Report on World Happiness Index

Pratyusha Bala, Tanmay Purohit , Kundan Kumar, Shem Basumatary

20/11/2022

Report by GROUP-18

Course : MTH208A

Course Instructor : Dr. Dootika Vats

Indian Institute of Technology,Kanpur

About the Data:

The **World Happiness Report** is a publication that contains articles and rankings of national happiness based on global survey data to report how people evaluate their own lives in more than 150 countries worldwide, which the report also correlates with various quality of life factors. This year marks the 10th anniversary of the World Happiness Report, and we have chosen to analyze the world happiness report to find out what factors influence happiness across the world.

Our dateset **World.happiness.report** contains world happiness reports of five year (2015 to 2019) containing -

1. Overall ranking
2. Country name
3. Happiness Score
4. GDP per capita
5. Life expectancy
6. Freedom of life choices
7. Generosity
8. Government corruption

Other than this we have also scraped

– **Continent** dataset which contains

1. country name
2. Continent

– **poverty** dataset which contains

1. Country
2. percentage poverty

– **Population.density** data set which contains

1. Country
2. Population Density

– **literacyrate** dataset which contains

1. Total literacy rate
2. Male literacy rate
3. Female literacy rate
4. Gender Gap
5. Year

Since data of different countries are missing therefore Out of 150 countries data we have choosen 140 countries and their attributes in different datasets for our observation.

Obtaining the data

Dataset which we have taken as csv files are –

1. world.happiness.report from 2015 - 2019

Dataset which we have taken as shp files, for making the interactive map are –

1. World Countries.shp

And Other web- scraped datasets are as stated formerly.

Biases/Misinformation in the data :

1) The disparity between factors : **Life expectancy** and **GDP per capita** present in the World Happiness Report have been computed from ready made data of *World Health Organisation* and *World Bank* respectively. While for the other factors they survey typically 1000 people in a country annually. The surveyors state that they are unable to reach out to many countries for the annual survey. In countries where they have been able to conduct annual surveys on a recurrent basis , the sample size spikes to 3000. So, there is some selection bias present in the data.

2) For each year , the number of countries present in the dataset is not constant. Rather , some countries are featured only in a few datasets and not in all of them .

3) Data is missing for a lot more African and Asian countries than European and North American countries. This gives us a partial understanding about some areas while we get a more comprehensive understanding about other areas.

Some Interesting questions to ask from the data:

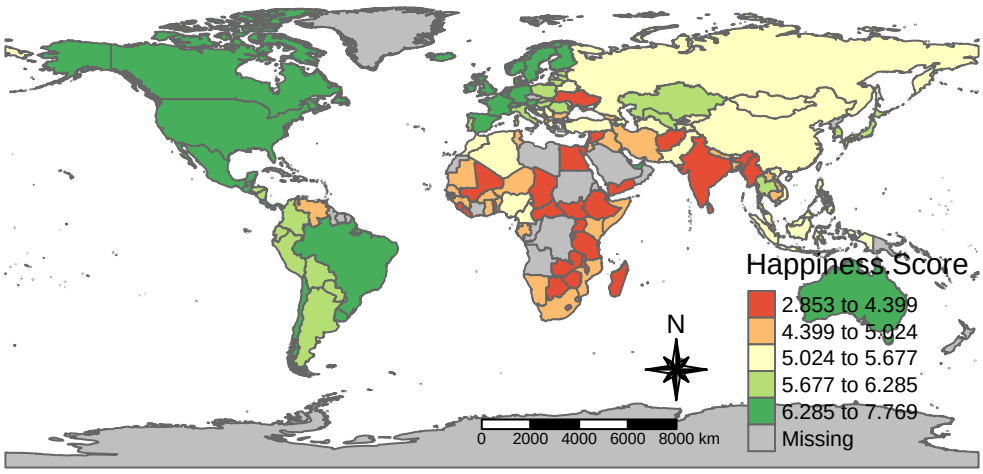
From the dataset we get the following Interesting questions :

- What are the attributes which best quantify happiness ?

- Are all factors given in the report necessary & reliable quantifiers of happiness or are they biased in any way ?
- Are there any paradoxes contrary to our logical thinking observed through data visualization ?
- Are there any other factors which can improve the report ?

