Analysis on World Happiness

ETC5513 Assignment 4

Yu Luo, Khushee Thakker, Yunqi Chen

1 Introduction

Countries are ranked each year based on a **Happiness Index**. This index rates the happiness of countries on a scale from 0 to 10. The happiness index is also called the ladder score or the happiness score.

A column is added for measuring the happiness rank. Happiness rank is given for the ladder score. The greater the ladder score, the higher the rank.

We use the six variables to help us understand the sources of variations in happiness among countries and over time.

2 Happiness Score by Region

This section is created by Khushee Thakker.

- We have collected happiness data from World Happiness Report for 2020. It has global happiness that ranks 156 countries by how happy their citizens perceive themselves to be.

- We grouped the countries into different regions in the world to observe which region is the happiest.

- From the box plot 2.1 we see that the happiness scores are quite different among these regions.

- The group determined that the **"happiest"** countries were located in **Western Europe**, particularly Finland and Denmark.

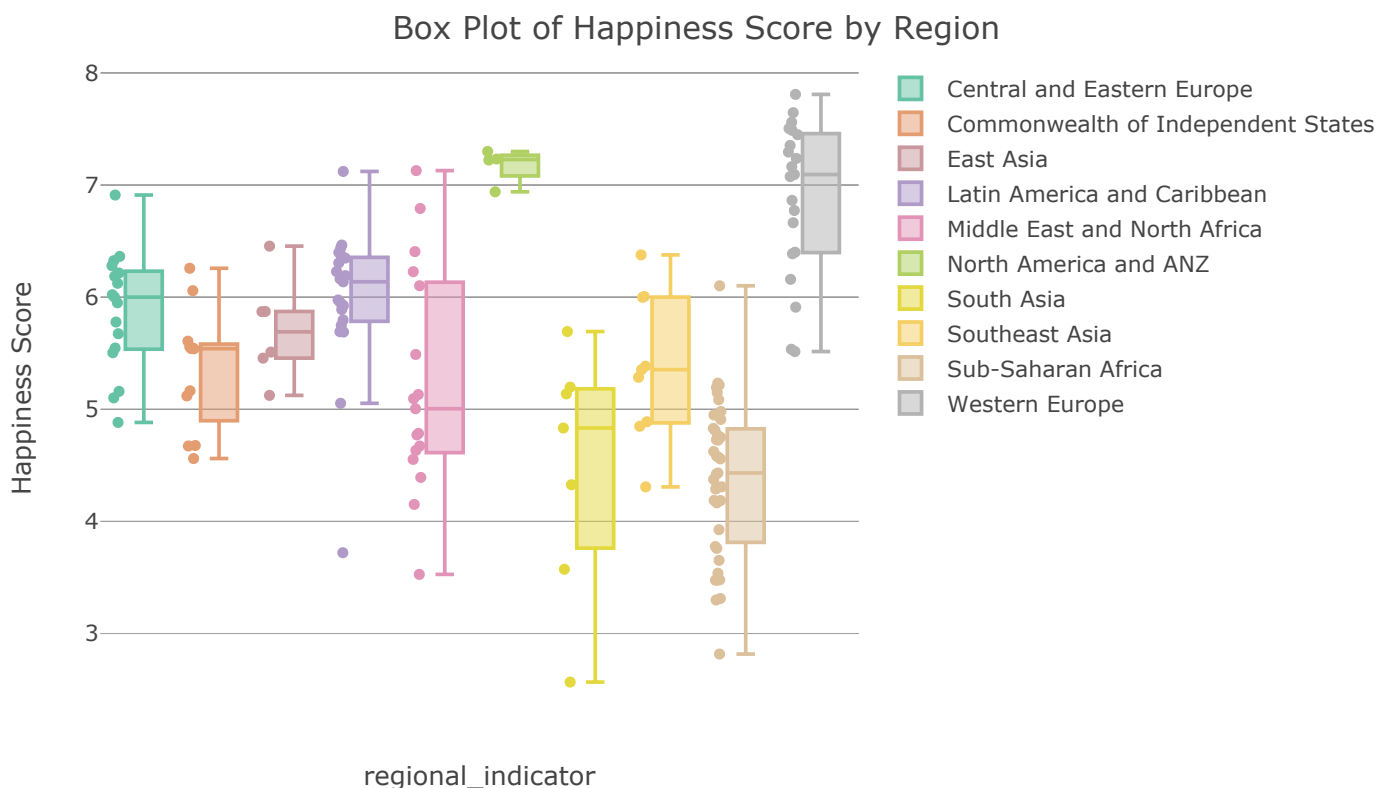- Meanwhile the **"least happiest"** countries were located in **Sub-Saharan Africa** and **South Asia**.

Figure 2.1: Happiness Score by Region

3 The Six Variables Might Affected Happiness

3.1 GDP

This subsection is created by Yu Luo.

Table 3.1 shows the average Logged GDP per capita of Regions from high to low. The top 3 regions with the highest GDP are North America and ANZ, Western Europe and East Asia, their GDP is all higher than 10.

The average GDP for all the countries is 9.2957058, and we found three regions' average GDP are lower than this level: Sub-Saharan Africa, South Asia, Commonwealth of Independent States. We divided into two groups, the regions that regional average GDP higher than global, and regions with regional GDP lower than global, to see how is the relationship between the Logged GDP per capita and the Ladder Score of countries in these regions.

Table 3.1: Average Logged GDP per capita of Regions from High to Low

Regional indicator	Average
North America and ANZ	10.71
Western Europe	10.69
East Asia	10.32
Central and Eastern Europe	9.98
Middle East and North Africa	9.71
Southeast Asia	9.37
Latin America and Caribbean	9.30
Commonwealth of Independent States	9.21
South Asia	8.56
Sub-Saharan Africa	7.90

In Figure 3.1, We found the countries in Western Europe and North America and ANZ, and most of the countries in East Asia and Central and Eastern Europe, have a higher score on either Ladder Score and GDP. In the Middle East and North Africa and South Asia, they are polarized, several countries are with high GDP and high Ladder Scores, several got low scores either on Ladder Score or GDP.

