**2024 World Happiness Index Report**

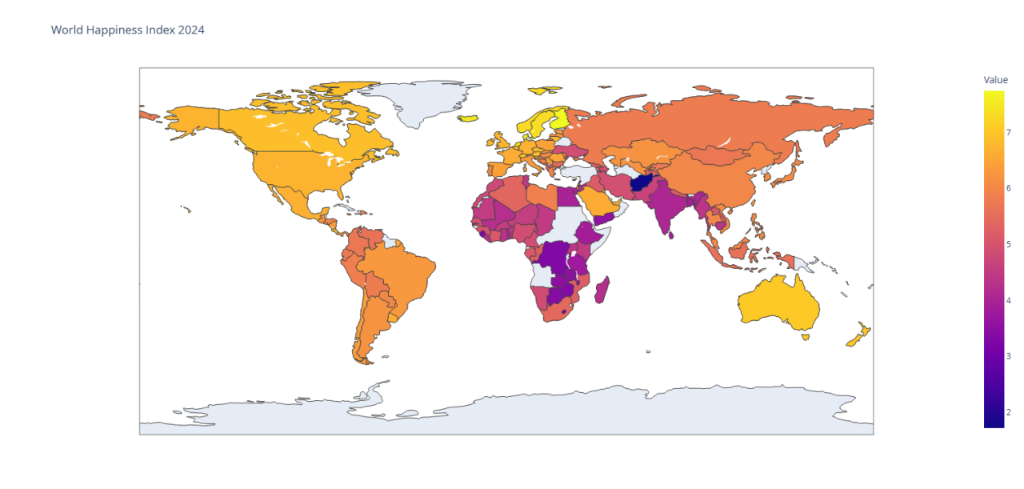
**1.Dataset Description**

This report is based on the 2024 World Happiness Index dataset, which includes the happiness rankings and related indicators of 143 countries worldwide. Key indicators include happiness score, economy (GDP per capita), social support, healthy life expectancy, freedom to make life choices, generosity, and perceptions of corruption. These data provide us with a comprehensive perspective to deeply understand the happiness levels of different countries. By analyzing these indicators, we can identify the key factors influencing national happiness levels and provide valuable references for policymakers. Each country in the dataset has detailed happiness scores and corresponding indicator values, enabling us to conduct multi-dimensional analysis and comparisons.