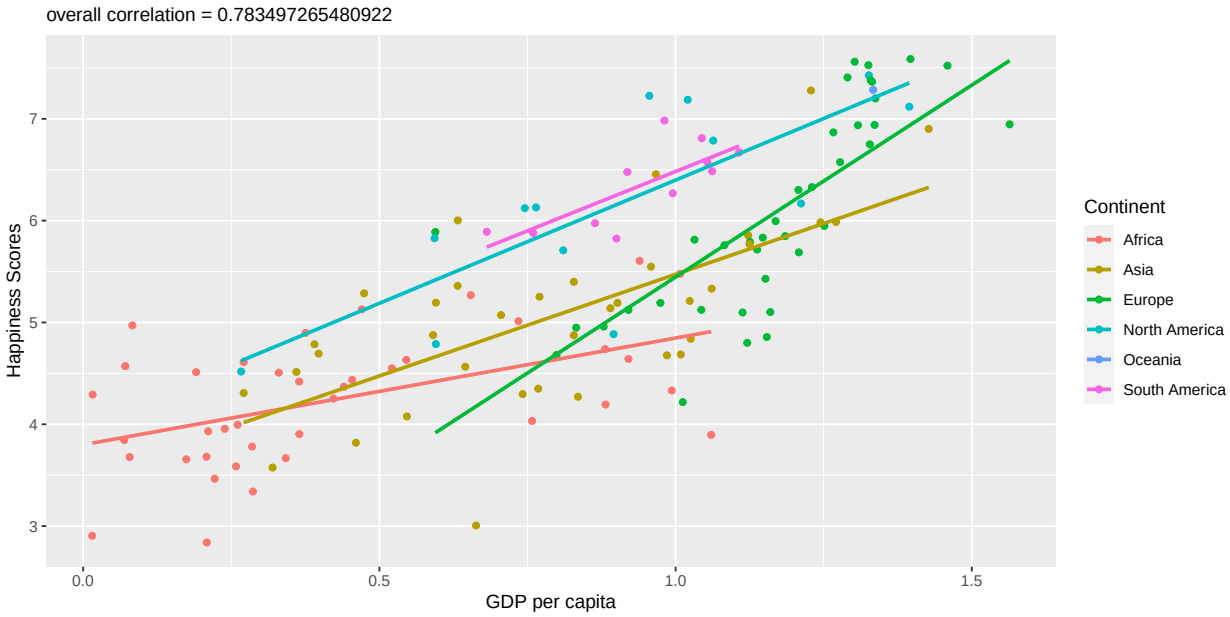
Important visualizations:

Here we have a world map showing the overall distribution of the Happiness Index over the countries in 2019. We can see that countries in the Nordic region like Finland, Sweden have high scores, whereas countries in Africa and middle-east have very low happiness score.

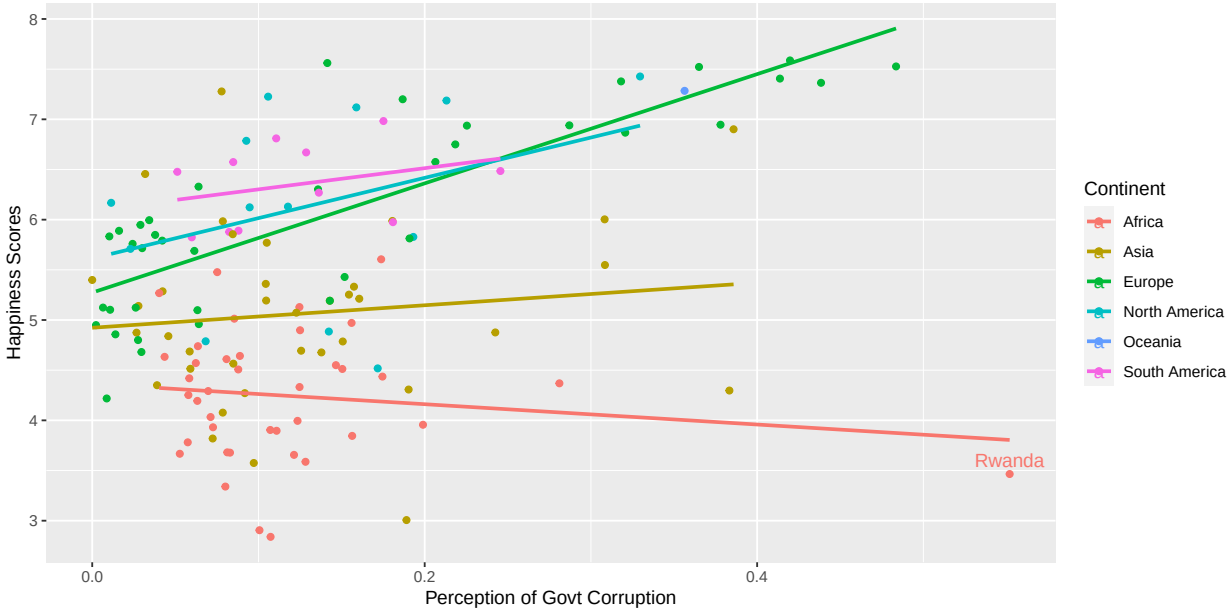


This is largely due to the wealth difference in these regions. Interestingly, Russia has relatively neutral score, even though it has a relatively high GDP.

