

How Different Factors Affect the Happiness of Countries From 2015 to 2019

STAT 5740

Team 10: Yunge (Healy) Li and Haocheng Wang

1. Introduction

World happiness score is an index about the quality of daily life of the people in different countries. It is evaluated base on multiple different aspects of a country, including GDP per capital, social support, life expectancy, government corruption, etc. Each of the scores of those variables comes from answers to the main life evaluation question asked in the poll. The respondents are asked to imagine a ladder with the best possible life for them being a 10 and the worst possible life being a 0 and to rate their own current lives on that scale. The data of poll comes from the Gallup World Poll.

The first report was published in 2012, and the fifth report was published at the United Nations at an event celebrating International Day of Happiness March 20th in 2017, which is a meaningful day. Over 150 countries in different regions are researched in this year. The world happiness report continues gaining influences around the world in these years. More and more governments, organizations and civil societies start to use this report to inform their policy-making decisions. The fields that are affected by this report includes but not limited to economics, psychology, and health. This report review population happiness in global scale and express how the progress of technology and society influence the personal and national variations in happiness

Our study looks at the variation of the happiness score for countries in different regions between 2015 and 2019 and how the factors affect the score and rank. We choose these two years because they are far enough from each other, which makes the score changes easier to catch and study. We firstly compare the score of countries in different regions by calculating the means in 2015 and 2019, and then further explore how the factors affect the overall score.

2. Data

Sources

Our data are derived from the Gallup World Poll and we obtained from the website. Our project used two datasets from the original data, which are the world happiness score 2015 and world happiness score 2019. This report is not conducted

every single year. Additionally, the number of countries been researched, and the factors been considered are slightly different within each year.

Our analysis is based on a single dataset created by an outer match merge of the datasets from 2015 and 2019. Countries which are in at least one of the reports are included in this merged dataset. We will analyze on all the countries no matter their score or rank.

Variables

The dataset includes the name of the country, the region that each country belongs to, their rank, overall score, and the specific score for health in 2015 and 2019. The sub-score (health, etc.) is come from the poll in different countries, and the overall score is calculated based on the sub-score. Countries are ranked base on the overall score from high to low. Except from the variables that the dataset originally has, we created variables of the score changes and rank change of all countries from 2015 to 2019.

Summary

There were 158 countries at total participated in the research in 2015 and 2019 and they are categorized base on the region. Their characteristics are summarized in Table 1, including mean and standard deviation. The countries which participate in the 2015 report have a mean score of 5.376, while the mean score of those countries (a couple of countries that participate in 2015 did not participate in 2019) which participate in the 2019 report is 5.434.

Table One

Region		Australia and New Zealand	Central and Eastern Europe	Eastern Asia	Latin America and Caribbean	Middle East and Northern Africa	North America
Sample Size		2	29	6	22	20	2
rank_2015	Mean	9.500	79.00	64.50	46.91	77.60	10.00
	StdDev	0.707	26.67	24.69	28.72	43.21	7.071
score_2015	Mean	7.285	5.333	5.626	6.145	5.407	7.273
	StdDev	0.001	0.570	0.554	0.729	1.101	0.218
Health_2015	Mean	0.920	0.719	0.877	0.704	0.706	0.884
	StdDev	0.016	0.081	0.154	0.109	0.107	0.031
rank_2019	Mean	9.500	70.50	64.83	52.00	84.89	14.00
	StdDev	2.121	27.37	24.49	31.30	44.51	7.071

Region		Australia and New Zealand	Central and Eastern Europe	Eastern Asia	Latin America and Caribbean	Middle East and Northern Africa	North America
score_2019	Mean	7.268	5.572	5.689	5.943	5.237	7.085
	StdDev	0.056	0.591	0.476	0.741	1.060	0.273
Health_2019	Mean	1.031	0.809	0.953	0.817	0.751	0.957
	StdDev	0.007	0.073	0.168	0.106	0.143	0.117
Score_change	Mean	-.017	0.227	0.063	-.195	-.094	-.188
	StdDev	0.054	0.410	0.195	0.637	0.274	0.055
rank_change	Mean	0.000	-8.00	0.333	4.450	4.368	4.000
	StdDev	1.414	19.20	12.23	24.83	10.00	0.000

