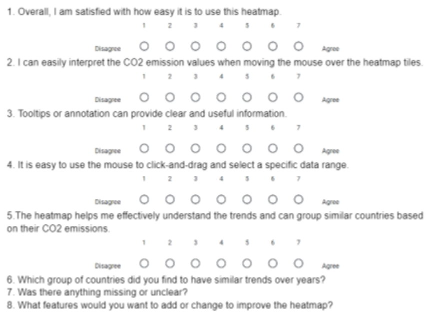
1. Overall, I am satisfied with how easy it is to use this heatmap.
Disagree 1 2 3 4 5 6 7 Agree
2. I can easily interpret the CO2 emission values when moving the mouse over the heatmap tiles.
Disagree 1 2 3 4 5 6 7 Agree
3. Tooltips or annotation can provide clear and useful information.
Disagree 1 2 3 4 5 6 7 Agree
4. It is easy to use the mouse to click-and-drag and select a specific data range.
Disagree 1 2 3 4 5 6 7 Agree
5. The heatmap helps me effectively understand the trends and can group similar countries based on their CO2 emissions.
Disagree 1 2 3 4 5 6 7 Agree
6. Which group of countries did you find to have similar trends over years?
7. Was there anything missing or unclear?
8. What features would you want to add or change to improve the heatmap?

Figure 2

Apriori of this evaluation process, the interactive features implemented for the heatmap allowed users to click, drag, and select data sections on the heatmaps. Moreover a dropdown window was leveraged to facilitate switching between different socio-economic measures. Figure 1A in the appendix shows the preliminary interactive visualization produced. As for the evaluation process, we designed an assessment where users first engaged in a 10-minute exploratory session of the heatmap, followed by taking a survey. Several features were added based on user feedback. To address difficulties in the accurate data selection, a dropdown list for selecting countries is added. Upon the selection of a country the heat map displays the chosen country along with the ten countries that precede and follow it in alphabetical order. In addition, the color scheme was revised to improve the visual clarity of the heatmap. The Thermal color scale was found not to be suitable for this particular scenario, as the intensity of certain hues and wide color range caused confusion among the users. Indeed we found out that the thermal color scale often resulted in users reporting some countries to have similar trends, when however this was not accurate. The latter became clear as, a-posteriori, we tried to validate their answers leveraging the line chart. The mentioned countries appeared indeed to actually display fairly dissimilar trends.

Lastly, from the feedback received we also realized how the possibility to only select one country caused some difficulties. Indeed, users were unable to specifically confront two or more countries of interest when not fairly close in the alphabetical order . Hence we added an additional dropdown list to choose multiple countries to be confronted with the "target" one.

B. Visualisations

The heatmap (Figure 2) was produced by applying min-max scaler to normalize the socio-economic measures (i.e. CO2 emission related to GDP, CO2 emission related to Purchasing Power Parity, and CO2 emission related to Population). This normalization ensures that the socio-economic measures can be effectively compared across all 155 countries and over the time period between 1971 and 2022. The x-axis shows the year, while the y-axis shows the name of the country. Each tile on the heatmap represents a CO2 emission value for a particular country in a given year. Color intensity, among all retinal variables, is a crucial one for our visualization as it allows users to visually perceive the annual trends of CO2 emissions across different countries. A sequential color scale called Reds, has been selected to represent the socioeconomic measures (i.e. CO2/GDP, CO2/PPP, CO2/GDP PPP). The scale ranges from low to high values, with darker shades of red representing higher levels of CO2 emission and lighter shades representing lower levels of CO2 emission. Instead of using complex color scales like Turbo or Thermal, this scale utilizes shades of a single color, which can be very effective in distinguishing similar countries, as similar color shades are assigned to countries with comparable characteristics. As a result, the ability to find comparable CO2 emission is considerably improved.

C. Interactions

Hovering the mouse over tiles on the heatmap can display additional information in tooltips, including year, country and CO2 emission values. In addition, users have the ability to choose a specific data range of interest by utilizing the click-and-drag of the mouse. Such action results in a zoom-in over a specific section of the heatmap which consequently allows for a more fine grained analysis.