Attributes like GDP, life expectancy and freedom of life choices show strong positive correlation to the happiness score as expected. Whereas, poverty index is negatively correlated to it. All plots shown are for 2015.

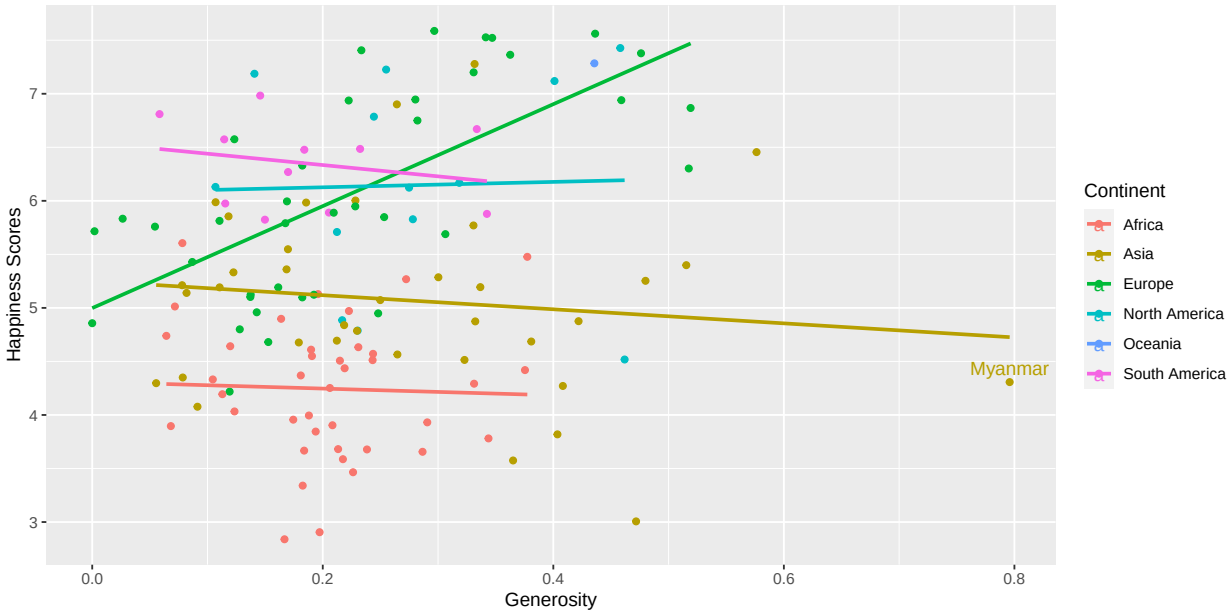


Interestingly, **perception of govt. corruption** that measures trust in govt. is positively correlated to scores for richer countries, but for Africa it shows exact opposite trend. We can see only one country, Rwanda, is influential enough to negate the correlation for Africa.



Regarding this, we researched that various governance indicators indicate that Rwanda performs relatively well in terms of control of corruption, compared to many African countries. The country has also achieved significant progress over the last years in terms of government effectiveness and transparency of the regulatory framework. In spite of these efforts, corruption remains prevalent in the country resulting in low happiness index.(source)

Again, for Africa, Asia and South America, it seems the more generous people are, the less happy they are. Notice that these continents consists mostly of third world countries.



In Asia,Myanmar is the influential outlier, with low happiness, and high generosity index. The reason was that Myanmar was undergoing a political transition during 2010s, with military dictatorship voluntarily giving way to a democratic government and the reform process culminated in 2015.This national disorder might be the reason for a low happiness score, while still topping global index of charitable behavior. It proves that generosity is not limited to the wealthy.(2)

Furthermore, disaster hit nations rank highly on some measures showing the resilience of human kindness.

Conclusions :

We have shown through our various data visualizations how the World Happiness Report has as many cons as it has pros. It does a good job at giving us a decent understanding of which countries/regions are leading in terms of happiness and which regions are struggling , and for the most part the final rankings are what one would expect them to be.

However , surveying only a couple thousand people to represent a country can be misleading. It might work to a large extent for countries with lower populations but for a country like India with a large and extremely diverse population , it is hard for us to imagine that the same sample space would work well in this case as well.