(Continued)

Region		Southeastern Asia	Southern Asia	Sub-Saharan Africa	Western Europe	All
Sample Size		9	7	40	21	158
rank_2015	Mean	81.22	113.1	127.9	29.52	79.49
	StdDev	39.92	26.57	22.86	29.27	45.75
score_2015	Mean	5.317	4.581	4.203	6.690	5.376
	StdDev	0.950	0.571	0.610	0.825	1.145
Health_2015	Mean	0.677	0.541	0.282	0.909	0.630
	StdDev	0.159	0.123	0.146	0.027	0.247
rank_2019	Mean	85.11	115.9	124.8	21.35	77.57
	StdDev	29.99	29.99	22.14	21.95	45.35
score_2019	Mean	5.274	4.527	4.315	6.898	5.434
	StdDev	0.618	0.793	0.606	0.680	1.111
Health_2019	Mean	0.745	0.617	0.412	1.014	0.732
	StdDev	0.188	0.149	0.156	0.024	0.239
Score_change	Mean	-.044	-.054	0.145	0.159	0.056
	StdDev	0.478	0.384	0.721	0.192	0.511
rank_change	Mean	3.889	2.714	-4.25	-6.35	-1.80
	StdDev	19.72	16.49	26.46	7.555	19.85

3. Specific Questions

a. Specific Question 1 (Healy Li)

Since there are too many countries and their region are classified too specifically, I will just pick two regions and compare their overall mean score in 2015 and 2019. My specific question is: does the score of countries in Europe generally higher than which of countries in Asia? Does this result change in 2019 compare to 2015?

- **Methods**

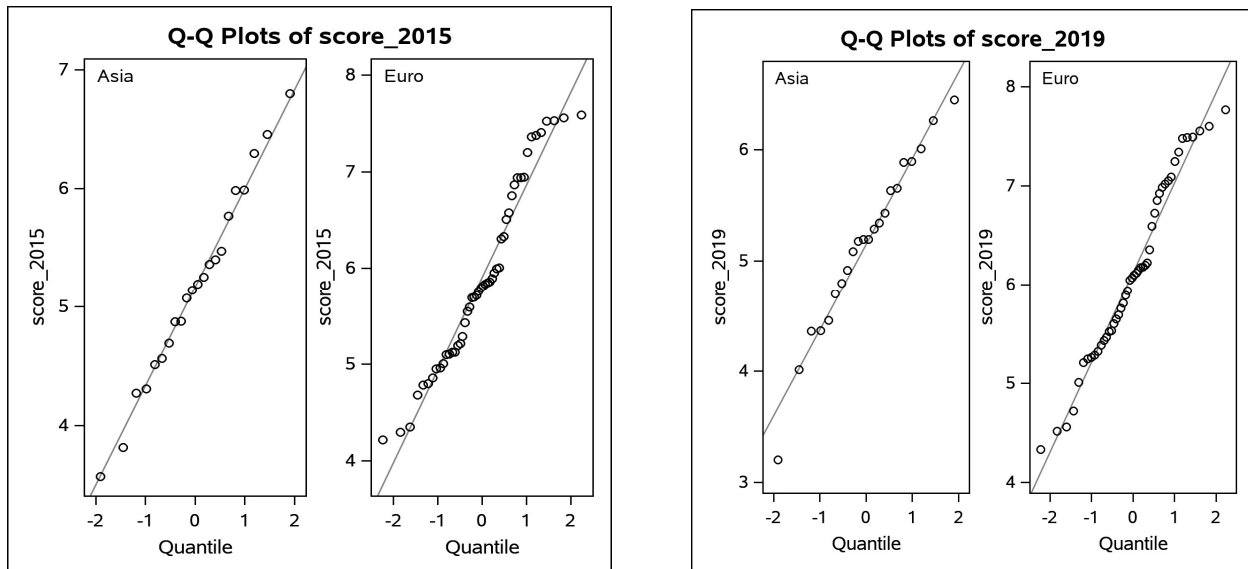
This question compares the mean of happiness score in different years (2015, 2019) across the two groups — European countries and Asian countries. I used two sample t-test to analyze this question. The table I created summarize the difference in mean between Asian countries and European countries, including estimated mean difference, t-test for null hypothesis (difference greater than zero), and the 95% one-sided confidence interval.