The interface (Figure 2) features three dropdown lists: "Select an emission index", "Select the 'target' country","Select other countries to compare".

By default, the heatmap displays trends of all 155 countries in comparison to Italy using the CO2/GDP emission index. However, users can dynamically update the heatmap by changing the target country and its comparison countries using the first two dropdown lists.

The last dropdown list also allows you to select "all" in which case all countries will be used in comparison to the target one. The analysis can then be refined either using the click-and-drag action or, for more targeted results, listing the countries to be compared with the target in the last dropdown list.

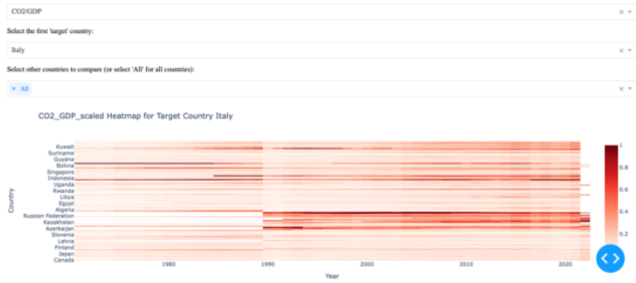

Figure 3 (figure 2A appendix for a zoomed version)

D. Interpretation of Results

Figure 2 displays the heat map obtained once a socio-economic measure, specifically CO2 GDP in this case, has been selected but no specific country is (hence the default ones are used). The high level visualization produced allows the user to get a first rough understanding of how to spot similar and dissimilar trends. Indeed in Figure 1 it can be seen that countries plotted nearby Libya seem to have similar trends as the colors stay consistent along the X axis. With the click and drag functionality we can zoom in into that area which shows the details for those trends (Figure 3A in the appendix). Following the same method it can be noticed how the countries around Indonesia have very dissimilar trends among each other.

Utilizing the two dedicated dropdown lists a user has the possibility to further refine the investigation. For instance as we focus on an analysis of Australia for the index CO2 GDP from figure 2 it can be seen that Australia exhibits a very similar trend to that of New Zealand. Indeed the light gradation of red that stays consistent along the X axis for both countries suggests that CO2 GDP, throughout the time, has stayed low for both countries. On the other hand China and India both appear to have considerably different trends to that of Australia. With significantly higher values and a clearly decreasing pattern for China. Such findings are cross validated (Figure 3) confronting the finding of the heatmap with the line chart. Indeed, the line chart allows for a more detailed comparison of the timeseries. Such a procedure allowed us to establish the effectiveness of the heat map as we found that conclusions drawn from the latter were also backed up by the findings of the line chart. Moreover it should be noticed how adopting a fitting distance metric, like it has been done for the line chart, allows to find grouping of clusters with considerably less effort compared to a method solely based on a visual exploration, as is the case for the heatmap presented.

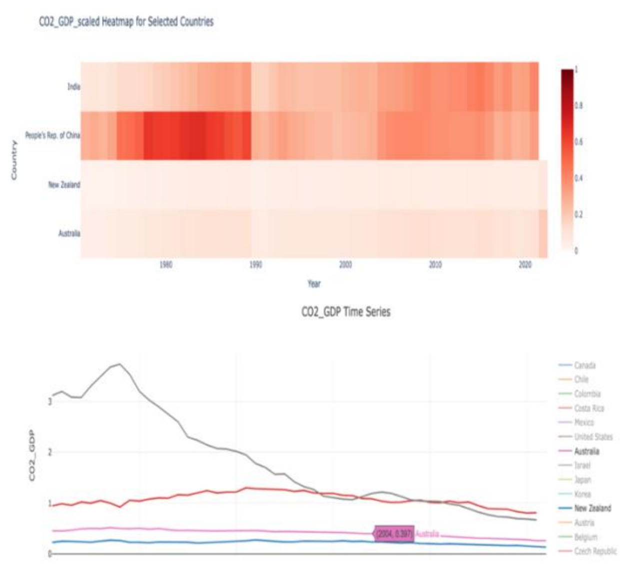

Figure 4

E. Line Graph

As a preliminary measure, a distance matrix is computed using the euclidean distance. Such a metric is computed among all the possible pairs of countries. The euclidean distance of 2 time series is computed point wise. Such metric is later on discussed as was used, down the line, to add some interactive functionalities.

