Happiness Score with 2019 data

Narendra Sahu

07/06/2020

Table of Contents

knitr::opts\_chunk$set(echo=FALSE, warning=FALSE, message=FALSE, error=FALSE)  
# Context

The World Happiness Report is a landmark survey of the state of global happiness. The first report was published in 2012, the second in 2013, the third in 2015, and the fourth in the 2016 Update. The World Happiness 2017, which ranks 155 countries by their happiness levels, was released at the United Nations at an event celebrating International Day of Happiness on March 20th. The report continues to gain global recognition as governments, organizations and civil society increasingly use happiness indicators to inform their policy-making decisions. Leading experts across fields – economics, psychology, survey analysis, national statistics, health, public policy and more – describe how measurements of well-being can be used effectively to assess the progress of nations. The reports review the state of happiness in the world today and show how the new science of happiness explains personal and national variations in happiness. There are three parts to my report as follows:

The happiness scores and rankings use data from the Gallup World Poll. The scores are based on answers to the main life evaluation question asked in the poll. This question, known as the Cantril ladder, asks respondents to think of 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 scores are from nationally representative samples for the years 2013-2016 and use the Gallup weights to make the estimates representative. The columns following the happiness score estimate the extent to which each of six factors – economic production, social support, life expectancy, freedom, absence of corruption, and generosity.

# Cleaning

Now we can load our dataset and see the structure of happiness variables. Our dataset is pretty clean, and we will implement a few adjustments to make it looks better.

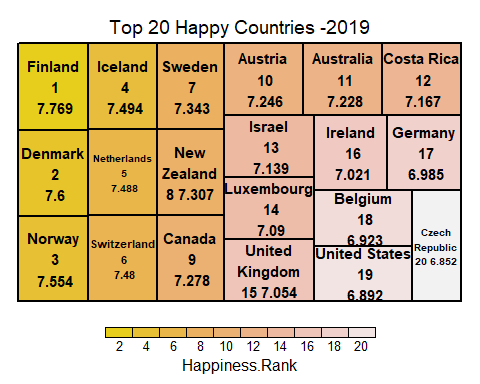
## 'data.frame': 156 obs. of 9 variables:  
## $ Overall.rank : int 1 2 3 4 5 6 7 8 9 10 ...  
## $ Country.or.region : Factor w/ 156 levels "Afghanistan",..: 44 37 106 58 99 134 133 100 24 7 ...  
## $ Score : num 7.77 7.6 7.55 7.49 7.49 ...  
## $ GDP.per.capita : num 1.34 1.38 1.49 1.38 1.4 ...  
## $ Social.support : num 1.59 1.57 1.58 1.62 1.52 ...  
## $ Healthy.life.expectancy : num 0.986 0.996 1.028 1.026 0.999 ...  
## $ Freedom.to.make.life.choices: num 0.596 0.592 0.603 0.591 0.557 0.572 0.574 0.585 0.584 0.532 ...  
## $ Generosity : num 0.153 0.252 0.271 0.354 0.322 0.263 0.267 0.33 0.285 0.244 ...  
## $ Perceptions.of.corruption : num 0.393 0.41 0.341 0.118 0.298 0.343 0.373 0.38 0.308 0.226 ...

Next we add another column to the dataset which is continent based on the countries given.

# Visualization

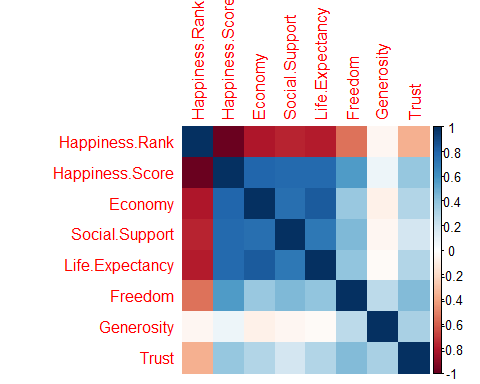
In this section, we will see different variables and find out how they correlate with each other and happiness score.