- **Results**

I estimate that the mean happiness score of Asian countries is 0.7355 point less than which of European countries in 2015. And I am 95% confidence that Asian countries' scores are at least 0.36 lower than European countries' score in this year. Similarly, the estimated mean score of Asian countries is 0.9753 point less than which of European countries in 2019. And I am 95% confidence that Asian countries' scores are at least 0.62 lower than European countries' score in this year. Therefore, we can conclude that Asian countries' scores are statistically significantly lower than European countries' in both years, and the difference was getting larger.

In 2015, we evaluate 22 Asian countries and 50 European countries, the estimates and test result of countries in Asia mildly violate normality, while which of countries in Europe violate the normality relatively excessive, but within a reasonable range. The following q-q plots indicate that the observed distributions in both Europe and Asia are basically close to a normal distribution. The result from 2019 is also the same.

Summary of T-test



Variable	Method	Estimate	Lower Bound	Upper Bound	T-Value	DF	P-Value	CI
score_2015	Satterthwaite	-0.7355	M	0.3602	-3.29	45.956	0.0010	(M,-0.36)
score_2019	Satterthwaite	-0.9753	M	0.6219	-4.63	47.457	<.0001	(M,-0.62)

b. Specific Question 2 (Haocheng Wang)

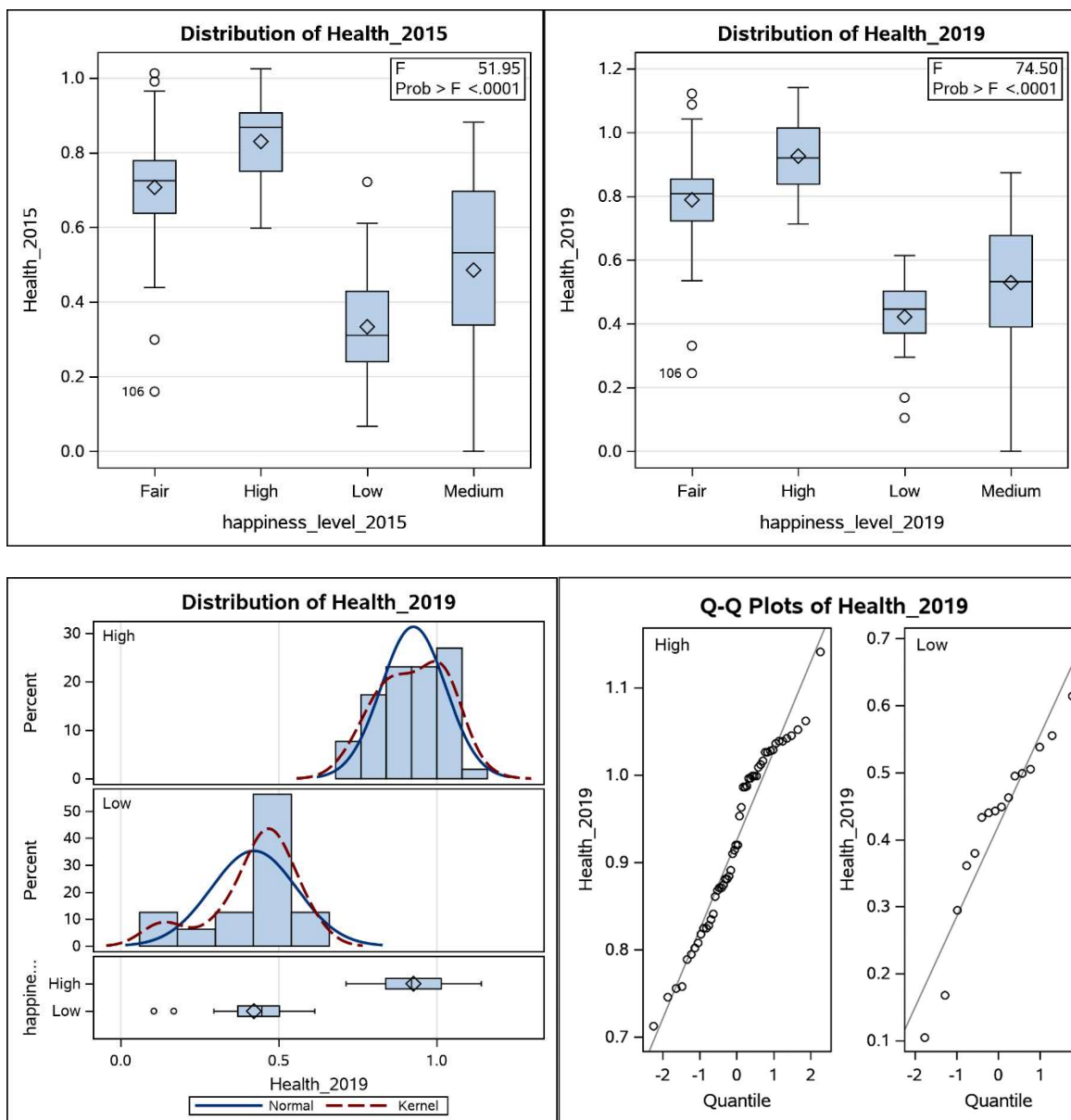
Does the mean health index differ from different happiness scores? Is an individual's higher happiness score associated with a significantly healthier life?