A line chart is used in order to highlight different trends in the socio-economic indexes of the different countries. The X axis takes on the time (i.e. years) whilst on the Y axis is depicted one of the socio economic indexes available from the dataset. The index, hence the time series to be depicted, can be dynamically chosen using a drop down window. Different countries are depicted with different colors to make the lines easier to be distinguished. Moreover, a “hover-over” functionality is added making it possible to visualize the name of the country (as well as the precise value of the index) to which the line belongs. A further dynamic functionality implemented allows to dynamically remove a given country from the visualization by double clicking on its name in the legend. Or to only show trends of selected countries. This was considered useful as at times countries with some extreme values cause the range of values adopted by the Y axis to stretch. This, in turn, results in many overlappings and an ineffective visualization.

As a first attempt we plotted the time series of all the available countries at once. Such an approach however resulted in numerous overlapping to the point that the visualization was not useful to convey any sort of information. The successive efforts have focused on determining a meaningful approach that could be applied in order to reduce the number of time series plotted at once whilst retaining the capability of spotting similar trends among countries (Figure 4A appendix). As previously mentioned, the most significant approach was found to be the computation of a distance matrix. Leveraging such information, the final visualization obtained introduces two additional functionalities. Firstly, it is possible to select the country upon which to find countries with similar trends. Next, it is required to select the N number of countries that are going to be depicted in the plot. Such countries are chosen based on their Euclidean distance from the “target country”, in order to pick the most similar ones. Following the same approach we also made it possible to plot the N most dissimilar countries to the target one. (Figure 5A appendix)

III. Task 3

We selected Australia as the target country for our analysis. In order to determine the years with the lowest and highest emissions, we standardized and summed all the attributes of each year for comparison. The year with the lowest sum was the year with the lowest emissions, while the year with the highest sum was the year with the highest emissions. Therefore, we found that the year with the lowest emissions was 1972, and the year with the highest emissions was 2010. The reason for using standardized summation is because there is a significant difference between data with different attributes. Some data have a minimum data size of 1000+, while others only have a minimum data size of 100+. If the data is directly added without any data processing, the influence of some attribute data will be ignored, resulting in subjective errors. Therefore, using standardized summation is more fair.