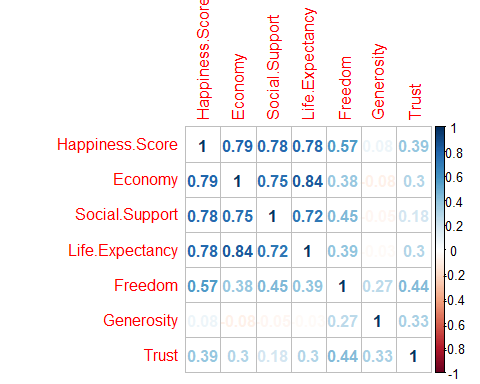
## Top 20 Happy countries



## Correlation plot

Lets see the correlation between numerical variables in our dataset.

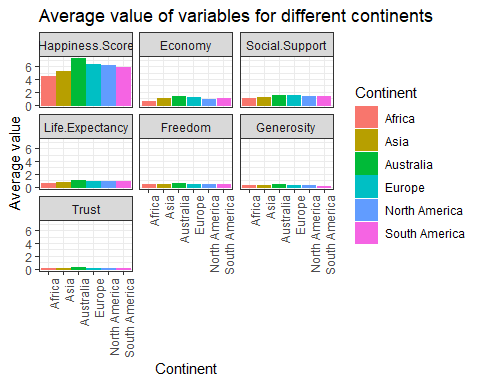




According to the above cor plot, Economy, life expectancy, and family or social life play the most significant role in contributing to happiness. Ggenerosity have the lowest impact on the happiness score. Trust or perception of corruption in a government is not highly correlated to happiness.

## Comparing different continents regarding their happiness variables

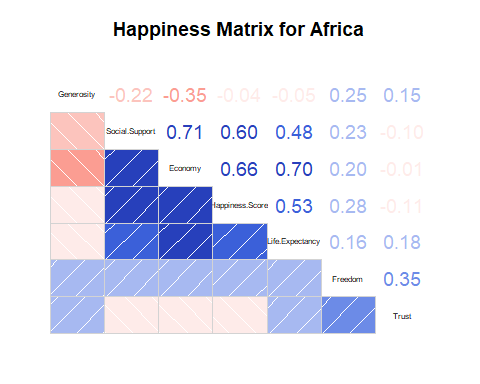
Lets calculate the average happiness score and the average of the other seven variables for each continent. Then melt it to have variables and values in separate columns. Finally, using ggplot to show the difference between continents.



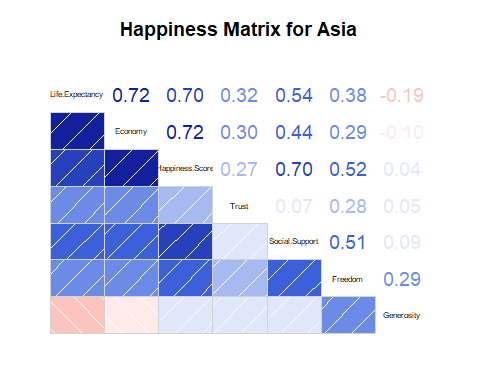
We can see that Australia has approximately the highest average in all fields after that Europe, North America, and South America are roughly the same regarding happiness score and the other seven factors. Finally, Asia and Africa have the lowest scores in all fields.

## Correlation plot for each continent

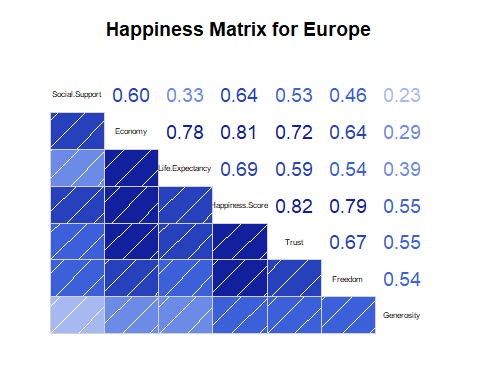
Lets see the correlation between variables for each continent.



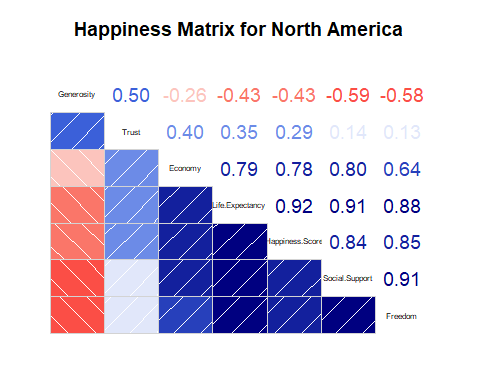
**Correlation between “Happiness Score” and the other variables in Africa:**  
Economy > Family or Social Support > Life Expectancy > Freedom  
There is no correlation between happiness score and trust as well as generosity.