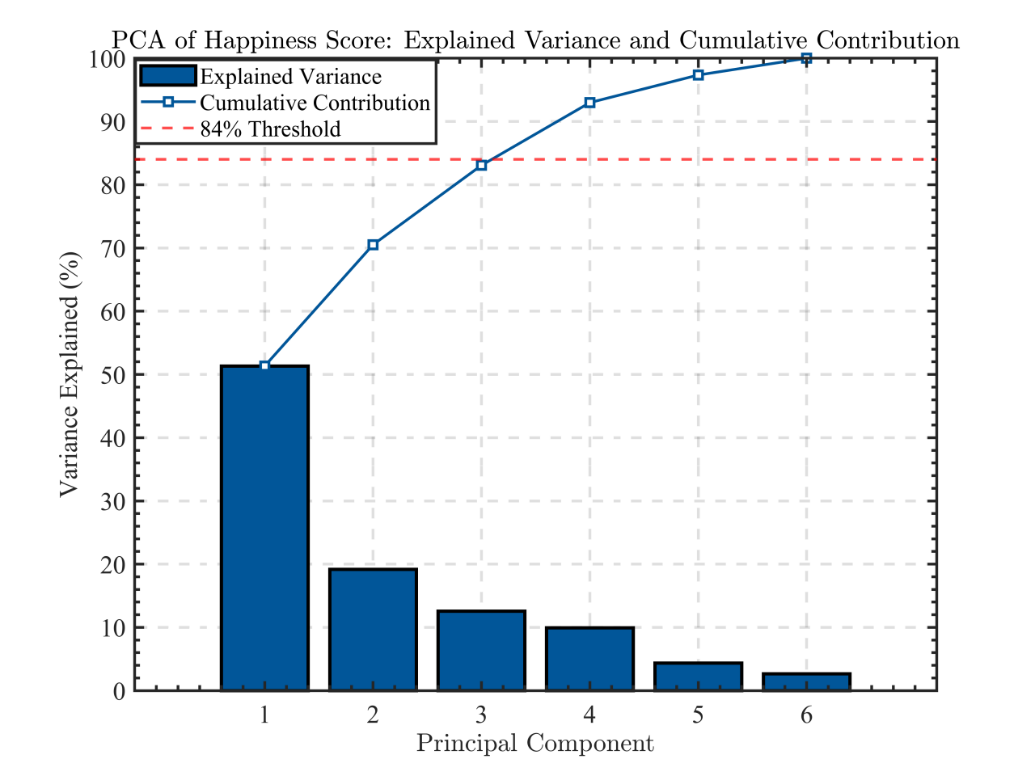
**2.Visualization Presentation**

According to the visualization analysis requirements, we used five different visualization methods to present the data, allowing for a more comprehensive understanding of the global happiness index distribution and its underlying driving factors:

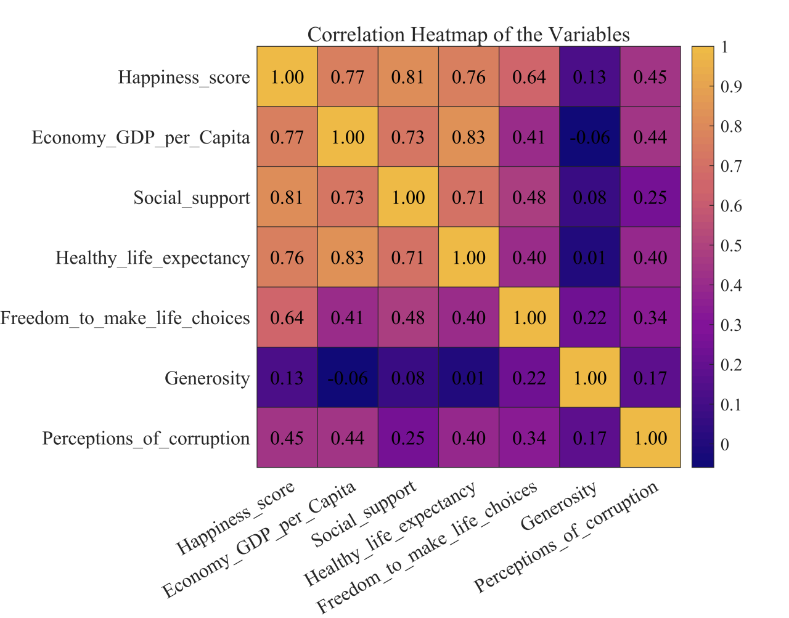
1. **World Map Heatmap:** Displays the happiness score distribution of 143 countries. The intensity of colors represents the level of happiness scores, providing an intuitive view of global happiness level differences. Countries with higher happiness scores are mainly concentrated in Northern and Western Europe, while those with lower scores are mostly in Africa and South Asia.