Some parameters included in the report are very abstract and hard to quantify and thus do not give as much of an understanding of Happiness Index that other parameters such as “GDP per capita” do. Since Happiness Index is a weighted average of the parameters present in the report, we also tried to see if any other factors not given in the report influenced happiness in a significant way and indeed we found a few factors that had a higher correlation with happiness than a few factors present in the report. Thus, if these factors were given relative values and used in the calculation for happiness index , it could possibly give us a more comprehensive view of happiness across the world.

References:

1) kaggle : https://www.kaggle.com/

2) worldpopulation : https://worldpopulationreview.com/

3) thehindu : https://www.thehindu.com/news/international/explained-what-is-the-world-happiness-report/article65240772.ece?homepage=true

4) pinterest : https://in.pinterest.com/

5) Continent : https://statisticstimes.com/geography/countries-by-continents.php

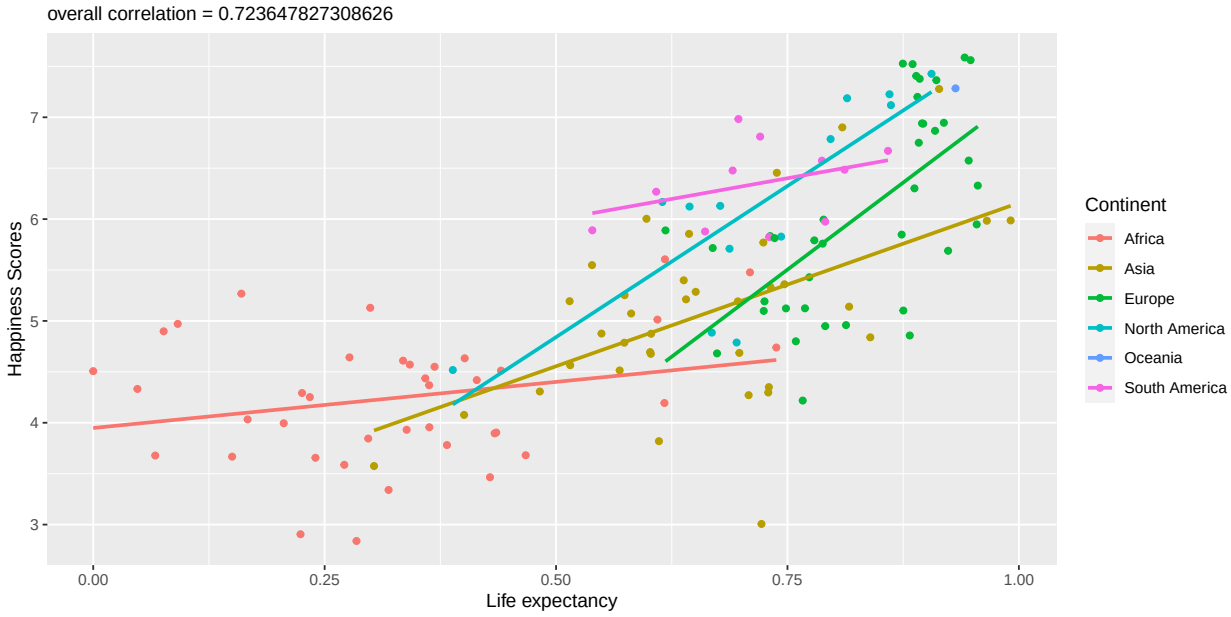
6) Literacy rate : https://en.wikipedia.org/wiki/List_of_countries_by_literacy_rate

