

A Study of Factors that Affect Happiness Scores Across the World

*Varun Agarwal, Sreya Chakrabarti, Arpitha Kashyap
Indiana University*

Abstract

The economics of happiness has been vastly studied since the beginning of the 20th century and various factors have been identified that help explain happiness index [1]. With an increase in social awareness and concern across the world, topics such as mental health and happiness of individuals have become important factors in various fields of research. For this study, we use 'The World Happiness' data set from Kaggle [2]. This report serves as a report to understand happiness across the world. The data set provides information about some economic and social factors which can estimate the happiness scores. The main goal of the project is to visually explore happiness scores across the world and to understand which factors play major role in altering happiness scores. Such an understanding can form a basis for informed and effective policy-making around the world. We plan to visually analyze the information provided in this dataset to form an understanding about the factors which have a high correlation with happiness scores across the world. Building over our visualizations, we would like to extend our dataset to incorporate more relevant socio-economic factors which can further provide a better estimate of happiness scores per country. We will also briefly discuss how happiness reports can be used in policy-making.

Introduction

Happiness reports have been studied in many research works to understand how they can relate to policy-making. For a long time, public policy has ignored factors that can be explained by findings in modern psychology. Policy makers in currently developed countries have mostly kept their focus on economic growth. Although, empirical studies through surveys have concluded that despite massive economic growths, happiness scores have barely changed [3]. In our opinion, social welfare and happiness can help make better decisions in policy-making because, after all, the policies are made to benefit the people. Even public health policies have been shown to largely affect happiness. [4] claims that "People's expectations for health standards influence their reported health and associated happiness". They also reiterate the fact that human wellness has been largely linked only to economic conditions, although, it can be explained better through health, family and happiness. This is not to say that economic conditions are not conducive to increase in happiness.

Only recently, studies in social welfare and happiness are being accepted by policy makers and economists. The World Happiness Reports, which first came out in 2012, are a good indicator of the adaptation of social welfare metrics, other than economic metrics, being considered in decision making process of public and health policies. We believe economic conditions only lead to better social welfare if they are accompanied by improvements in public health, family and community policies.

In our report, we discuss several visualization techniques, and how they can be applied to our data set to produce insightful results about the data. Our observations are based on the many informative visualization techniques that we have utilized.

Background and related work

The first world happiness report was something that drew international attention as it defined certain paradigms that related closely to the well-being and happiness of the nations. This report was released on April 1st, in the year 2012. This detailed report gives us an insight on the state of world happiness, causes of happiness as well as misery, and policy implications which are highlighted by case studies [5].

The second world happiness report was released in 2013, and since has been released on an annual basis. The year 2014 however, there was no report released. Experts in fields such as economy, survey analysis, national statistics and psychology talk about how progress of a nation can be assessed effectively by using certain factors or measurements of well-being.

A lot of recent work has been done upon the same subject which we wish to present an analysis upon. Economics of happiness has gained substantial support and interest with researchers, and is continually being researched for more practical applications.

[6] visually explores life satisfaction amongst people across the globe from 2005 to 2016. It presents a detailed discussion of correlation of happiness with income, life expectancy, mental health, and daily life events.

The report entails a brief discussion about the authenticity of happiness scores, and whether happiness can really be “measured”. [7] uses world happiness reports created by UN to construct an interesting world map with happiness scores on a quantitative scale.

Also, many users have worked upon the same dataset, and have posted their works on Kaggle. One such work [8], draws a thoughtful parallel between terror attacks and happiness scores, which can be used to further extend our analysis.

The listed works do provide good analysis on happiness scores and their correlation with socio-economic factors. Although, we feel that better visualization techniques, such as using better color palettes, can be applied to create more informative and accurate visualizations. We plan to build on the pre-existing works and further extend the analysis by using regression and classification techniques to predict happiness scores.

The factors or variables that are used for this report or dataset are those that are found in the research literature. They seem to explain differences in life evaluations on a national level. There are few factors like inequality or unemployment that are not present in the dataset since international data is unavailable for some countries.