While what we found interesting and different from other Region is in Latin America and Caribbean, some countries even their GDP is lower than the average level, their Ladder Score is high. On the contrary, some countries in the Middle East and North Africa, they got higher GDP score, however, their Ladder Scores are low.
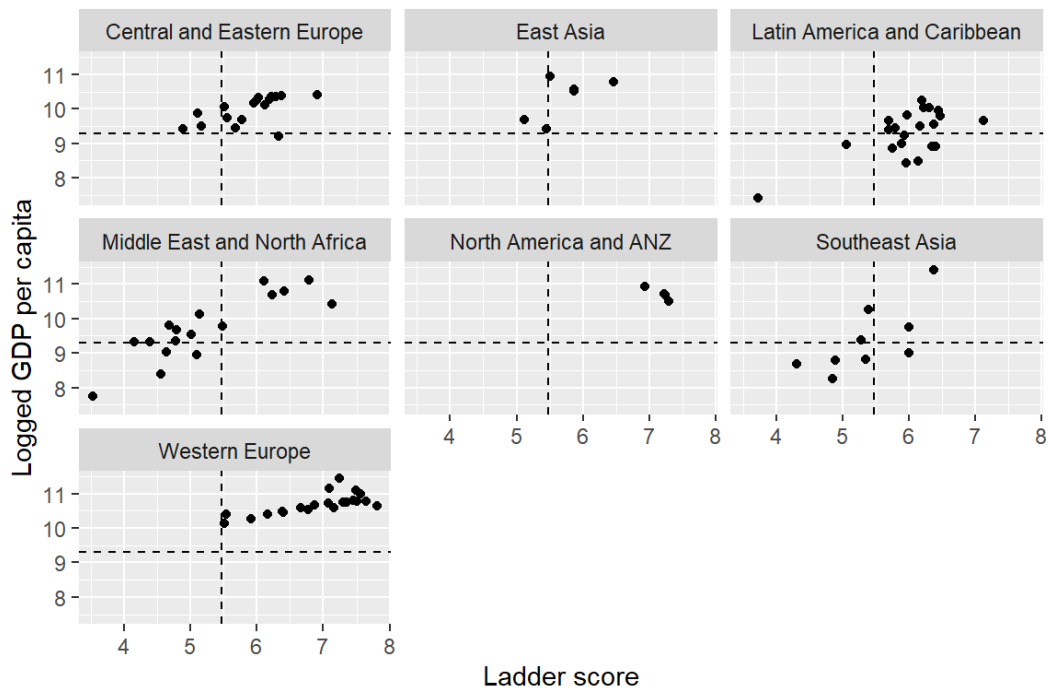
Figure 3.1: Ladder Score vs GDP for the Regions with Regional GDP Means Higher than Total Means

Below Figure 3.2 shows the regions whose regional GDP scores are lower than the average GDP of all the countries, it indicates most countries in these regions, their Ladder Scores are also lower.
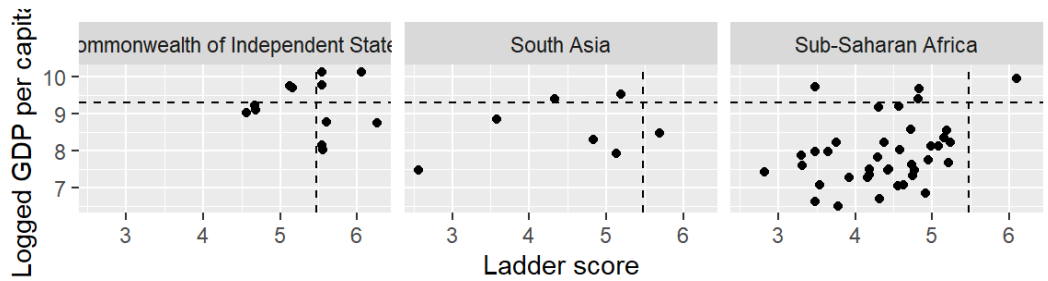

Figure 3.2: Ladder Score vs GDP for the Regions with Regional GDP Means Lower than Total Means

3.2 Social Support

This subsection is created by Yu Luo.

The average Social Support of regions from high to low is shown in Table 3.2.

The same method as GDP, we calculated the average Social Support of all counties, it is 0.8087211.

Table 3.2: Average Social Support of Regions from High to Low

Regional indicator	Average
North America and ANZ	0.93
Western Europe	0.91
Central and Eastern Europe	0.87
Commonwealth of Independent States	0.86
East Asia	0.86
Latin America and Caribbean	0.86
Southeast Asia	0.82

Regional indicator	Average
Middle East and North Africa	0.80
South Asia	0.71
Sub-Saharan Africa	0.68

Either the high Social Support group (shows in Figure 3.3) or low Social Support group (Figure 3.4), we nearly found no country has low Social Support score but got high Ladder Score. The relationship between Social Support and Ladder Score is close to a positive correlation. The better Social Support, the more possible it to get a better Ladder Score; on the opposite, as shown in Figure 3.4, the worse Social Support, the lower the Ladder Score.

But not all the high score Social Support countries also got high Ladder Scores. In Central and Eastern Europe, Commonwealth of Independent States, Middle East and North Africa, South Asia, and Sub-Saharan Africa, some countries with high Social Support scores got low Ladder Scores.
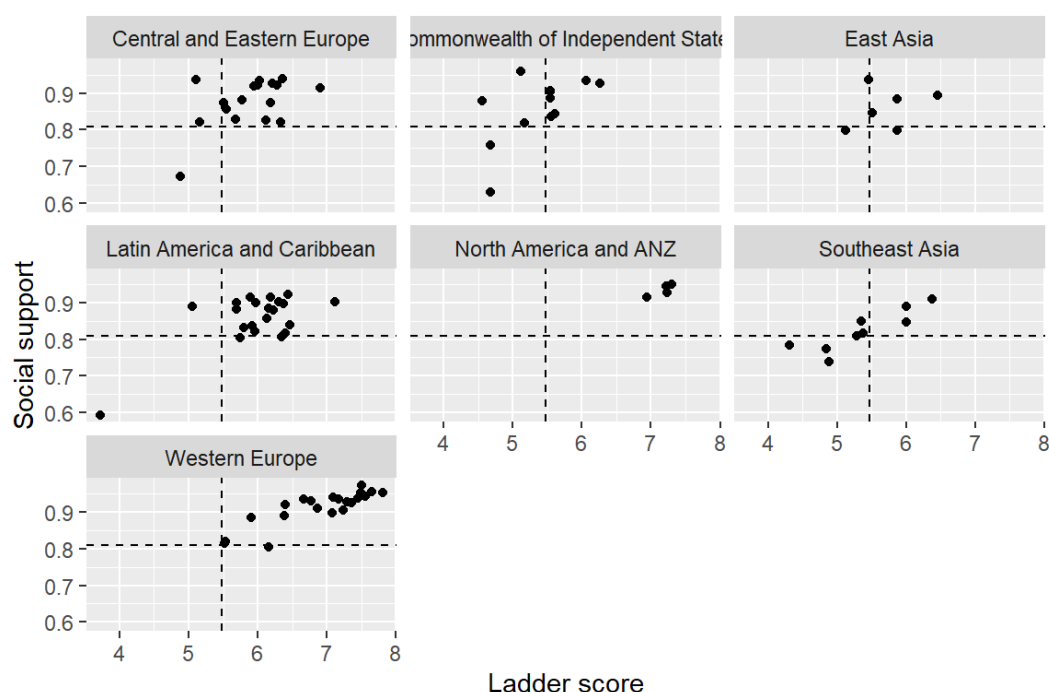