A. Interactive Visualization

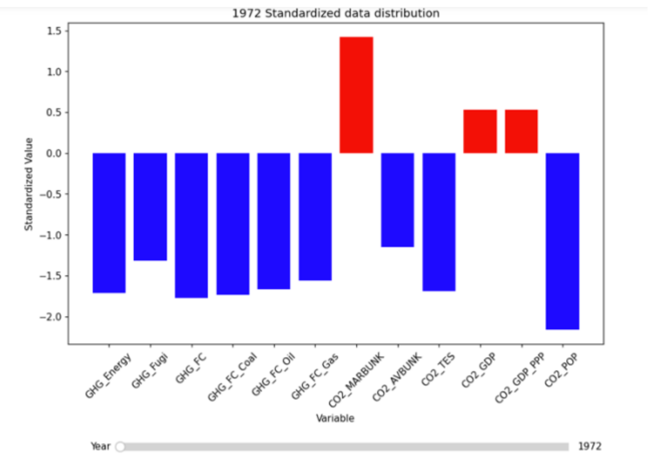

fig.1

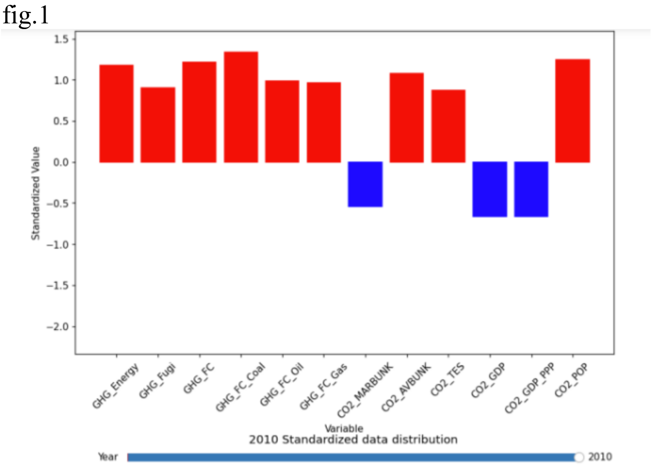

fig.2

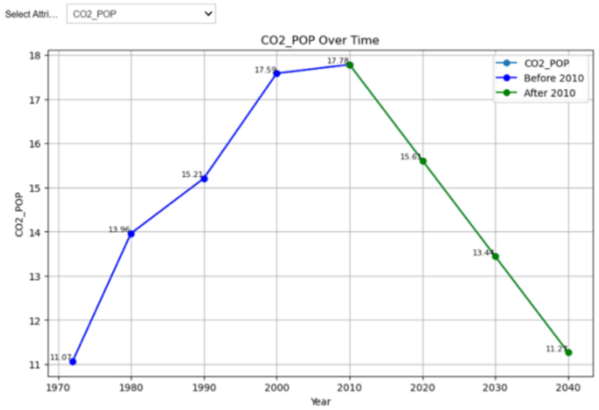

fig.3 Display the corresponding line chart by selecting different attributes (divided into actual data before 2010 and planned data after 2010)

B. Evaluation

1. 1. For CO2_POP, which time periods have significant differences?
1) 1972-1990 2) 1980-2000 3) 1990-2010
2. Can you intuitively observe the changed value?
1) Yes 2) No
3. If you want to further show the change in emissions from the lowest to the highest, which way is better?
1) Drag and drop year interaction 2) Select year interaction
4. Can you intuitively observe the predicted value?
1) Yes 2) No
5. If you want to show the prediction line of 12 attributes, which way is better?
1) Select the property display line figure 2) Select the fold among the 12 lines
6. Can you make a plan for each attribute from the forecast trend graph?

The improvement of fig.1&fig.2 based on feedback from the survey questionnaire includes: (1) Add different colors to the graphics. The initial default color was blue. Although the difference in the positive and negative directions of the graph was significant, in order to more conveniently meet the influence of visual variables, red was applied to positive values and blue was applied to negative values. (2) Increase drag and drop interaction. In order to more conveniently and intuitively display the changes from the lowest emission year to the highest emission year, drag and drop interaction is implemented. By dragging the year with the cursor, not only can it pause in any year from 1972 to 2010, but also changes and trends can be observed through quick drag and drop.

After questionnaire evaluation, we made some key improvements to the visual chart - fig.3. First, we introduced color to clearly distinguish between actual data before 2010 and projected data after 2010, and added a corresponding label description for each data point. This improvement enables viewers to more intuitively understand the differences between planned and actual data, helping them to assess future trends more fully.

Second, more critical details are provided by the way data points are added for each data point. This gives the audience a better understanding of the magnitude and trends of change. This also helps support more detailed data analysis and decision making.

The interactive feature of the drop-down selection enables the viewer to actively explore the predicted values of different properties. This interactive tool allows users to customize the data they want to explore based on the properties they are interested in.

C. Visualization

1. Fig.1&Fig.2

As shown in the figure, this is the density distribution map of the years with the lowest and highest emissions after standardization, where blue represents negative values and red represents positive values. In order to further visualize the process of change, interactive dragging was used to drag and drop the years from 1972 to 2010. However, when visually displaying two images, it can be seen that the attributes that were positive in 1972 became negative in 2010. The horizontal axis represents the names of different attributes, while the vertical axis represents the standardized values.