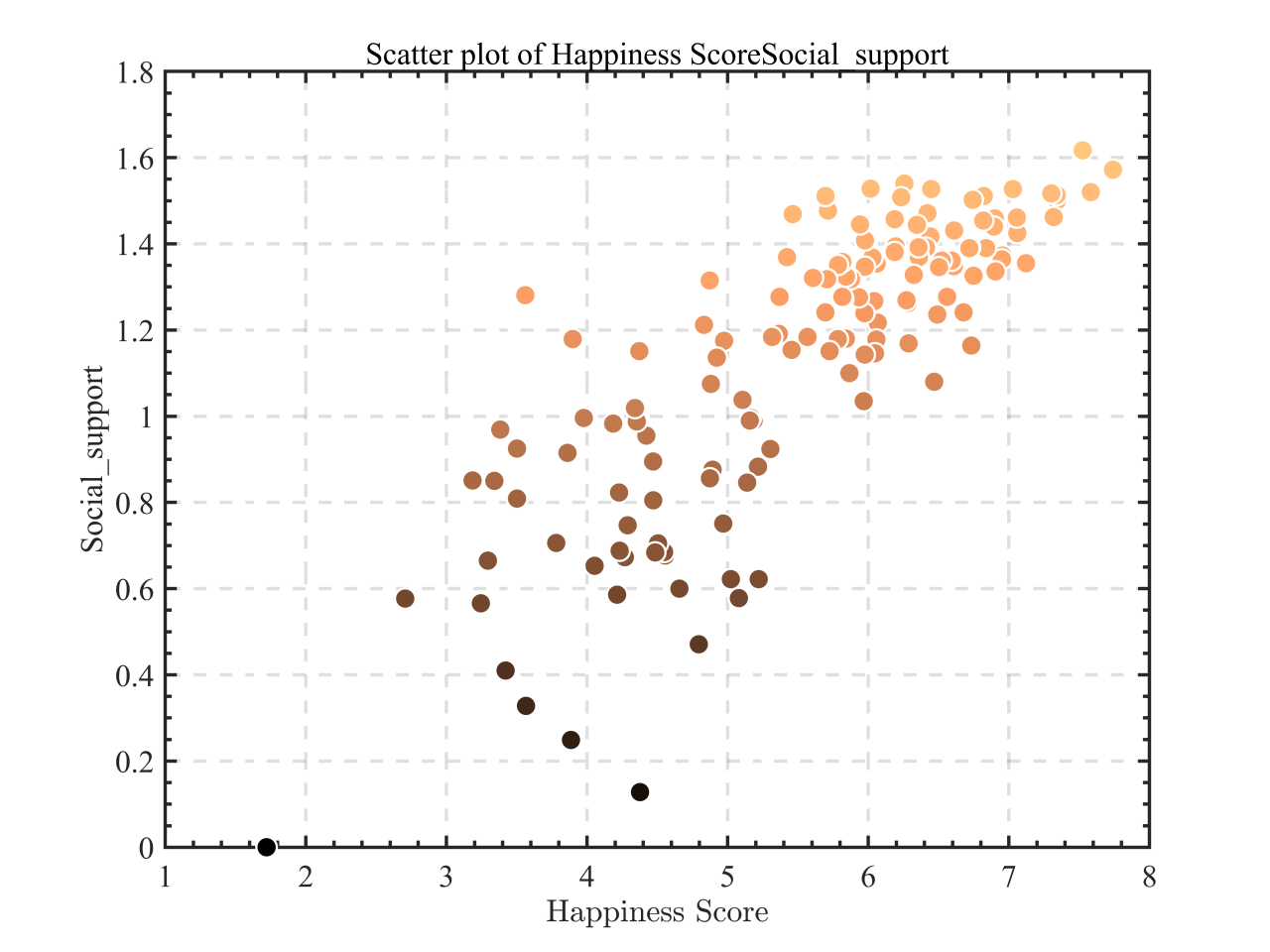
1. **Principal Component Analysis (PCA):** Conducts PCA on happiness scores, plotting bar charts of principal components and cumulative contribution line charts.We have found that after taking the first three components, the contribution rate to happiness has already reached 85%, so we believe that the principal components can be taken as the first three for data dimensionality reduction.Principal Component 1:(0.5117,0.4797, 0.5032,0.3736,0.0733, 0.3316)PrincipalComponent2:( -0.2401,-0.1161,-0.1987,0.3296,0.8369,0.2837) Principal Component 3:( 0.0486,-0.3930,-0.0262,-0.2362, -0.2429, 0.8530). PCA helps us better understand the data structure and identify the factors contributing most to happiness scores.