The factors or variables are intended to illustrate correlation rather than the clean causal estimates. This is because some of the data which are drawn from the same survey sources are correlated with each other. They are also likely to be two-way relations between life evaluations and the chosen variables in some cases. For example, healthy people are overall happier, but happier people tend to be overall healthier [5].

Research goals and hypothesis

The major research goal is to visualize how various factors contribute to a nation's happiness scores, and whether these factors differ from region to region. We desire to go beyond the original

dataset that we obtained from Kaggle, to find more socio-economic factors which could help in building an informative and accurate predictive model for happiness indices across the world.

Another factor which interests us is to check whether the economic status of individuals in a country is most conducive to a higher happiness index. We hypothesize that the economic status of individuals might be a major factor in predicting the happiness index, however, it would be good to see otherwise. We also want to observe the changes in happiness index for countries over a period of three years (from 2015 to 2017) to evaluate the major causes of the change. Again, we want to go beyond our dataset to identify more socio-economic and cultural factors which might better predict the happiness index.

According to [9], luxury conveniences and contentment with lifestyle were also major factors during life evaluation. We feel that these factors would also be closely related to the happiness index. This would come into picture during a later part of our analysis.

We also try to measure happiness score among countries over time and across social backgrounds to avoid “happiness traps”. For example, the U.S. in recent decades is facing Government Corruption issues which declines life satisfaction. The larger perspective of the goal will be to promote a public dialogue around the index. This is to assure that people share their own feelings, understand and appreciate how they could increase happiness in their countries and hence contribute to the betterment of world happiness.

There is a reason behind why we use people’s actual life evaluations rather than some index of factors. This is to make sure people understand their importance and how they influence well-being by attaching fundamental importance to the evaluations that people make of their own lives. This provides a reality and power that no constructed index could ever have given them. For this kind of a report that strives for objectivity, it is important that the rankings are dependent on the data collected from population-based samples of individuals. It is not based on what we think might influence the quality of people’s lives [5].

The average scores simply reflect what individual respondents report to the Gallup World Poll surveyors [5]. Since our data is extracted from population-based samples in each country, we can present confidence regions for our estimates. This provides a way to see if the rankings are based on differences big enough to be statistically meaningful.

Happiness is increasingly found to be the actual measure of social progress and the goal of public policy. Our research considers how and why life evaluations are used as our central measure of subjective well-being among nations. This will be followed by our efforts to explain the differences in national average evaluations, across countries and over time.