Regarding the final visualization effect, the content of the horizontal and vertical coordinates is to display the values of different attributes, as well as to view the standardized values of the 12 attributes. Then, through the strong contrast between the tones of red and blue, it helps the viewer to more clearly identify the negative and positive values. The final drag and drop interaction button can stay in different years from 1972 to 2010 to observe the situation based on the personalized thoughts of the observer, and can also quickly drag and drop to view the trend of changes.

2. Fig.3

The year is the X-axis. The Y-axis is the metric associated with that attribute. Each property may involve different units of measure, so the Y-axis measurement depends on the properties of the property. For example, if you wanted to look at per capita carbon emissions over years, then the X-axis would be years and the Y-axis would be "CO2_POP" (in tons or other appropriate units). This will help understand trends in Australia's per capita carbon emissions for planning purposes.

We mainly use line charts to make plans. First, line charts can visualize trends, making trends and patterns visible at a glance. This helps decision-makers make sense of data and identify events more easily. For example, if there is any significant growth or decline in a broken line, this may indicate that this period needs special attention. Second, line charts are particularly useful for time series data, such as annual data. They can be used to track changes over time, helping decision makers better understand historical data and make predictions about future trends. Not only that, it also helps with goal setting. Based on the trends and analysis of line charts, decision makers can better set goals. This helps with resource planning to achieve specific outcomes.

The chart uses a pull menu widget to select the properties to draw. Its options are different attributes in the data. When the user selects different properties in the drop-down menu, the interactivity automatically draws the corresponding line chart. The benefits of interactivity include dynamic visualization, data communication, and easy comparison. Users can select different properties according to their needs to dynamically view the changing trend of properties, without manually changing parameters. Comparing the relationships between different attributes allows us to better explore the data to understand the plans we are making. Not only that, but by plotting two broken lines before and after 2010, users can also more easily compare trends over different periods, which helps analyze the impact of the measures developed on the data.

D. Emission reduction plan

Reducing greenhouse gas emissions and using more renewable energy is a key action needed to mitigate the effects of climate change. In order to protect the environment, Australia has developed a long-term emission reduction plan. Starting with the change of "CO2_POP" variable allows everyone - from individuals to households and businesses - to work to control the rate of global warming and climate change.[1] From the line chart, we can see that the rate of per capita carbon emissions has decreased significantly from 1980 to 1990 and from 2000 to 2010. Assuming we reduce per capita carbon emissions as Australia did in 1980-1990, we expect to return per capita carbon emissions to 1972 levels within 30 years. Since 2015, the Australian government has launched a number of new energy development strategies, which has made Australia quickly become a global leader in new energy, and its low-carbon development has entered the fast track. Since 2017, Australia's per capita carbon emissions have continued to decrease, reducing the level of carbon emissions per capita to 15.37 tonnes in 2020 (about the same as 1985). [2] So the line chart predicts that our plan is feasible. We can reduce per capita carbon emissions in the following ways: Energy efficiency improvements. In the 1980s, some homes and businesses had begun to focus on energy, taking steps to reduce the use of electricity and natural gas. The same is true in the 21st century, "Changing the way we power our homes and industry." Get more energy from the sun and wind into our power system and try to stop using fossil fuels. "says Dr Simon Bradshaw(2023) . On August 25, 2009, the Australian Parliament passed the Climate Change Act to set the standard for the use of renewable energy, and determined that renewable energy will account for 20% of electricity demand by 2020.[3] Public transport and development. Beginning in 1980, some Australian cities began to develop more integrated public transport systems, encouraging people to use public transport in order to reduce personal car use. Carpooling and shared transportation are also becoming popular. [4]Today, bike sharing is common on Australian streets. This undoubtedly reduces CO2 emissions from the aspect of life. Recycling and garbage sorting. In the 1980s, some local governments began to promote recycling and waste sorting to reduce landfill and incineration and reduce the environmental burden. Today, Australia is taking the issue more seriously. All Australian states and territories have recycling programs in place to encourage residents to recycle recyclable materials such as paper, plastic, glass and metal. [5]Residents and businesses can also be required to implement garbage sorting programs. This helps to reduce the amount of waste going to landfill and incineration and increase the recovery rate of resources. Environmental education and awareness. In 2001, Australia established the National Environmental Education Network, which consists of federal, state and territory environmental and education departments. On the basis of the implementation strategy put forward in "For a Sustainable Future: National Action Plan for Environmental Education", a set of standardized school environmental education action plan - Sustainable Schools Plan is proposed.[6] Today, thousands of schools in Australia participate in the Sustainable Schools Program. According to statistics, sustainable schools reduce waste collection by 80%, water consumption by 60%, and energy consumption by 20%. Therefore, we should continue to organize such education and publicity activities. Raise public awareness of carbon reduction and encourage them to take action.