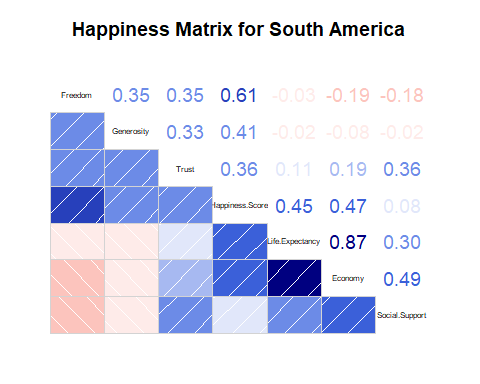
**Correlation between “Happiness Score” and the other variables in Asia:**  
Economy > Family or Social Support > Life Expectancy > Freedom > Trust There is no correlation between happiness score and generosity.



**Correlation between “Happiness Score” and the other variables in Europe:**  
Trust > Economy > Freedom > Life Expectancy > Family or Social Support > Generosity  
In Europe, Trust or perception of corruption is top factor in determining happiness, while in all other continets, this relation is insignificant.



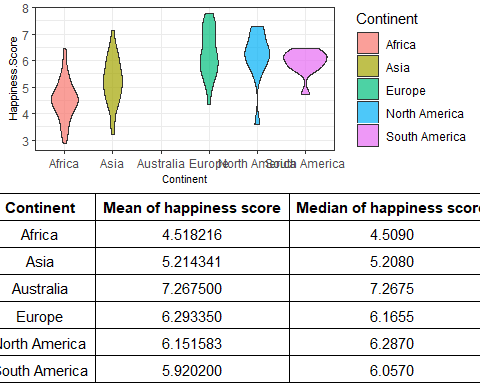
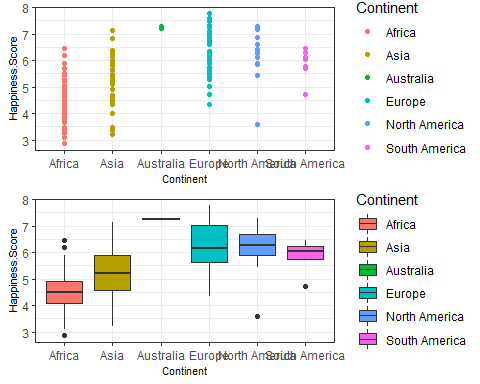
**Correlation between “Happiness Score” and the other variables in North America:**  
Life Expectancy > Freedom > Family or Social Support> Economy > Trust  
There is an inverse correlation between happiness score and generosity.



**Correlation between “Happiness Score” and the other variables in South America:**  
Freedom > Economy > Life Expectancy > Generosity > Trust The family or Social Support is the least significant factor in South America.

## Happiness score comparison on different continents

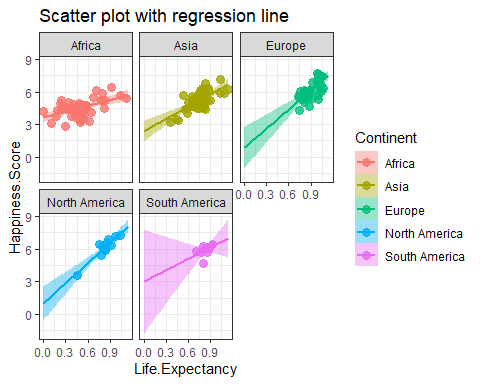
We will use scatter plot, box plot, and violin plot to see the happiness score distribution in different countries, how this score is populated in these continents and also will calculate the mean and median of happiness score for each of these continents.