- **Methods**

I am using ANOVA test to find out the mean health index over different groups of happiness level, both in 2015 and 2019. The distribution of health index will help us to compare the health level in different levels of happiness scores. And the results of ANOVA test will give us evidence of whether the mean health score differ from different happiness score. Also, I conducted a t-test specific to the analyze on health score correspond to low and high happiness levels in 2015

• Results

From the distribution boxplot of health index both in 2015 and 2019, we can tell that a higher happiness level is associated with a higher health index. And the p-values in both tests are lower than 0.0001, which provides strong evidence to support the idea that the mean health index differ from different levels of happiness scores. Also, the distribution plot indicated a normal distribution both for High and Low levels of happiness score. Based on the test results, it satisfies my assumption that mean health index differ from different happiness scores. And it satisfies my assumption that a higher happiness score is associated with a significantly healthier life. And by the Q-Q plots of the health score in 2019, regression residuals do not show a clear nonlinear pattern, which corresponds to my assumption.



4. Discussion

By analyzing the happiness scores and related factors collected from people in different countries, we have a better understanding of the sense of well-being across different cultures and how lifestyles, specifically health in our research, are impacted by happiness. We found that Asian people's happiness scores (95% CI: [4.8, 5.5]) are noticeably lower than those in Europe countries (95% CI: [5.9, 6.4]). However, we are not able to identify the specific factors contributing to this result. In the future, more data are needed for us to conduct further research to identify the factors contributing most the happiness different across different regions. Also, our study of the data gives us clear vision of the connection between happiness and health. We find out that they are very positively related. By splitting happiness scores into different levels, our tests indicated higher happiness level is strongly corresponds to better health score. The latest data in 2019 included 52 observations of high level of happiness score, with health scores in 95% CI: [0.90, 0.95], and 16 observations of lower level of happiness scores, with health scores only in 95% CI: [0.35, 0.49]. Further data of other factors related to happiness would provide us with more insights on how happiness score is affected in more ways.

5. References

- Kaggle. 2015. World Happiness Report. Accessed Oct. 13, 2020.
<https://www.kaggle.com/unsdsn/world-happiness?select=2015.csv>
- Kaggle. 2019. World Happiness Report. Accessed Oct. 13, 2020.
<https://www.kaggle.com/unsdsn/world-happiness?select=2019.csv>
- BSAS Institute, Inc. (2018) SAS Studio 3.8 (Enterprise Edition) for Linux [Computer software].