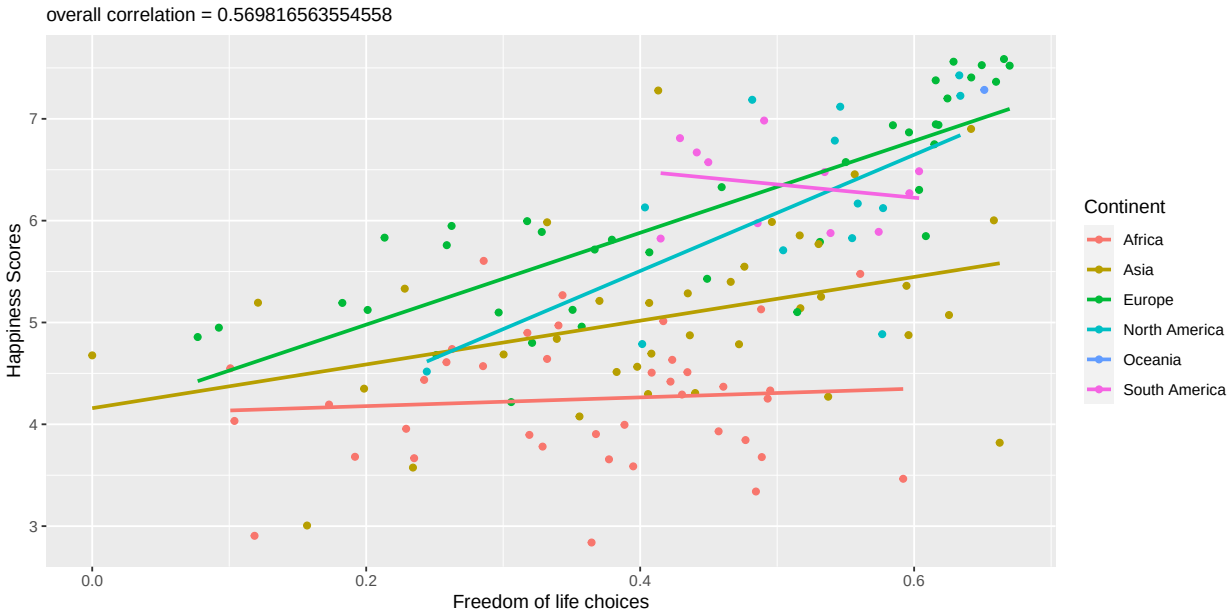
7) Poverty rate : https://www.indexmundi.com/g/r.aspx?v=69

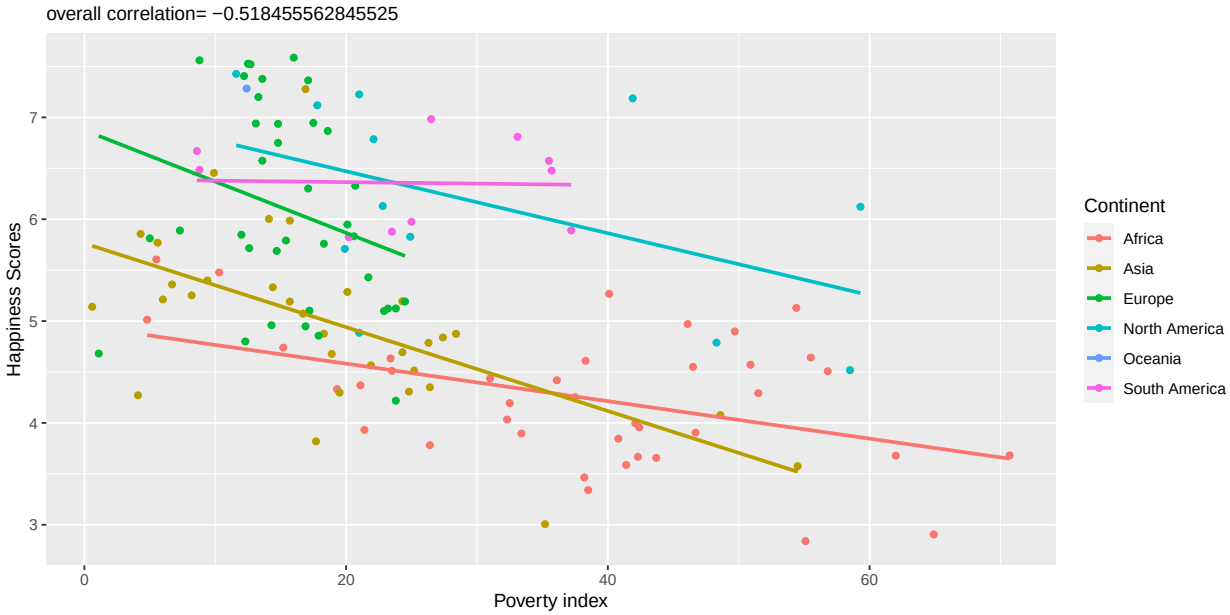
8) World Countries: http://tapiquen-sig.jimdo.com. "Shape downloaded from http://www.efrainmaps.es. Carlos Efra n Porto Tapiqu n. Geograf a, SIG y Cartograf a Digital. Valencia, Spain, 2020." Based on shapes from Enviromental Systems Research Institute (ESRI). Free Distribuition.

Appendix :

Some plots for the year 2015 for reference




overall correlation = 0.569816563554558
Happiness Scores
Freedom of life choices
0.0
0.2
0.4
0.6
3
4
5
6
7
Continent
Africa
Asia
Europe
North America
Oceania
South America


overall correlation= −0.518455562845525
Happiness Scores
Poverty index
0
20
40
60
3
4
5
6
7
Continent
Africa
Asia
Europe
North America
Oceania
South America