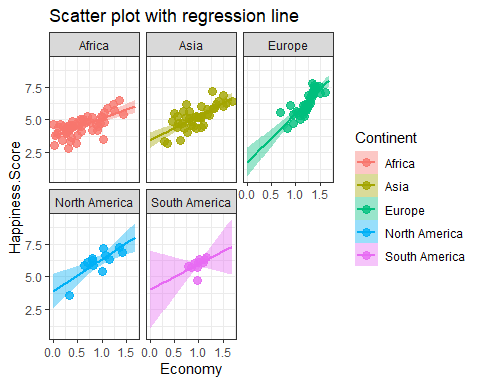
As we have seen before, Australia has the highest median happiness score. Europe, South America, and North America are in the second place regarding median happiness score. Asia has the lowest median after Africa. We can see the range of happiness score for different continents, and also the concentration of happiness score.

## Scatter plot with regression line

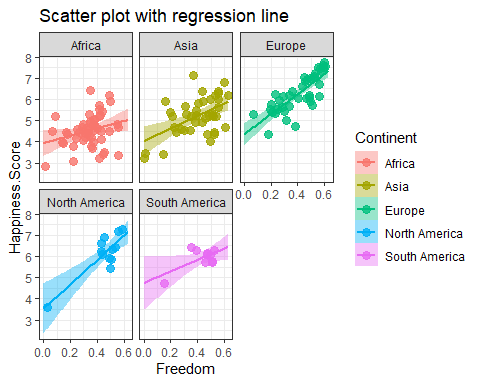
Lets see the correlation between happiness score and the other seven factors in the happiness dataset for different continents by creating a scatter plot.



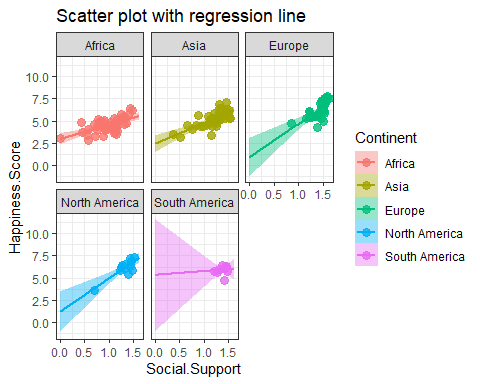
The correlation between life expectancy and happiness score in Europe, North America, and Asia is more significant than the other continents.



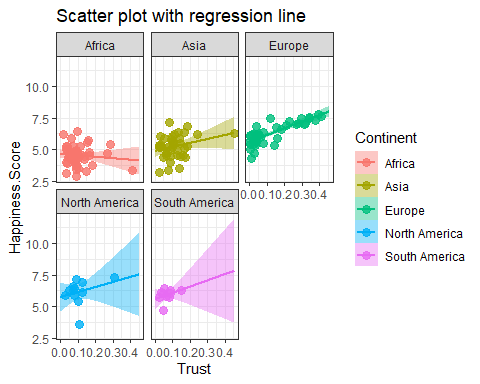
We can see pretty the same result here for the correlation between happiness score and economy. Africa has the lowest relationship in this regard.



Freedom in Europe and North America is more correlated to happiness score than any other continents.

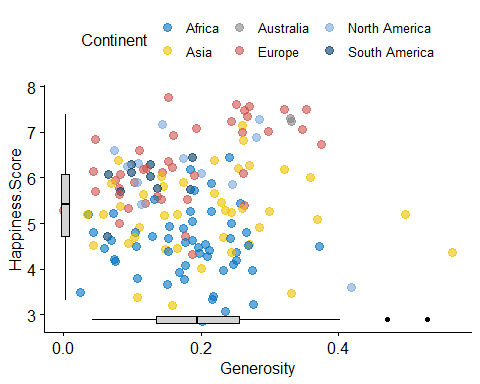


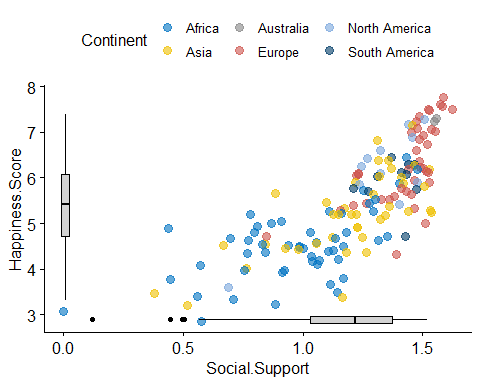
In South America with increase in the family or Social Support score, the happiness score remains constant.