Figure 3.3: Ladder Score vs Social Support for the Regions with Regional Social Support Means Higher than Total Means
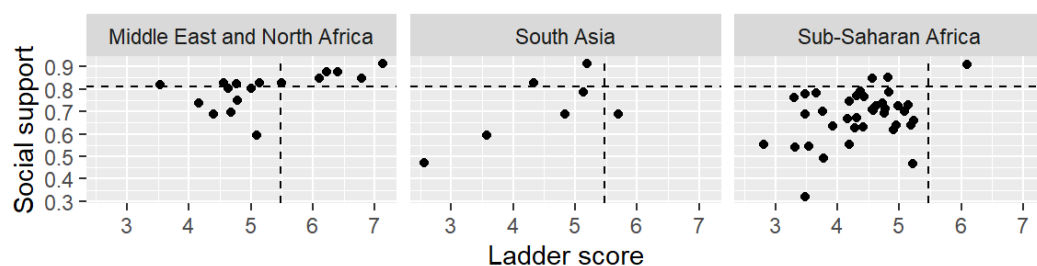

Figure 3.4: Ladder Score vs Social Suppor Lower than Total Means

3.3 Healthy Life Expectancy

This subsection is created by Yu Luo.

Table 3.3 shows the ranking of Social Healthy Life Expectancy score of regions.

Table 3.3: Healthy life expectancy

Regional indicator	Average

Regional indicator	Average
Western Europe	72.86
North America and ANZ	72.18
East Asia	71.09
Central and Eastern Europe	68.15
Latin America and Caribbean	66.72
Middle East and North Africa	65.31
Commonwealth of Independent States	64.73
Southeast Asia	64.71
South Asia	62.45
Sub-Saharan Africa	55.09

In the high Healthy Life Expectancy score countries group, it shows a positive correlation in Southeast Asia, Middle East and North Africa and North America and ANZ. But we found in the Commonwealth of Independent States, Central and Eastern Europe, Western Europe, some countries are almost at the same level of Healthy life Expectancy, their Ladder Scores arrange from a big difference. See that in Figure 3.5.
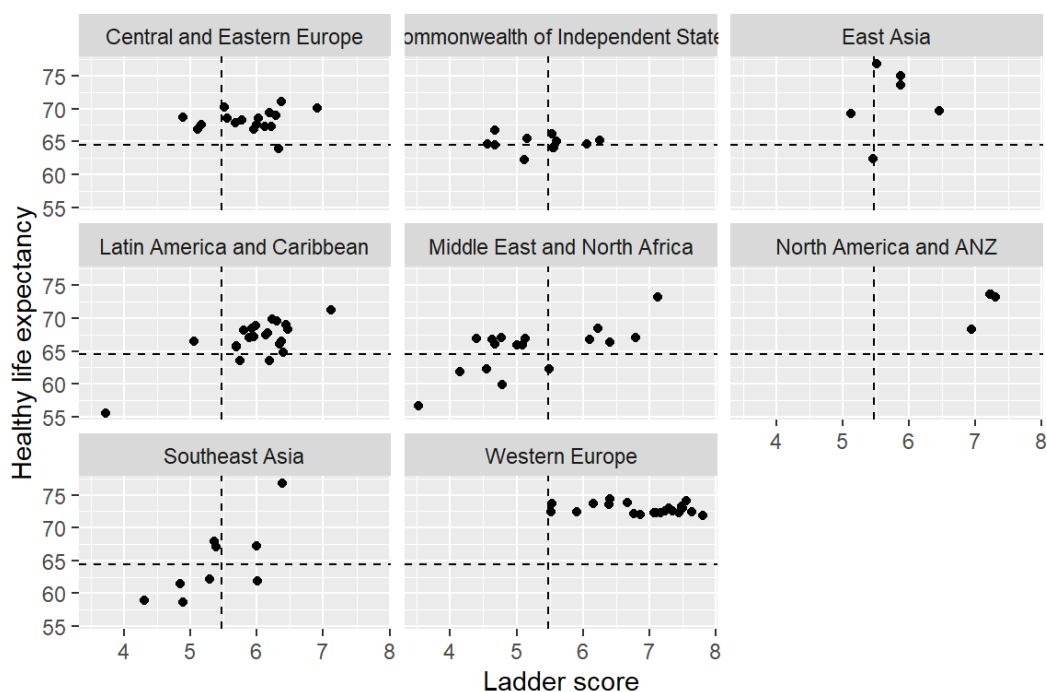

Figure 3.5: Ladder Score vs Healthy Life Expectancy Higher than Total Means

Figure 3.6 shows, in the low Healthy Life Expectancy region, they also have a low Ladder Score, while they don't look have a special relationship between Healthy Life Expectancy and Ladder Score.
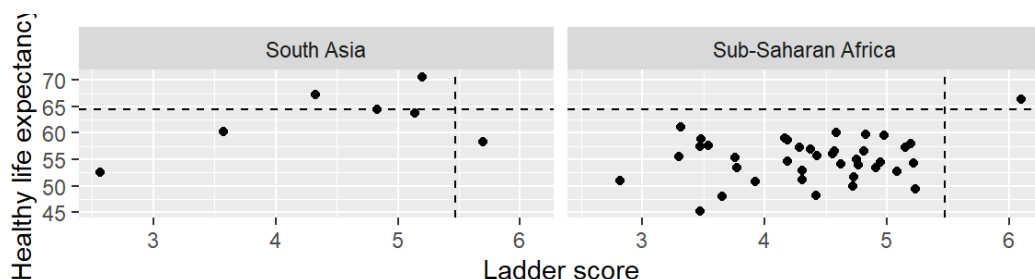
Figure 3.6: Ladder Score vs Healthy Life Expectancy Lower than Total Means

3.4 Freedom to make life choices

This subsection is created by Yunqi Chen.