References

[1] H. Zhou, X. Yuan, H. Qu, W. Cui, and B. Chen, "Visual Clustering in Parallel Coordinates," *Computer Graphics Forum*, vol. 27, no. 3, pp. 1047–1054, May 2008, doi: https://doi.org/10.1111/j.1467-8659.2008.01241.x.

[2] J. Blaas, C. P. Botha, and F. H. Post, "Extensions of Parallel Coordinates for Interactive Exploration of Large Multi-Timepoint Data Sets," *IEEE Transactions on Visualization and Computer Graphics*, vol. 14, no. 6, pp. 1436–1451, Nov. 2008, doi: https://doi.org/10.1109/tvcg.2008.131.

[3] Ł. Halik, "The analysis of visual variables for use in the cartographic design of point symbols for mobile Augmented Reality applications," *Geodesy and Cartography*, vol. 61, no. 1, pp. 19–30, Nov. 2012, doi: https://doi.org/10.2478/v10277-012-0019-4.

[4] Vaclav Smil (2017). Energy Transitions: Global and National Perspectives .

[5] Hannah Ritchie and Pablo Rosado (2017) - "Fossil Fuels" Published online at OurWorldInData.org. Retrieved from: https://ourworldindata.org/fossil-fuels [Online Resource]

[6] Cassidy, E. (2019). Which Countries Use The Most Fossil Fuels? Retrieved from: https://blog.resourcewatch.org/2019/05/02/which-countries-use-the-most-fossil-fuels/

[7] Gençsü, I., & Hino, M. (2015). Raising Ambition to Reduce International Aviation and Maritime Emissions. Retrieved from http://newclimateeconomy.report/workingpapers/wp-content/uploads/sites/5/2016/04/NCE-Aviation-Maritime_final.pdf

[8] Schinas, O., & Bergmann, N. (2021). Emissions trading in the aviation and maritime sector: Findings from a revised taxonomy. Title of the Journal, Volume(Issue), Page numbers. doi: 10.1016/j.clscn.2021.100003

[9] Balsamo, M., Montagnaro, F., & Anthony, E. J. (2022). Socio-economic parameters affect CO2 emissions and energy consumption – An analysis over the United Nations Countries. Title of the Journal, Volume(Issue), Page numbers. doi:10.1016/j.cogsc.2022.100740

[10] Phil.T(2023, October 24), Australia's net-zero emissions plan for 2050, *SBS News*.

[11] Emily.H(2016, August), Climate change—reducing Australia's emissions, Parliamentary Library Briefing Book - 45th Parliament, Parliament of Australia.

[12] Wuwei.Z(2021, November 1st), What climate policies have countries historically introduced?, The Paper News.

[13] Australian Government(2022). Australia's Long-Term Emissions Reduction Plan.

[14] Dongbao.D(2010, May), "The Review and Recent Developments of Australia's Climate Change Policy." World Environment .

[15] Yong.F (2020, April 2nd) ,This is how Australia cultivates its citizens' ecological awareness, *Guangming Daily*.

References

[16] Ratanamahatana, C.A., & Keogh, E.J. (2004). Making Time-Series Classification More Accurate Using Learned Constraints. *SDM*.