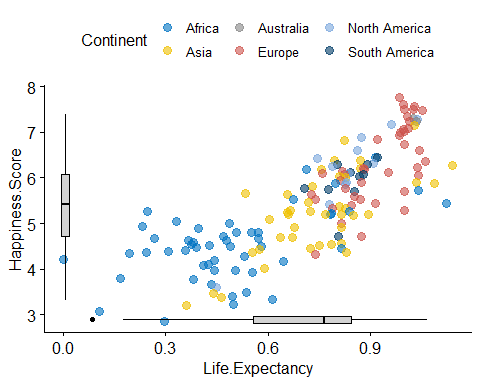


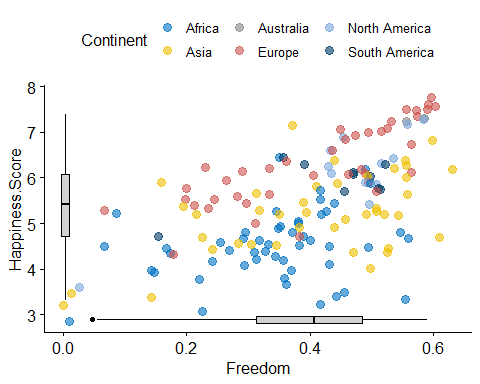
## Scatter plot colored by Continents

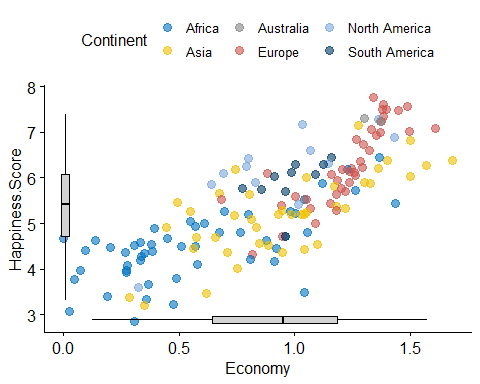
The following is just another way of seeing happiness score distribution on different continents when taking the correlation of happiness score with different variables into account.

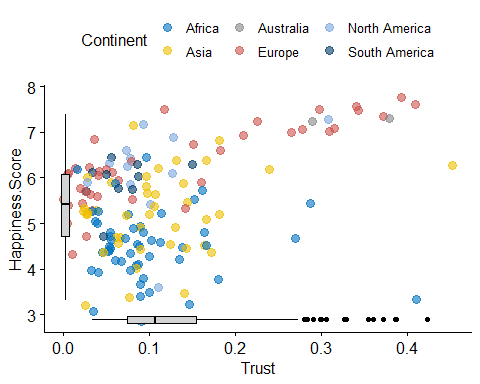




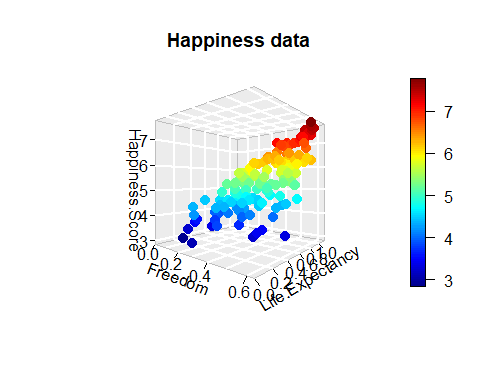




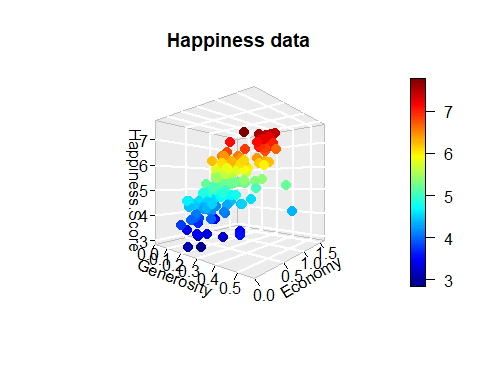




## 3D Plot



According to this plot, the higher the life expectancy and freedom scores, the higher will be the happiness score.



The higher economy score and the lower generosity score will lead to the higher level of happiness.

With an increase in the economy score and the happiness score, trust remains constant. This is the trend for happiness scores below 5. After this point, we can see that the impact of trust on happiness score increases gradually.