Freedom to make life choices is the national average of binary responses to the GWP question, "Are you satisfied or dissatisfied with your freedom to choose what you do with your life?".

Table 3.4: The Average Freedom to Make Life Choices Score by Region

Region	Average Score
Southeast Asia	0.9126666
North America and ANZ	0.9070455
Western Europe	0.8549933
Latin America and Caribbean	0.8306836
Commonwealth of Independent States	0.7840038
South Asia	0.7718054
Central and Eastern Europe	0.7702710
East Asia	0.7610266
Sub-Saharan Africa	0.7197015
Middle East and North Africa	0.7101706

The above table 3.4 is the average "freedom to make life choices" score among all the regions. It is divided into three groups by the average score greater than 0.80, 0.75-0.80, and lower than 0.75.
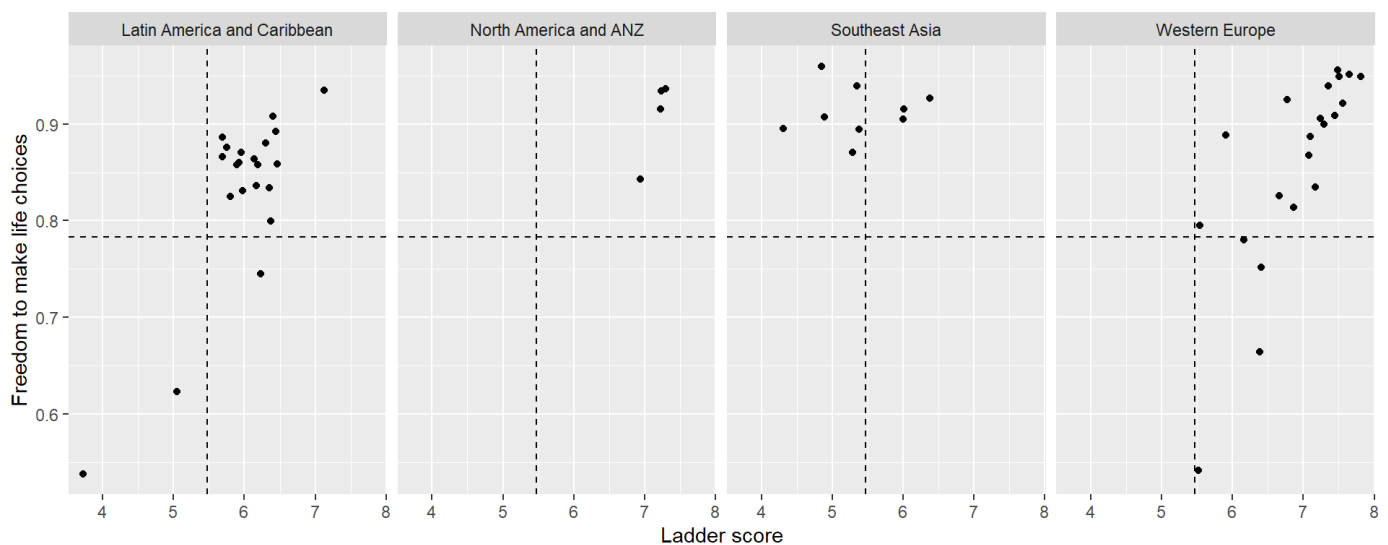

Figure 3.7: The Group of High Freedom to Make Life Choices Score

As the figure 3.7 above, in the group of scores higher than 8.0, except Southeast Asia, the ladder scores in the other three regions are higher than the average.
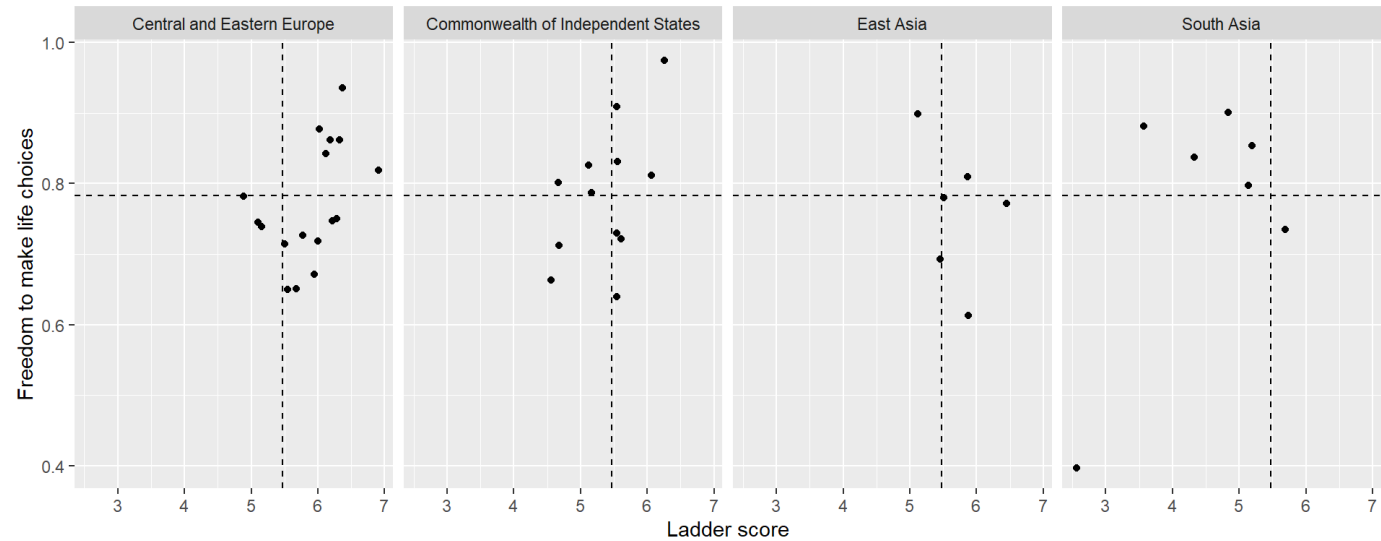
Figure 3.8: The Group of Middle Freedom to Make Life Choices Score