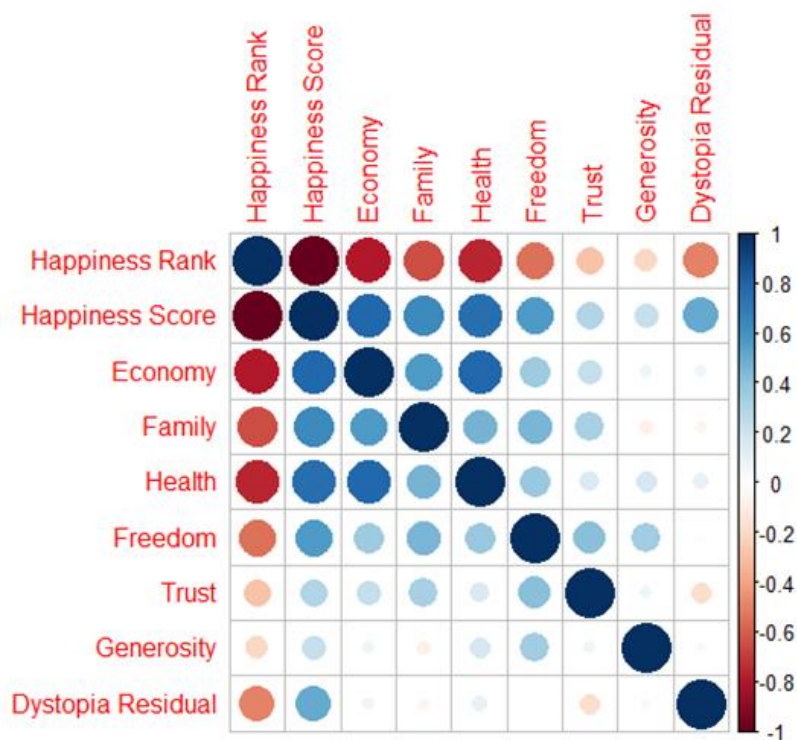
Process

To assess the relation of health and socio-economic factors with public happiness, we obtained a data set that presents the happiness ratings for about 155 countries, for a period of 3 years (2015 – 2017). The dataset uses data from the Gallup World Poll to inform the happiness scores and rankings. Happiness scores and rankings are based on the participant’s response to the life evaluation question in the poll. The question is based on the Cantril Self-Anchoring Striving Scale, as mentioned in [10]. The happiness scores in this dataset are compared against a hypothetical country called Dystopia, which is the baseline for happiness scores of all other countries. Dystopia

scores are computed by taking the lowest average scores of all the factors for all the countries during the period of 2014 to 2016 [5]. The happiness scores in our data set is explained by the sum of seven factors in our data set - GDP per capita, Social Support/Family, Health life expectancy at birth, freedom to make life choices, generosity, and trust on government (absence of corruption). Please note that the data set we have obtained from Kaggle does not present the time series data itself, rather it presents the results after regression on the factors below. Each of these factors have been measured as following, as per [5]:

- **GDP Per Capita** - It is measured in term of purchasing power parity, taken from the World Development Indicators released by the World Bank in August 2016. Note that GDP data for 2016 has been extended from 2015 to 2016 using country-specific forecasts, because its data was not available during the time of the World Happiness Report in 2017.
- **Health Life Expectancy** - It is constructed based on the data from World Health Organization (WHO). A time series for health life expectancy at birth was constructed to produce this data.
- **Social Support or Family Indicator** - It is computed by taking the average of the binary responses to the Gallup World Poll (GWP) question, "If you were in trouble, do you have relatives or friends you can count on to help you whenever you need them, or not?"
- **Freedom** - This is seen as the freedom to make life choices. It is measured by the GWP question "Are you satisfied or dissatisfied with your freedom to choose what you do with your life?"
- **Generosity** - It is measured on the basis of the GWP question, "Have you donated money to a charity in the past month?"
- **Perception of Corruption** - It is measured on the basis of two GWP questions "Is corruption widespread throughout the government or not?" and "Is corruption widespread within businesses or not?"
- **Dystopia Residual** - This is the Dystopia Happiness Score (1.85) + the Residual value, which is the value that remains unexplained by the explanatory factors

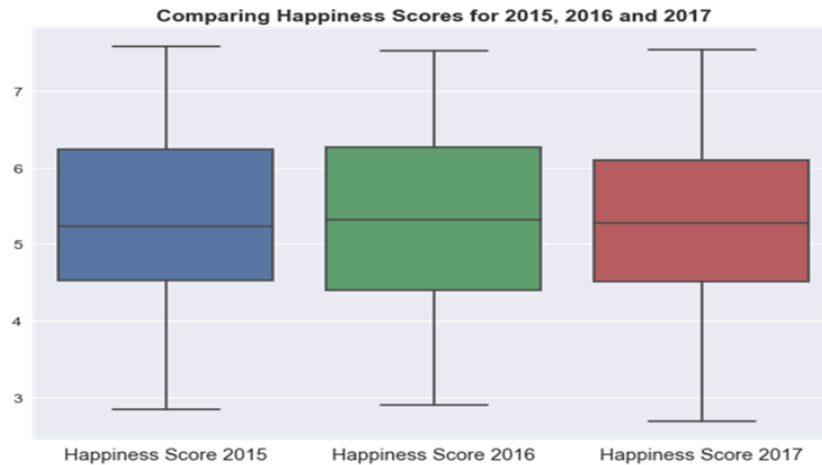
We have done a preliminary analysis on our data to better understand its structure and explore inter-dependencies. This analysis also serves as the basis for our results and insights that follow. To check for the correlation within our data, we have used the `corrplot()` in R. This is a better method than simply presenting the correlation values between variables, because of the ease of readability, and thus obtaining inferences from the chart.



