Team Members: Weiyu Song Xinchi Yao

This dataset contains Quality of Life indices for various countries around the globe, extracted from the Numbeo website.

The data provides valuable metrics for comparing countries based on several aspects of living standards,which can assist in decisions such as choosing a place to live or analyzing global trends in quality of life.

OBS: The code to generate this dataset is presented on: https://www.kaggle.com/code/marcelobatalhah/web-scrapping-quality-of-life-index

The data from kaggle website, the website is: https://www.kaggle.com/datasets/marcelobatalhah/quality-of-life-index-by-country

For the first visualization, our group chose to analyze the "Quality of Life Index". We displayed the "Quality of Life Index" of each country in the world map,

In this way, the relationship between the geographical location of each country and the "Quality of Life Index" can be intuitively seen.

This snippetreads data from a CSV file, extracts country and QOL information, and uses the pyecharts library to generate an interactive map of the world.

The colors of the map are colored according to the values of the quality of life index, and the color range is defined through the visual mapping option.

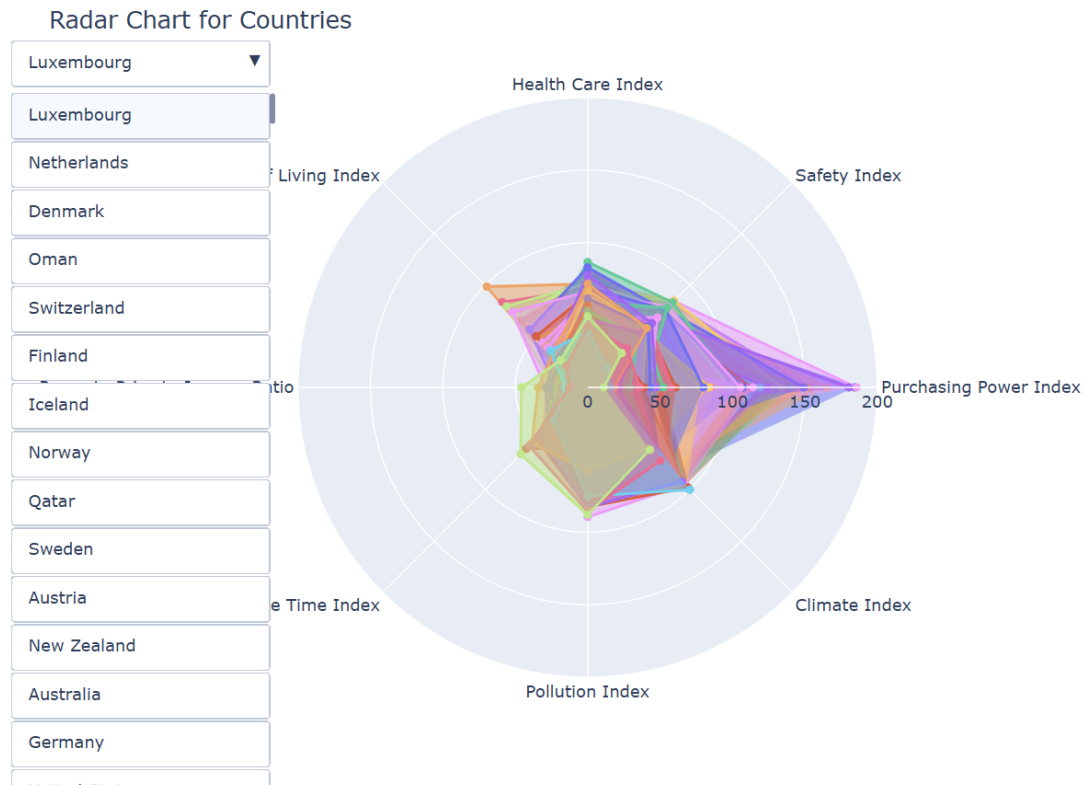Finally, the map is rendered and saved as an HTML file that can be viewed in a browser.