The above figure 3.8 is the group of scores of 0.75-0.80. The distribution of the points is generally surrounding the average line. However, in South Asia, most of the points are lower than the average.
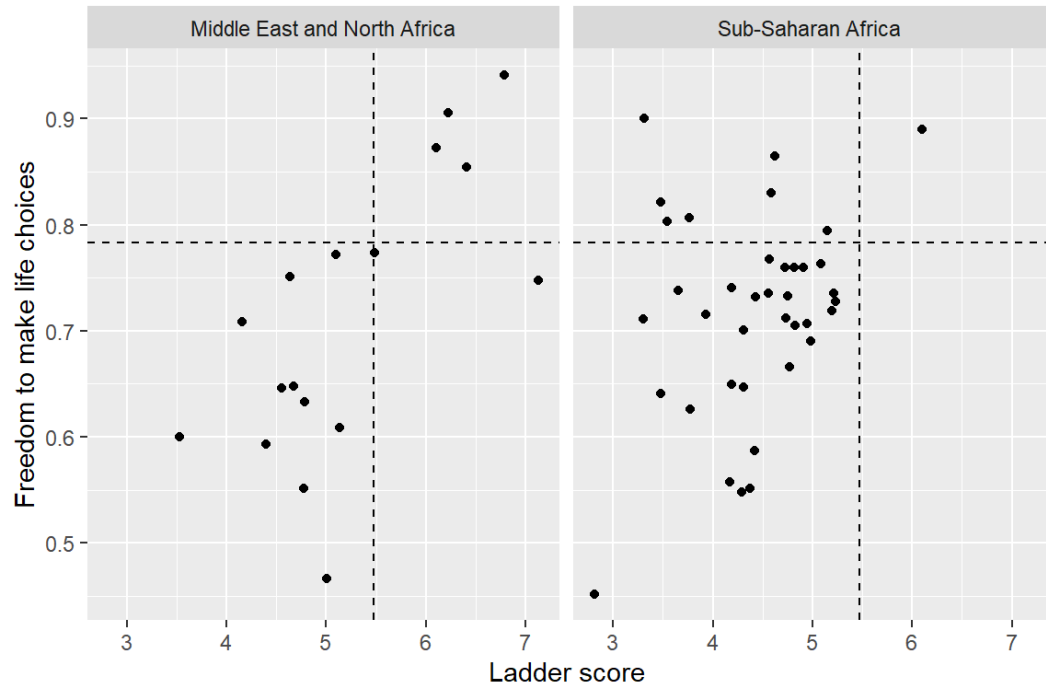

Figure 3.9: The Group of Low Freedom to Make Life Choices Score

From the above figure 3.9, in the group of scores less than 0.75, a tremendous amount of the points of Sub-Saharan Africa are on the left of the average line. However, the points are scattered evenly from high to low in the Middle East and North Africa.

In general, the ladder score is related to the "Freedom to make life choices" score in most of the regions, except Southeast Asia, and Middle East and North Africa.

3.5 Generosity

This subsection is created by Yunqi Chen.

Generosity is the residual of regressing the national average of GWP responses to the question, "Have you donated money to a charity in the past month?" on GDP per capita.

The table below 3.5 is the average generosity score among all the regions. It is divided into two groups by the average score over and below 0.

Table 3.5: The Average Generosity Score by Region

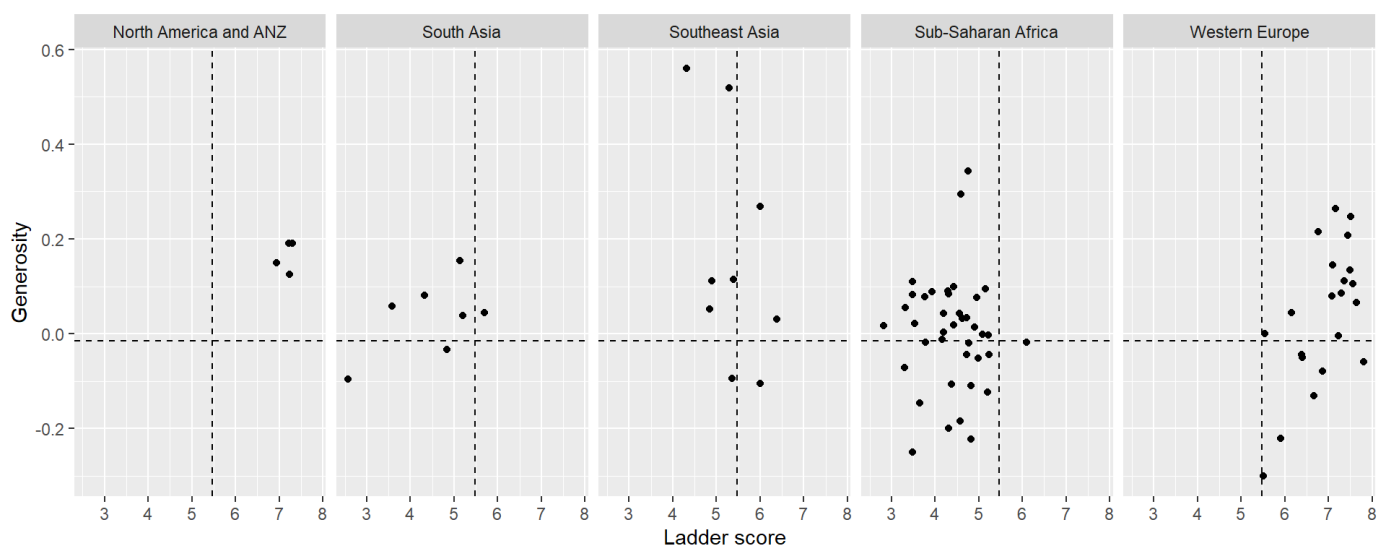| Region	Average Score
North America and ANZ	0.1641802
Southeast Asia	0.1618698
Western Europe	0.0389232
South Asia	0.0351059
Sub-Saharan Africa	0.0023831
Commonwealth of Independent States	-0.0412847
East Asia	-0.0519775
Latin America and Caribbean	-0.0719013
Middle East and North Africa	-0.0844265
Central and Eastern Europe	-0.1027124