It is clearly seen that Economy and health indices are very highly correlated with Happiness Score. This meets our assumptions and observations made from existing literature, that economy and health are the major contributors to Happiness index. We also observe that social support and freedom of choice of life are also highly correlated with happiness scores. Although, we should beware of concluding anything from the correlation plot, because it is widely known that correlation does not imply causation. This is why we need to further delve into better analysis and graphs to provide conclusive evidence.

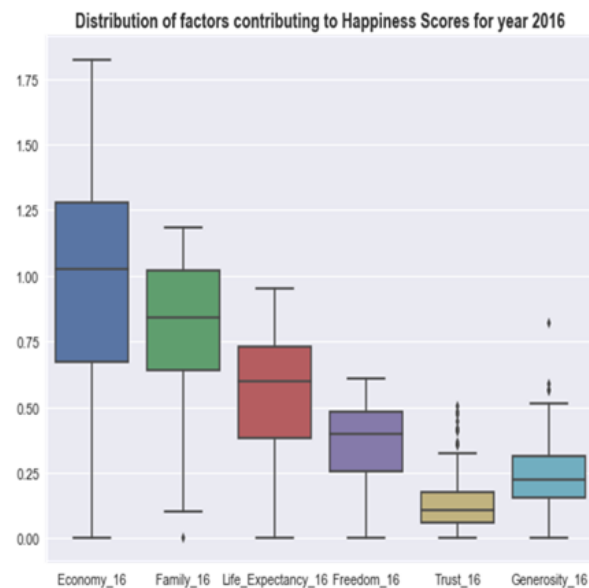
It is also interesting to see why trust on government and generosity indices have a comparatively lower correlation with happiness scores. Do people not consider corruption in business and government and how generous they feel as happiness motivators? We will revisit this question later.

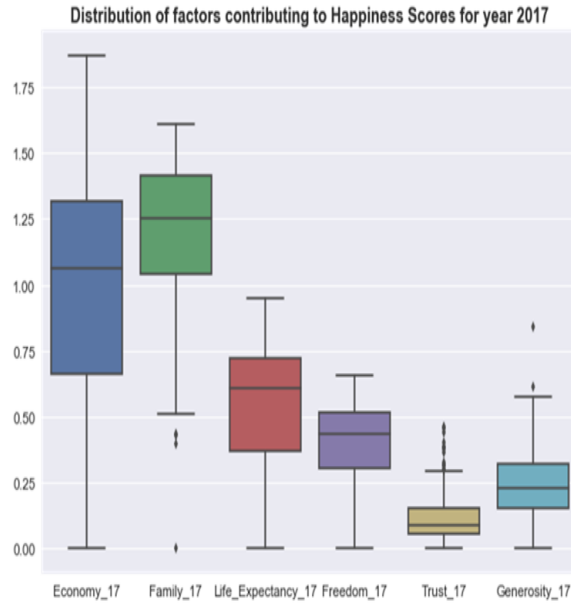
We now look at the distribution of happiness scores around the world by using boxplots. A clear indication of quartile values gives us a view of how happiness is divided across the countries. Having clear indication of median happiness scores for all the years can provide us insights of happiness scores within a year, and also allow us to compare the distribution of happiness scores between the years.



From the boxplots above, we can see that the distribution between the three years has not changed a lot. Although, we can see that the median happiness score for 2016 is higher than that for 2015 and 2017. For the year 2017, we can see that the values for the inter-quartile range is much shorter than that of 2015 and 2016.