The libraries used in this process are pyecharts and pandas
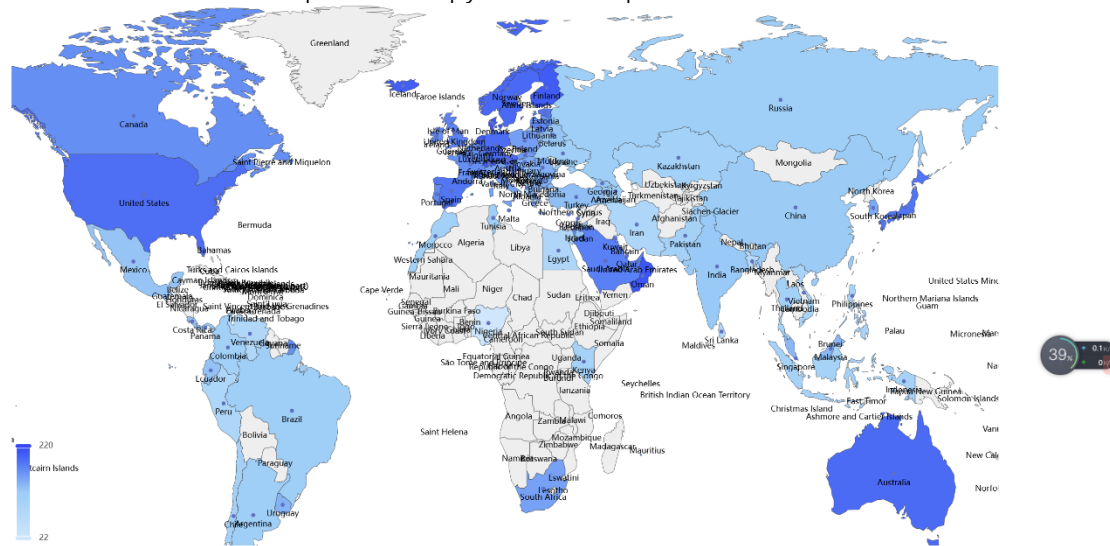
For the second visualization operation, we selected the radar chart to analyze the other 8 attributes except "Quality of Life Index", and selected the country to view by the drop-down selection method

After some simple processing of the data, we add an interactive menu that allows users to select different countries to display their corresponding radar chart through the drop-down menu.

Finally, we use the fig.write_html("radar_chart.html") method to save the chart as an HTML file for later viewing.

Radar Chart for Countries

| Luxembourg | ▼ |

Luxembourg
Netherlands
Denmark
Oman
Switzerland
Finland
Iceland
Norway
Qatar
Sweden
Austria
New Zealand
Australia
Germany

Health Care Index
Safety Index
Living Index
Purchasing Power Index
atio
Time Index
Climate Index
Pollution Index

The third visualization program is our implementation to generate a heat map, which is used to show the correlation between different quality of life indicators.

The code extracts the quality-of-life columns, computes the correlation matrix, uses Matplotlib and Seaborn to create a heatmap, displays the magnitude and direction of the correlation (red: positive correlation, blue: negative correlation, white: no correlation),and saves and displays the plot. Use Pandas for data processing and Matplotlib and Seaborn for plotting.