Figure 3.10: The Group of the Generosity Score Higher Than 0

From the above figure 3.10, in the group of the generosity score higher than 0, the ladder score in North America and ANZ, and Western Europe is higher than the average. On the contrary, in South Asia and Sub-Saharan Africa, the points overwhelmingly position left than the average line. The points for Southeast Asia is around the average line.
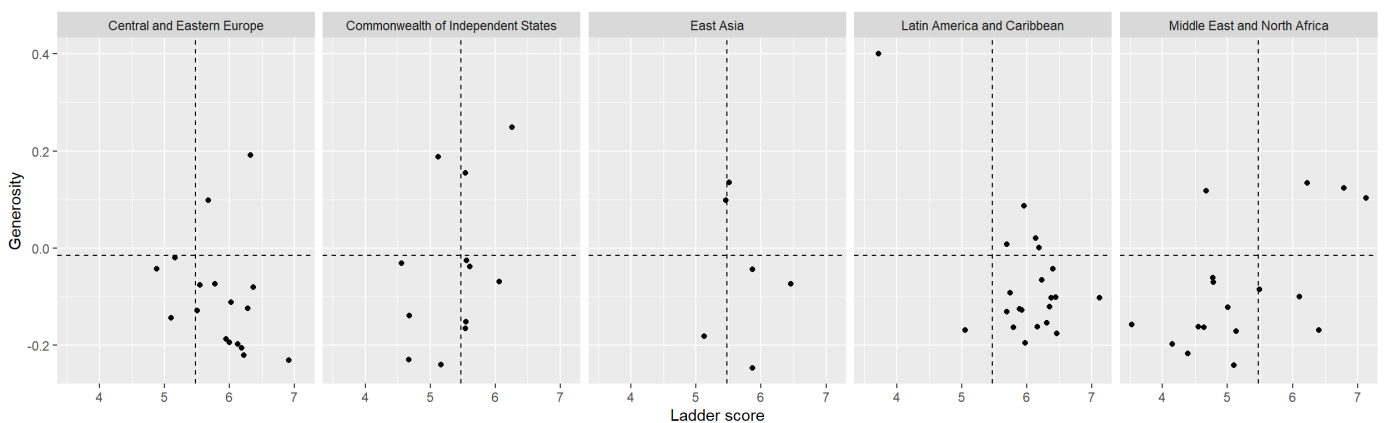

Figure 3.11: The Group of the Generosity Score Less Than 0

The above figure 3.11 indicates that in the group of the generosity score less than 0, except of Latin America and Caribbean, the ladder score distributed scattered around the average ladder score line.

In general, the generosity score does not show a relationship with the ladder score.

3.6 Perceptions of corruption

This subsection is created by Yunqi Chen.

Perceptions of corruption are the average of binary answers to two GWP questions: "Is corruption widespread throughout the government or not?" and "Is corruption widespread within businesses or not?" Where data for government corruption are missing, the perception of business corruption is used as the overall corruption-perception measure .

The table below 3.6 is the average "perceptions of corruption" score among all the regions. It is divided into three groups by the average score greater than 0.75, 0.70-0.75, lower than 0.70.

Table 3.6: The Perceptions of Corruption Score by Region

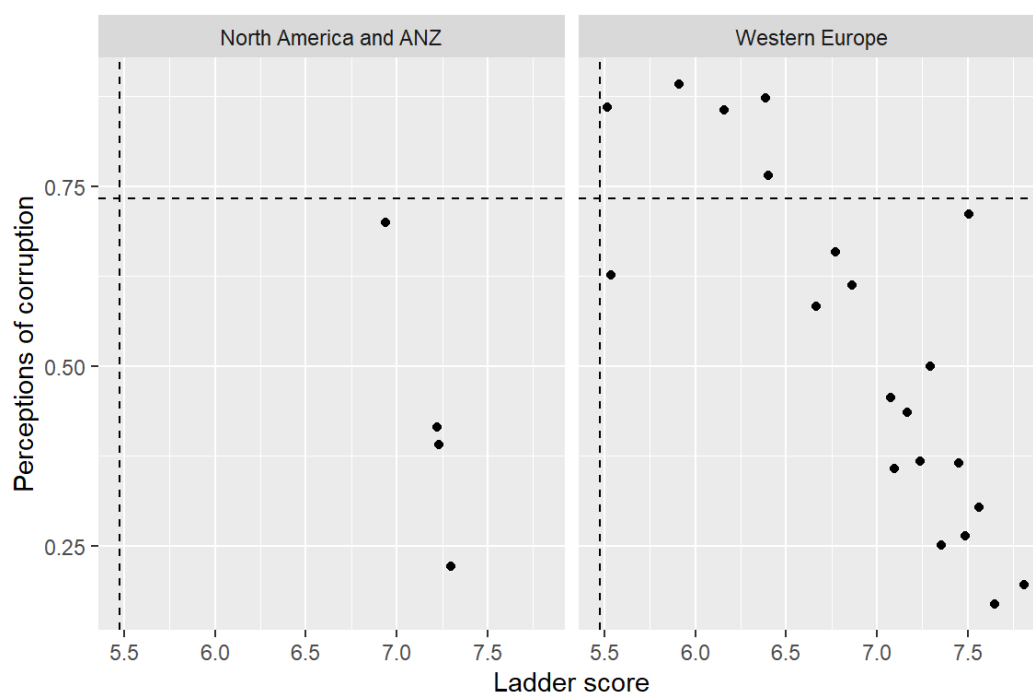| Region	Average Score
Central and Eastern Europe	0.8508587
Latin America and Caribbean	0.8015262
South Asia	0.7906557
Sub-Saharan Africa	0.7704044
Middle East and North Africa	0.7615096
Commonwealth of Independent States	0.7463870
Southeast Asia	0.7049204
East Asia	0.7022663
Western Europe	0.5287299
North America and ANZ	0.4317167