To further understand our data in depth, let us draw boxplots to understand the distribution of factors contributing to happiness scores for each year. These boxplots will help us understand how much a particular factor contributed to happiness scores, and will also allow comparison between the factors. This can give us an inference of which factor had the most contribution towards happiness.





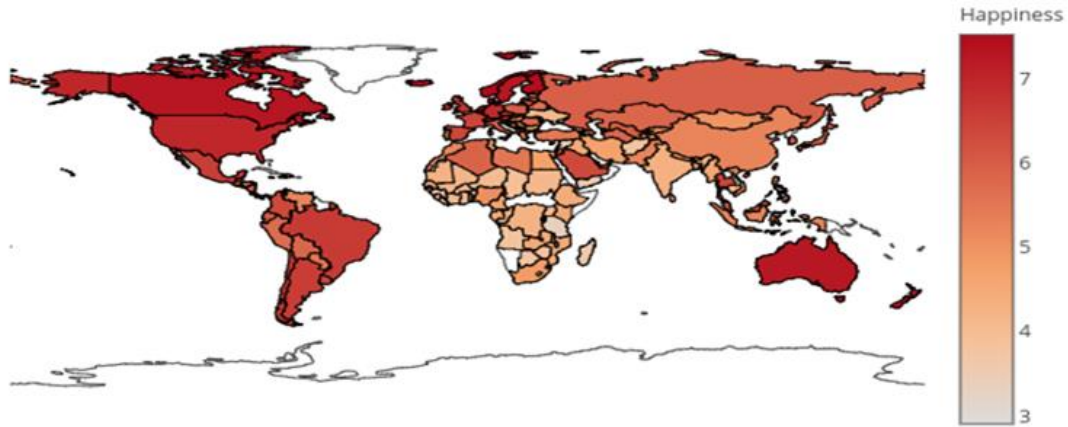
For the boxplots of year 2015, we can see that Family index was higher than every other factor, and explains the majority of the happiness score, too. Economy seems to have the second highest median value, but also has a higher range of values. Trust seems to have the lowest range of values, so it does not seem that the perception of corruption in government and business explains a lot of happiness scores. This could probably be because generally we observe polar views about the government (for and opposition), due to which its contribution is minimal.

Results

A preliminary analysis of the data has given us a good amount of insight to further explore upon. In order to have a visual sense of happiness around the globe, we turn to plotting a world map with happiness scores described by a quantitative color bar.