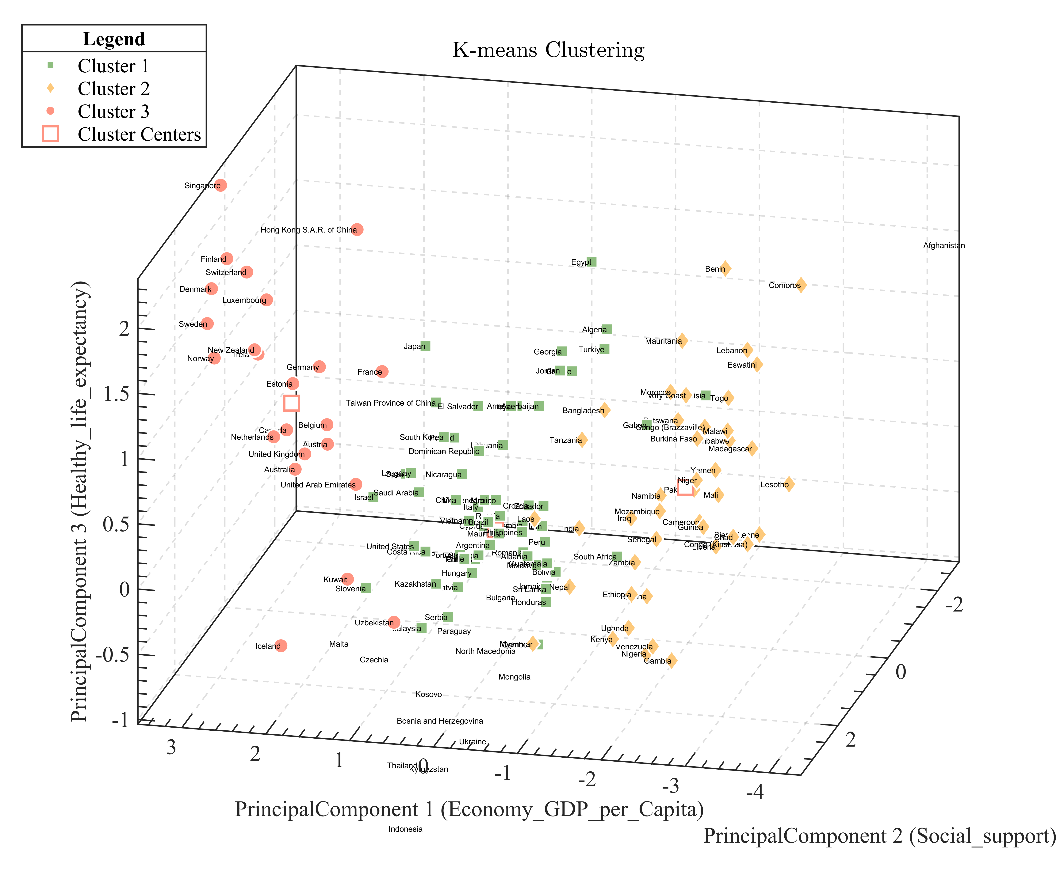
1. **Correlation Heatmap:** Analyzes the correlation among seven variables (economy, social support, healthy life expectancy, freedom to make life choices, generosity, perceptions of corruption, and happiness score), creating a 7x7 correlation heatmap. Happiness scores show strong positive correlations with economy, social support, and healthy life expectancy, while generosity and perceptions of corruption have relatively smaller impacts.



1. **Scatter Plot:** Shows the relationship between happiness scores and the most correlated variable (social support), indicating that social support significantly affects happiness scores. Countries with higher social support tend to have higher happiness scores.



1. **K-Means Clustering:** Uses the K-means algorithm to cluster countries based on PCA results, displaying the distribution of different groups in the principal component space. Clustering helps identify groups of countries with similar characteristics and further analyze their differences in economy, social support, and healthy life expectancy.