Figure 3.12: The Group of Low Perceptions of Corruption Score

From the figure 3.12, in the group of scores less than 0.70, not surprisingly, all the points are on the right of the average ladder score line.
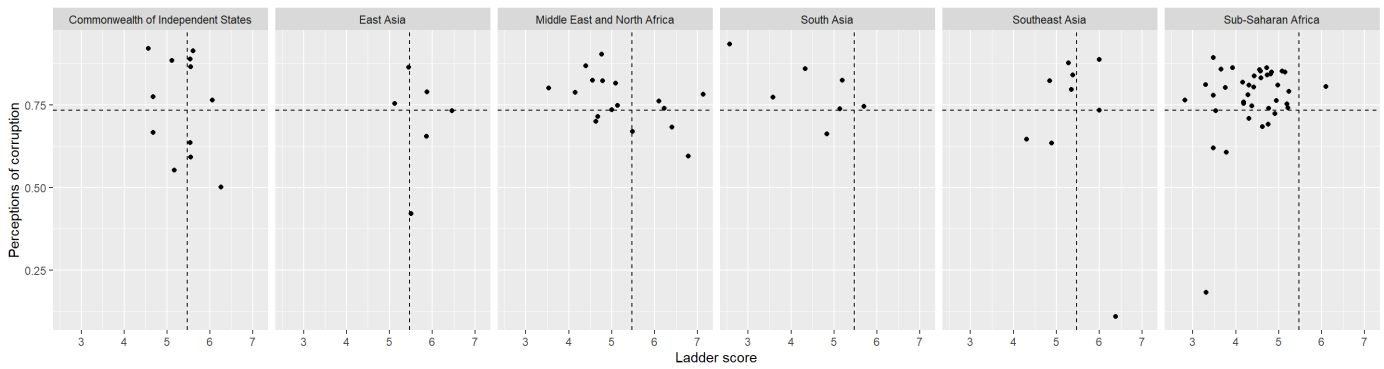
Figure 3.13: The Group of Middle Perceptions of Corruption Score

From the figure 3.13, in the group of scores with between 0.70 and 0.80, except South Asia and Sub-Saharan Africa shows lower ladder score, the other regions all indicate a middle ladder score.
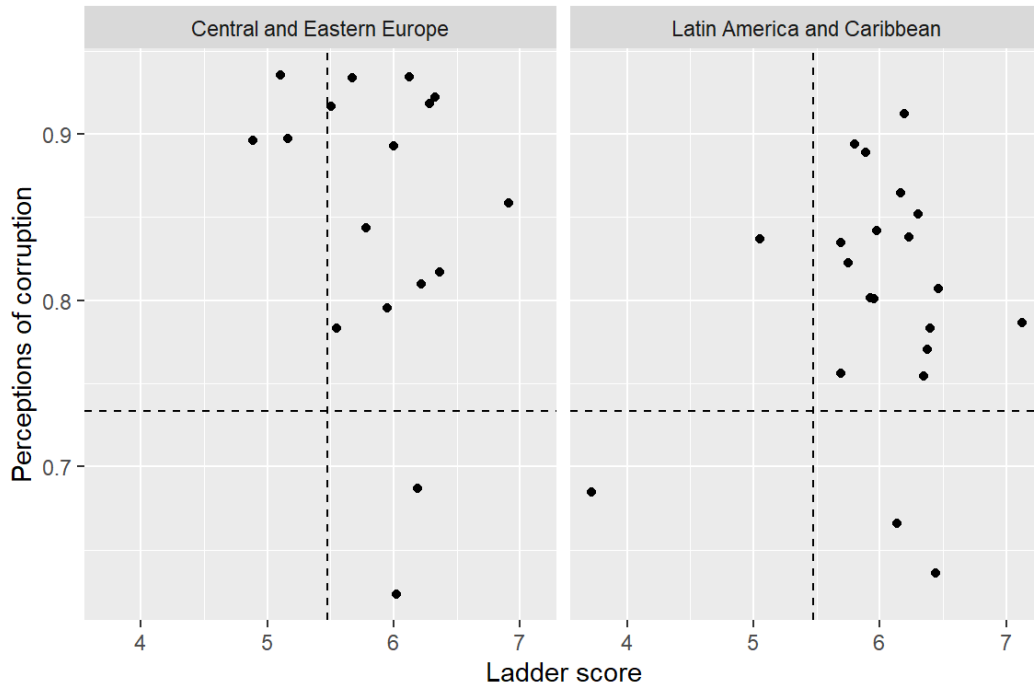

Figure 3.14: The Group of High Perceptions of Corruption Score

The above figure 3.14 is the group of scores greater than 0.80. In Central and Eastern Europe, and Latin America and Caribbean, most of the points are on the right of the average line.

In conclusion, as for the "perceptions of corruption" score, the lowest and the highest four regions all have a high ladder score. Most of the regions in the middle score groups show the middle ladder score.

4 Happiness score in differetent countries

This section is created by Khushee Thakker.

- In the above sections we have seen factors that are reflect happiness score in different regions.
- In the this section we will look happiness score with respect to countries countries.

4.1 Compare Top5, Middle5 and Worst5 Countries

- The below table 4.1 shows the Countries according to their happiness rank. We have divided them into levels, Top5, Middle5 and Worst5 Countries.

Table 4.1: Top 5, Middle5 and Worst 5 Countries

country_name	regional_indicator	Level	happiness_rank	score
Finland	Western Europe	TOP5	1	7.8087