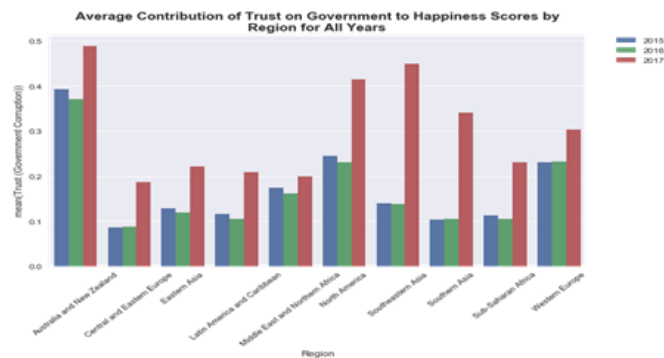
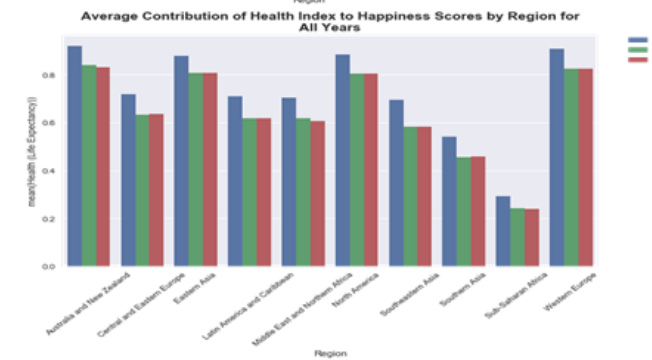
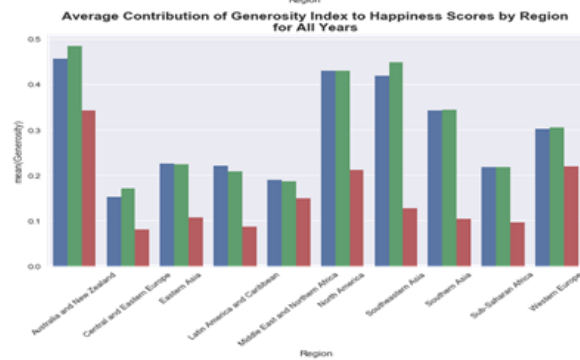
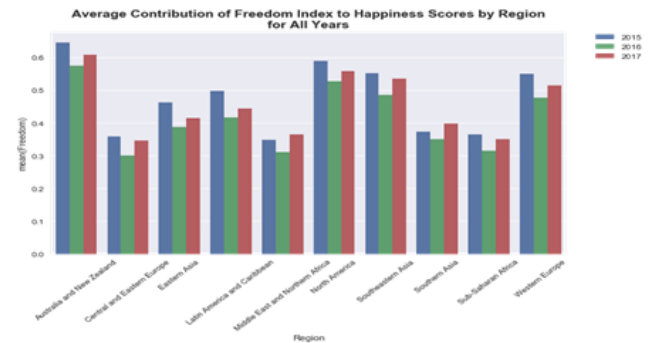
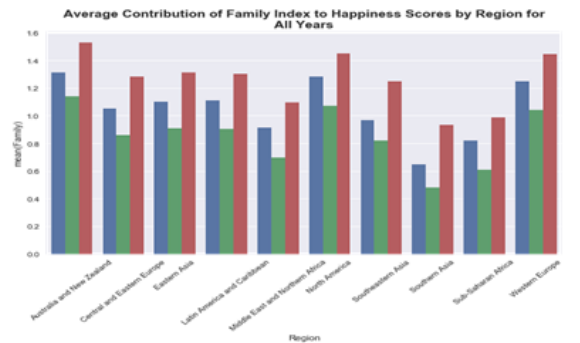
Happiness Scores Across the Globe (2017)

Global Happiness



The world map above provides a plethora of information about the happiness scores around the globe. Regions such as North America, Australia and New Zealand, and the Scandinavian countries seem to trump every other region in terms of their happiness scores. We have known very well that Scandinavian countries have very peaceful international policies and do not indulge in any kind of international agitation. Moreover, their public policies have been widely considered to be the best, which partly has to do with their smaller population. Asia seems to have moderate happiness scores, which is likely due to the fact that most Asiatic countries are still developing with respect to both social and economic factors. Africa seems to score lowest among all the regions. This indicates major under-development in African countries, in terms of their living(poverty), health issues, economy and more. We have widely observed in news the turmoil in African countries due to rebellious groups and corrupt authorities. These could be some of the major reasons for the lowest happiness scores across the world.

To further validate our observations and speculations, and to understand region-wise differences in happiness scores, we plot separate bar charts for each factor contributing to happiness. The bar charts show average region-wise scores of each factor for all the three years. This can help us identify, by comparison, the factors that stand out from those of other regions. Inclusion of the yearly factor further allows us to see if these factors differ in their explanation of happiness scores across years.