Heatmap of Quality of Life Indices
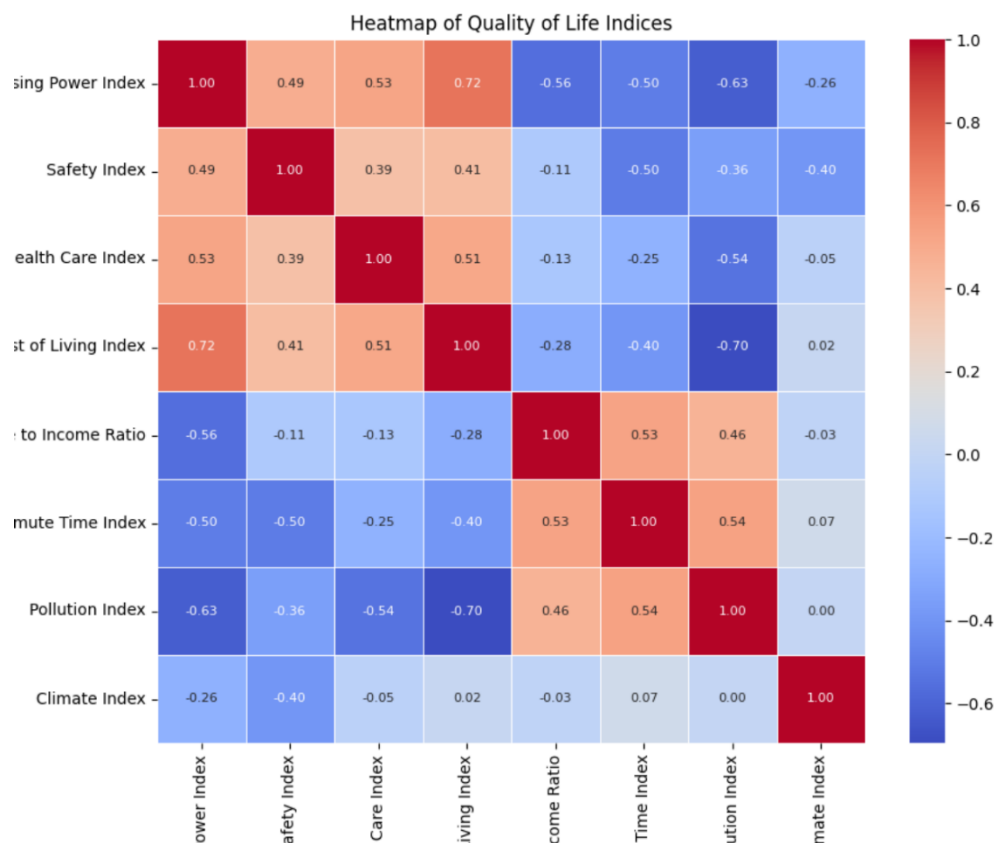
For the fourth data visualization program, we chose "scatter plot matrix".
Heat plots with Matplotlib and Seaborn, showing the magnitude and direction of correlation (red: positive correlation, blue: negative correlation, white: no correlation), and saving and displaying the images.
Use Pandas for data processing and Matplotlib and Seaborn for plotting.
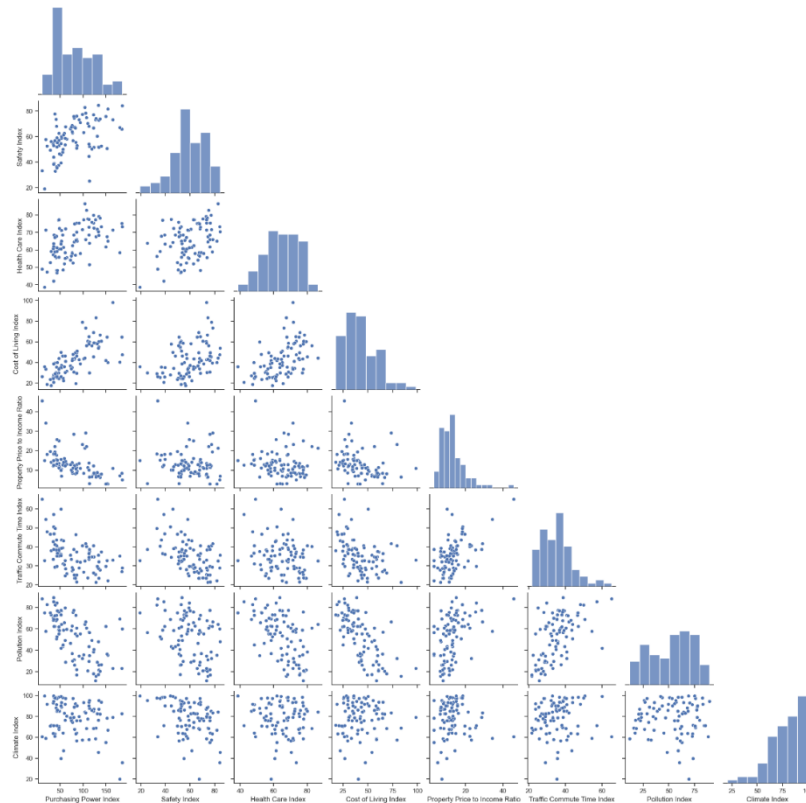
Below is a description of our dataset.