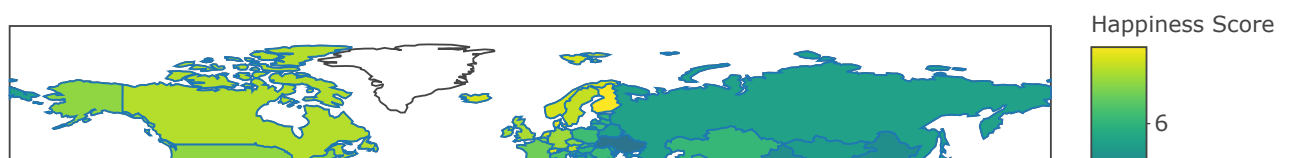| country_name	regional_indicator	Level	happiness_rank	score
Denmark	Western Europe	TOP5	2	7.6456
Switzerland	Western Europe	TOP5	3	7.5599
Iceland	Western Europe	TOP5	4	7.5045
Norway	Western Europe	TOP5	5	7.4880
North Cyprus	Western Europe	MIDDLE5	76	5.5355
Greece	Western Europe	MIDDLE5	77	5.5150
Hong Kong S.A.R. of China	East Asia	MIDDLE5	78	5.5104
Croatia	Central and Eastern Europe	MIDDLE5	79	5.5047
Libya	Middle East and North Africa	MIDDLE5	80	5.4888
Central African Republic	Sub-Saharan Africa	WORST5	149	3.4759
Rwanda	Sub-Saharan Africa	WORST5	150	3.3123
Zimbabwe	Sub-Saharan Africa	WORST5	151	3.2992
South Sudan	Sub-Saharan Africa	WORST5	152	2.8166
Afghanistan	South Asia	WORST5	153	2.5669

Figure 4.1: Compare top median and bottom five countries

- In the above figure 4.1, we have taken the top 5, middle 5 and the bottom 5 countries to understand the relations between 6 factors and happiness score.
- Country with higher happiness score have higher GDP. Happiest countries tend to be those with strong and stable economies.
- Health is particular important because it directly affects individuals in countries.
- We observe that the countries with highest happiness score have a better GDP, less corruption, more freedom to make choices, good health ratio, more generosity and receive social support.

4.2 Happiness Score for Countries
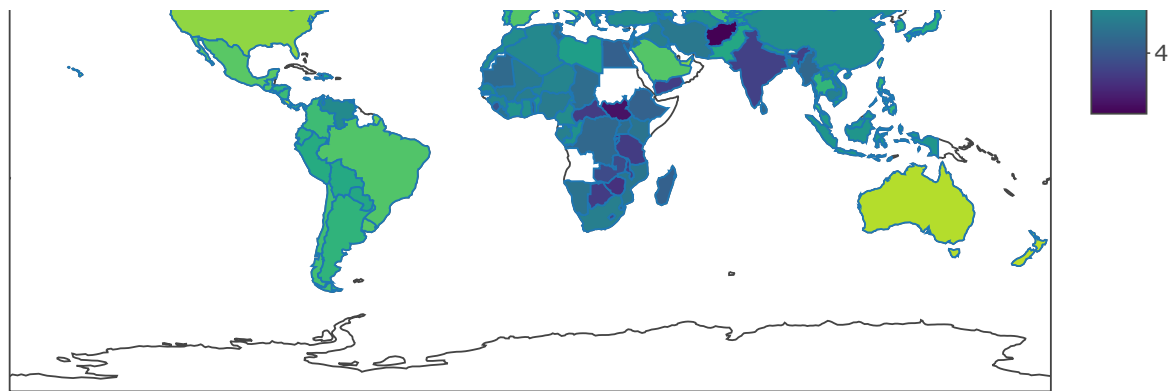
World Happiness Map, Year 2020

Figure 4.2: World Happiness Map

- The World Happiness report calculates how much of the Happiness Index can be explained by the key factors. Add them all up and you should get close to the actual number.

- From the above map 4.2 you can easily spot the happiest country, Finland.

- Finland is ranked 1st according to the happiness score it has. It is said one of the reason why Finnish people are so happy is that they do not have to deal with corruption and black money.

- Countries such as Afghanistan and South Sudan have least happiness score, one one the possible reasons could be the wars happening there.

5 Conclusion

This analysis illustrated that the world's happiest countries are primarily in Western Europe, North America, and Australia & New Zealand, the countries in these regions got the higher Ladder Scores.

The six factors included in the report: Logged GDP per capita, Social Support, Healthy Life Expectancy, Freedom to make life choices, Generosity, basically show a positive correlation between them and Ladder Scores. Most of the countries in high-factors regions got higher Ladder Scores than most of the countries in low-factors regions. But it is definitely a golden rule that once one country or region got high scores on these six factors, the country or region was the happiest one. We found many countries are in high-factor region countries and their factors are higher than some other countries, while their Ladder Score is lower than those countries. Also, we found in some regions, shows there perhaps is a relationship between the factors and Ladder Score, but in some regions, there is no relationship existing.

By analyzing these reports, we are able to decipher what makes countries and their citizens happier, thus allowing us to focus on prioritizing and improving these aspects of each country. It is through this that we are able to achieve the true pursuit of happiness, which we as human beings strive for.

6 References

- Arel-Bundock et al., (2018). countrycode: An R package to convert country names and country codes. Journal of Open Source Software, 3(28), 848, https://doi.org/10.21105/joss.00848 (https://doi.org/10.21105/joss.00848)

- C. Sievert. Interactive Web-Based Data Visualization with R, plotly, and shiny. Chapman and Hall/CRC Florida, 2020.

- Hadley Wickham and Jennifer Bryan (2019). readxl: Read Excel Files. R package version 1.3.1. https://CRAN.R-project.org/package=readxl (https://CRAN.R-project.org/package=readxl)

- Hadley Wickham (2019). stringr: Simple, Consistent Wrappers for Common String Operations. R package version 1.4.0. https://CRAN.R-project.org/package=stringr (https://CRAN.R-project.org/package=stringr)

- Hao Zhu (2021). kableExtra: Construct Complex Table with 'kable' and Pipe Syntax. R package version 1.3.4. https://CRAN.R-project.org/package=kableExtra (https://CRAN.R-project.org/package=kableExtra)

- H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.

- Sam Firke (2021). janitor: Simple Tools for Examining and Cleaning Dirty Data. R package version 2.1.0. https://CRAN.R-project.org/package=janitor (https://CRAN.R-project.org/package=janitor)

- Wickham et al., (2019). Welcome to the tidyverse. Journal of Open Source Software, 4(43), 1686, https://doi.org/10.21105/joss.01686 (https://doi.org/10.21105/joss.01686)

- World Happiness Report. 2020. The World Happiness Report 2020. [online] Available at: https://worldhappiness.report/ed/2020/ (https://worldhappiness.report/ed/2020/)

- Yihui Xie (2020). bookdown: Authoring Books and Technical Documents with R Markdown. R package version 0.21.

- Yihui Xie (2016). bookdown: Authoring Books and Technical Documents with R Markdown. Chapman and Hall/CRC. ISBN 978-1138700109