The bar charts above clearly show that Sub-Saharan African region constantly has the lowest average happiness scores across all the regions in the world. This also clarifies that Africa should be further divided into the Middle East & Northern Africa, and Sub-Saharan Africa, as both of these regions seem to have different average values. Sub-Saharan Africa seems to score the lowest in all the factors except for generosity and trust on government indices. It is surprising to see that Central and Eastern Europe scores lower than Sub-Saharan Africa in terms of trust on government and generosity. We can conjecture that the distribution is skewed because of the

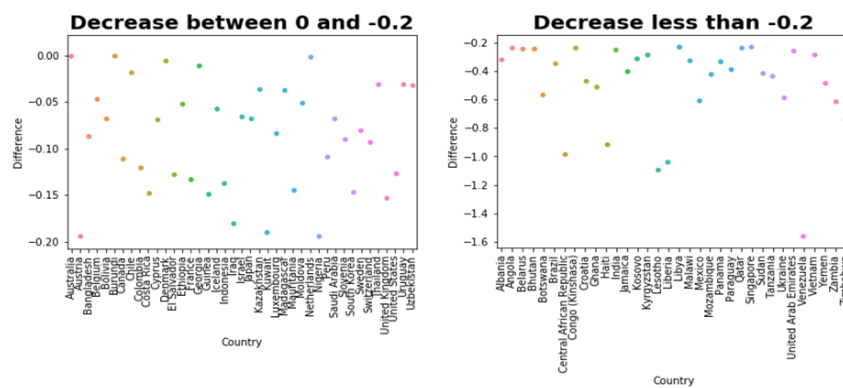
government crisis in Ukraine, but it still is open to debate. On the other hand, Western Europe and Australia & New Zealand seem to have higher averages than all other regions, which is what we had speculated. It is also interesting to see that the trust on government seems to have a higher contribution in explaining happiness scores for all regions in 2017, as compared to 2015 and 2016.

We further expand our analysis by dividing countries into two categories -

1. Countries that saw a decrease in happiness scores from 2015 to 2017, and
2. Countries that saw an increase in happiness scores from 2015 to 2017.

We do so to get a view of the extent of increase or decrease of happiness scores in two years.

Decrease of Happiness Score from 2015 to 2017

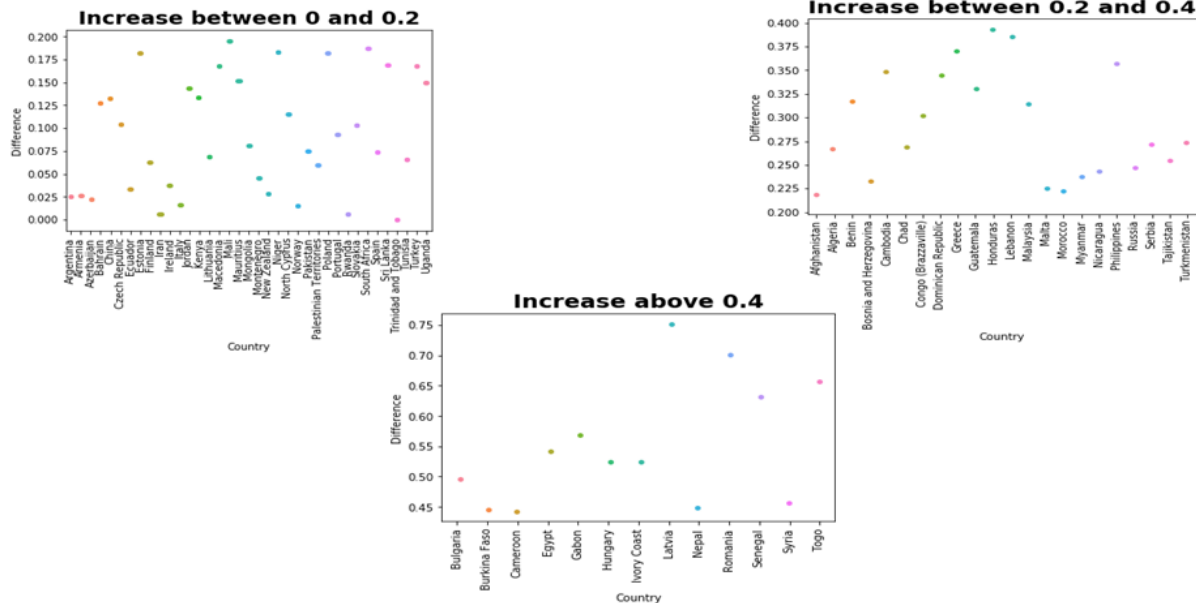


In the plots above, we can see the decrease of happiness score among countries from 2015 to 2017. Countries showing no difference in value indicates that the Happiness score has been approximately constant in the last 2 years. However, countries that have a negative difference indicate drop in happiness score from 2015 into 2017.