Overall quality of life ranking

There is a significant gap between Luxembourg, which has the highest QOL (220.1), and Nigeria, which has the lowest QOL (21.5).

European countries (e.g., Luxembourg, the Netherlands, Denmark, Switzerland) generally rank high,

indicating that these countries perform well in terms of overall quality of life.

The heatmap generated from this data can help us find the following interesting conclusions:

Positive correlation:

There may be a strong positive correlation between Purchasing Power Index and Quality of Life Index. The higher the purchasing power, the higher the quality of life usually is.

Safety Index and Quality of Life Index may also show a positive correlation, indicating the importance of a safe environment for quality of life.

Health Care Index and Quality of Life Index may have a positive correlation, reflecting the direct impact of medical conditions on quality of life.

Negative correlation:

There may be a negative correlation between Pollution Index and Quality of Life Index, the more severe the pollution, the lower the quality of life.

The Traffic Commute Time Index and the Quality of Life Index may show a negative correlation, with longer commute times likely to be associated with lower quality of life.

2. Relationships between specific metrics

There may be a positive correlation between Cost of Living Index and Property Price to Income Ratio, indicating that countries with a high cost of living are also likely to have a high price to income ratio.

Safety Index and Property Price to Income Ratio may show a negative correlation, and countries with poor safety environment may have high house price to income ratio, reflecting the trade-off between safety and housing affordability.

3. Anomalies or interesting patterns

The correlation between the Climate Index and the Quality of Life Index may be weak, suggesting that climate may have a less significant effect on quality of life than other factors.

The correlation between Purchasing Power Index and Cost of Living Index may be weak, reflecting that even if purchasing power is high, the cost of living is not necessarily low and may be affected by other economic factors.

4. Comparison between countries

Through the heatmap, it can be observed that some countries perform prominently on

specific indicators. For example:

With high scores on the Purchasing Power Index and Quality of Life Index, Luxembourg is likely to be a country with a high quality of life.

Qatar has a high score on the Purchasing Power Index, but a low score on the Pollution Index, which may reflect the contradiction between high purchasing power and environmental pollution.

Nigeria, which has the lowest Quality of Life Index score, may perform poorly on multiple indicators and deserves further analysis of the specific reasons.

Here are a few interesting observations from the pairplot:

1. Correlation between variables:

Positive correlation: There may be a positive relationship between some variables. For example:

There may be a certain positive correlation between "Purchasing Power Index" and "Cost of Living Index" There may also be a positive relationship between "Safety Index" and "Health Care Index".

Negative correlation: There may be a negative correlation between some variables. For example:

There may be a negative correlation between "Pollution Index" and "Climate Index" (the distribution of points in the scatter plot shows a downward trend).

There may be a negative correlation between "Property Price to Income Ratio" and "Purchasing Power Index".

2. Data distribution characteristics:

The histogram on the diagonal shows the distribution of each variable:

The distributions of "Purchasing Power Index" and "Safety Index" are relatively concentrated and may have a unimodal distribution.

The distributions of "Traffic Commute Time Index" and "Pollution Index" may be scattered and have large fluctuations.

The distribution of "Climate Index" may be skewed to one side.

3. Outliers and outliers:

In some scatterplots, there may be outliers (points that are far away from other points). For example:

In the scatter plot of "Property Price to Income Ratio" and other variables, there may be some outliers,

indicating that the values of some data points are very different from others.

Outliers may also exist in the scatter plots of "Traffic Commute Time Index" and "Pollution Index".

4. Nonlinear relationship between variables:

The relationship between some variables may not be linear. For example:

A scatter plot of "Cost of Living Index" and "Health Care Index" might show some kind of nonlinear pattern (such as a curved-line relationship).

The relationship between "Pollution Index" and "Climate Index" may also have nonlinear characteristics.