**3.Visualization Method Explanation**

1. **World Map Heatmap:** Uses color intensity to represent happiness score levels, helping us intuitively understand the global distribution of happiness levels. The heatmap clearly shows that countries with higher happiness scores are concentrated in Northern and Western Europe, while those with lower scores are mostly in Africa and South Asia.
2. **Principal Component Analysis (PCA):** Employs dimensionality reduction to transform multiple variables into a few principal components, facilitating a better understanding of the data structure.From our observations, we can see that the contribution to happiness has already reached 85% when taking the first three components, so we select the first three as the main components for data dimensionality reduction. Principal Component 1:（0.5117,0.4797,0.5032,0.3736,0.0733,0.3316) Principal Component 2:(-0.2401,-0.1161,-0.1987,0.3296,0.8369,0.2837) Principal Component 3:(0.0486,-0.3930,-0.0262,-0.2362,-0.2429,0.8530) PCA helps identify the factors contributing most to happiness scores.
3. **Correlation Heatmap:** Uses color intensity to represent the strength of correlations between variables, helping identify which factors most influence happiness scores. For example, economy and social support show strong correlations with happiness scores, while generosity and perceptions of corruption have relatively smaller impacts.
4. **Scatter Plot:** Displays the relationship between two variables through point distribution, visually showing the positive correlation between social support and happiness scores. Countries with higher social support tend to have higher happiness scores.
5. **K-Means Clustering:** Uses the K-means algorithm to group countries, displaying the distribution of different groups in the principal component space. Clustering helps identify groups of countries with similar characteristics and further analyze their differences in economy, social support, and healthy life expectancy.

**4.Libraries Used**

We used several Matlab libraries for data analysis and visualization, including:

1. Mapping Toolbox: Create a geobubble chart
2. Statistics and Machine Learning Toolbox: Plotting a heatmap, Perform K-means clustering and plot a three-dimensional scatter plot,Perform principal component analysis.

**5.Results and Significance**

Through analysis, we found:

1. Economy (GDP per capita), social support, and healthy life expectancy are key factors influencing happiness scores. These factors show strong positive correlations with happiness scores. Countries with higher economic levels generally have higher happiness scores, while social support and healthy life expectancy also significantly impact happiness scores.
2. Generosity and perceptions of corruption have relatively smaller impacts on happiness scores but still significantly affect certain countries. For example, generosity may positively influence happiness scores in some countries, while perceptions of corruption may reduce them.
3. Nordic countries such as Finland, Denmark, and Iceland rank at the top of the happiness index, closely related to their high economic levels, strong social support, and good healthy life expectancy. These countries' happiness scores are significantly higher than the global average.
4. K-Means Clustering reveals the distribution of different country groups in the principal component space, helping identify groups with similar characteristics. For example, some developing countries show similarities in economy and social support, while developed countries excel in healthy life expectancy and economic levels.

These results not only help us understand the global distribution of happiness levels but also provide policymakers with directions to improve national happiness. For instance, enhancing social support and healthy life expectancy may be effective ways to boost happiness scores. By analyzing the happiness indices and driving factors of different countries, policymakers can formulate more targeted policies to improve national happiness levels.

Through this report, we aim to provide readers with a comprehensive perspective on the current state of the global happiness index and its underlying driving factors. We believe these analysis and visualization results will offer valuable references for future research and policymaking. By deeply analyzing the driving factors of happiness indices, we can better understand the differences in happiness levels across countries and provide scientific evidence for improving happiness indices globally.

**6.remark**

|  |  |  |  |
| --- | --- | --- | --- |
| **Individual Contribution** | | | |
| **CWID** | **Name** | **Contribution (description)** | **Percent**  **Contribution** |
| **A20563409** | **Han Liu** | Data collection and submission, mathematical modelling | **33.3%** |
| **A20563422** | **Kun Liu** | Writing and revising of the paper | **33.3%** |
| **A20563411** | **Zhonglin Li** | Data preprocessing, programming analysis, and graphs | **33.3%** |
|  |  |  |  |

Data for this assignment are obtained from the publicly available dataset of the Kaggle website. All data download preprocessing, chart programming and paper drawing were done independently by the team. This data set is about the influencing factors of happiness in 143 countries worldwide.