IEEE conference templates contain guidance text for composing

IV. Appendix

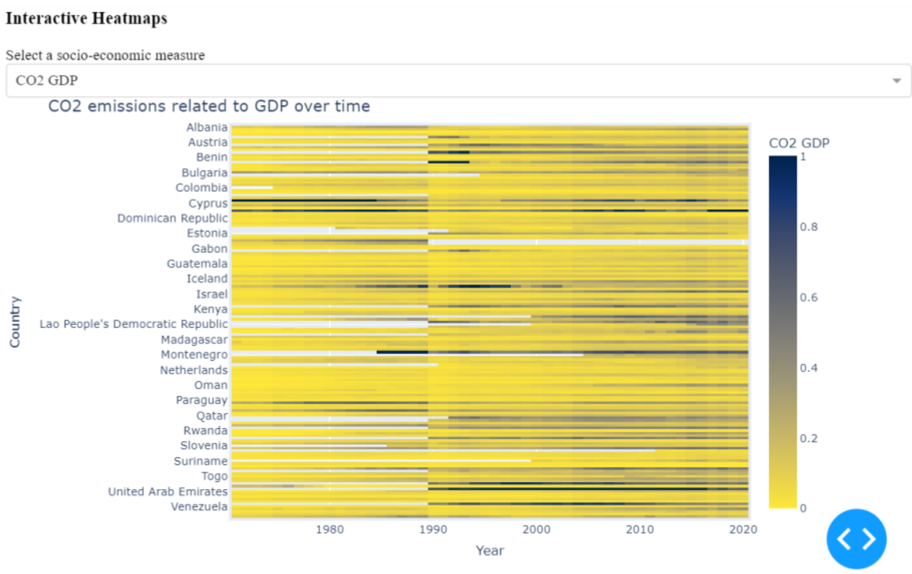

Figure 1A

Figure 2A

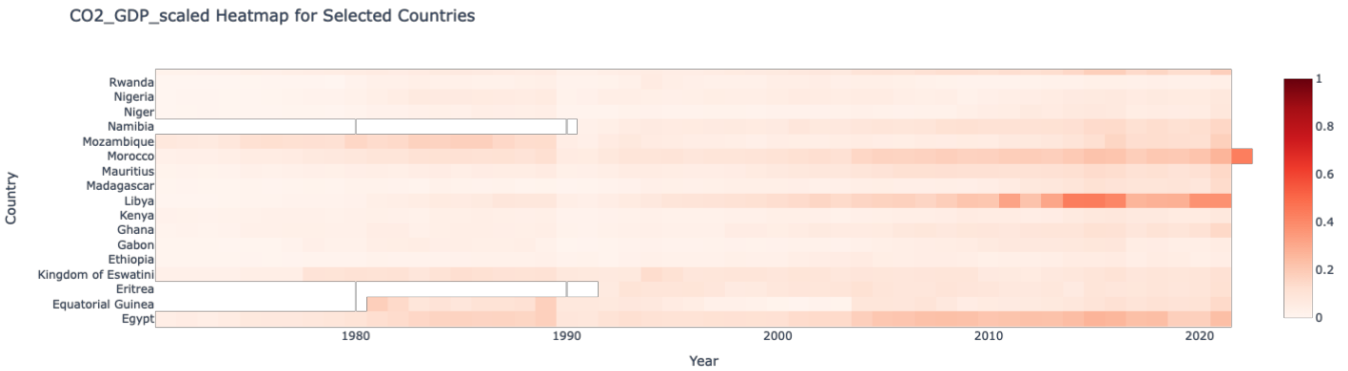

Figure 3A

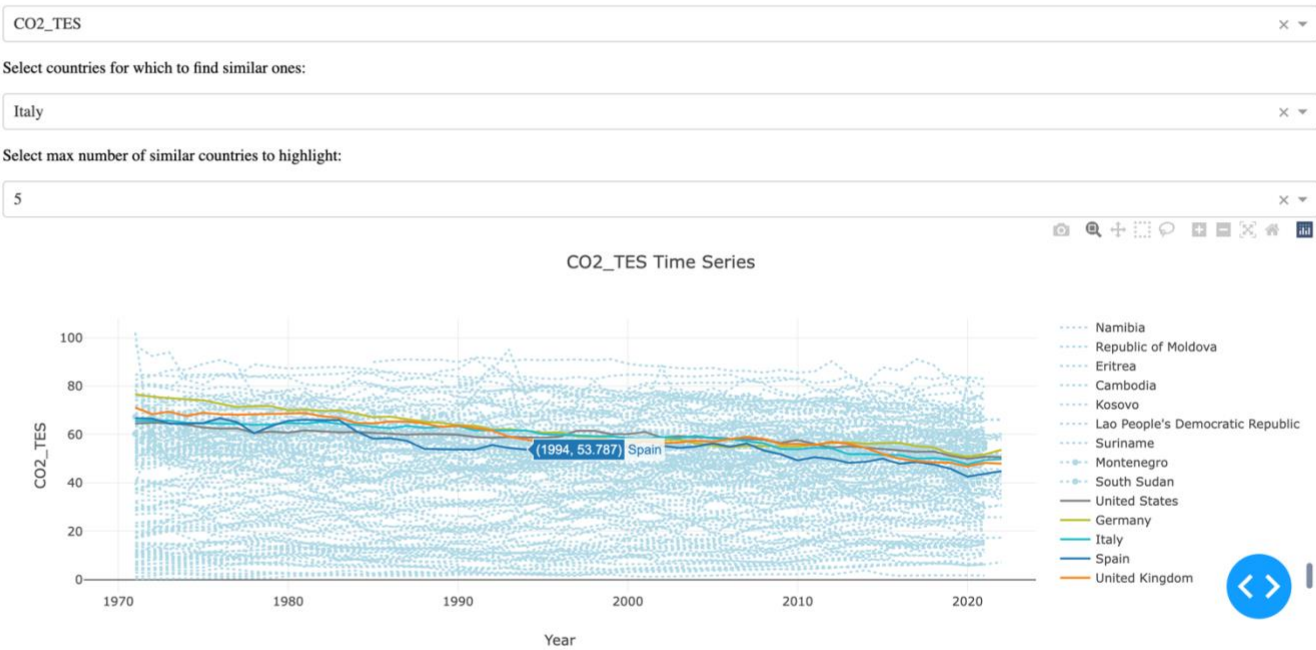

Figure 5A

CO2_TES

Select countries for which to find similar ones:

Italy

Explore trends for:

Most similar countries

Select max number of similar countries to highlight:

5

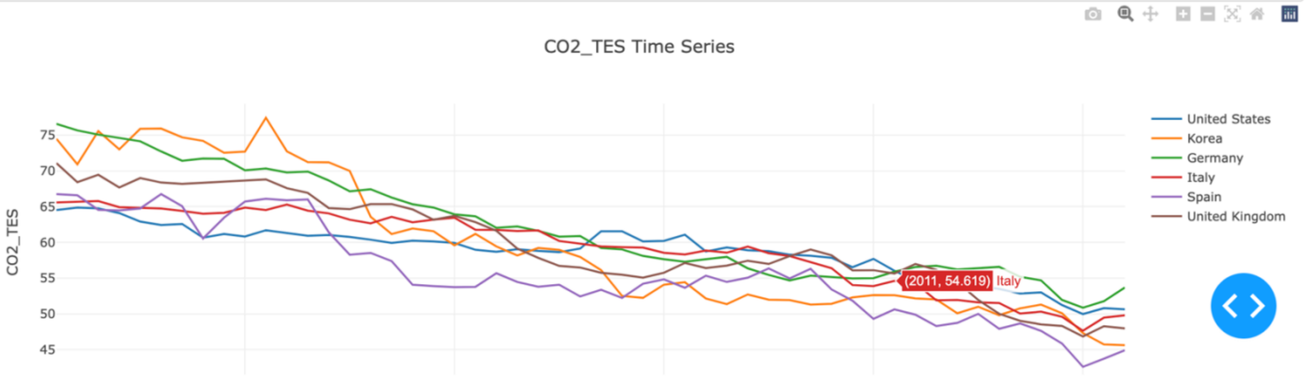

Figure 6A