From the scatter plots above, we can see that there are more countries that saw a decrease in happiness score between -0.2 and 0. There are much lesser countries that saw a more drastic decrease in happiness scores, which are lesser than -0.2. This gives us an insight into the range of decrease in happiness scores. A country with negative happiness score is more likely to be between 0 and -0.2 rather than being less than -0.2, which gives us the inference that happiness scores are not likely to change drastically. This is expected because results of even drastically good/bad public policies are generally not visible until after a few years.

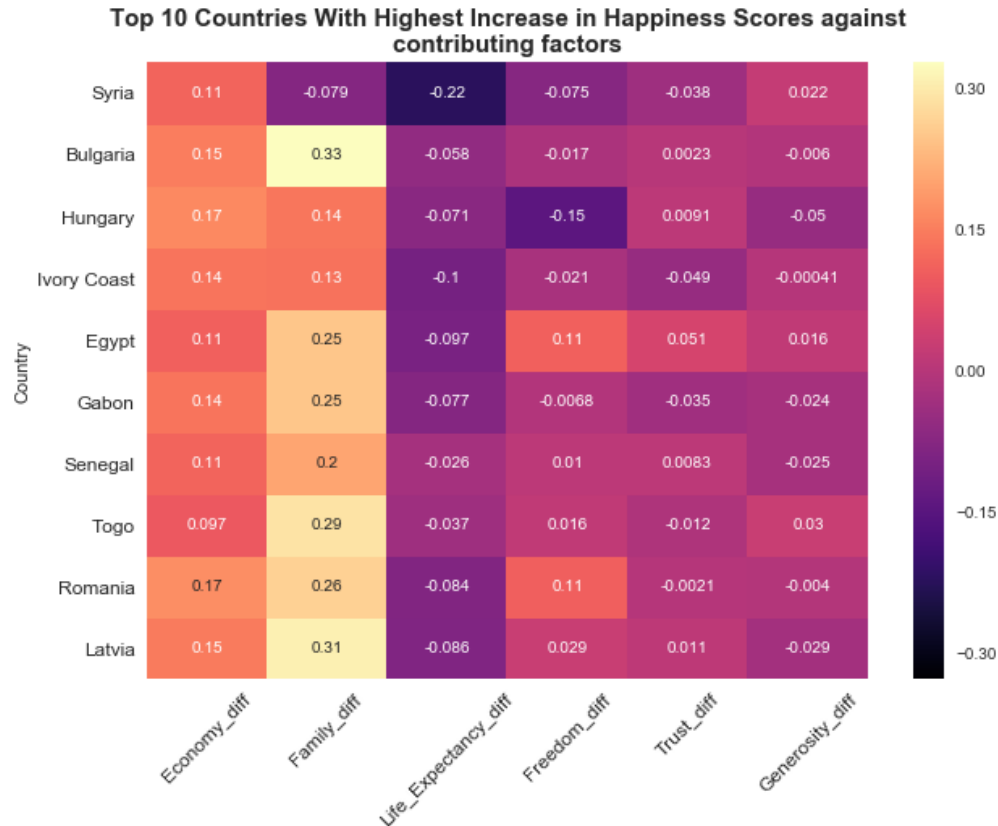
From the plot we can see that the country having the maximum decrease in Happiness score is Venezuela, which seems to almost be an outlier with a decrease in happiness by almost 1.6 points. This is a major decrease which we plan to elaborate on in future.

Increase of Happiness Score from 2015 to 2017

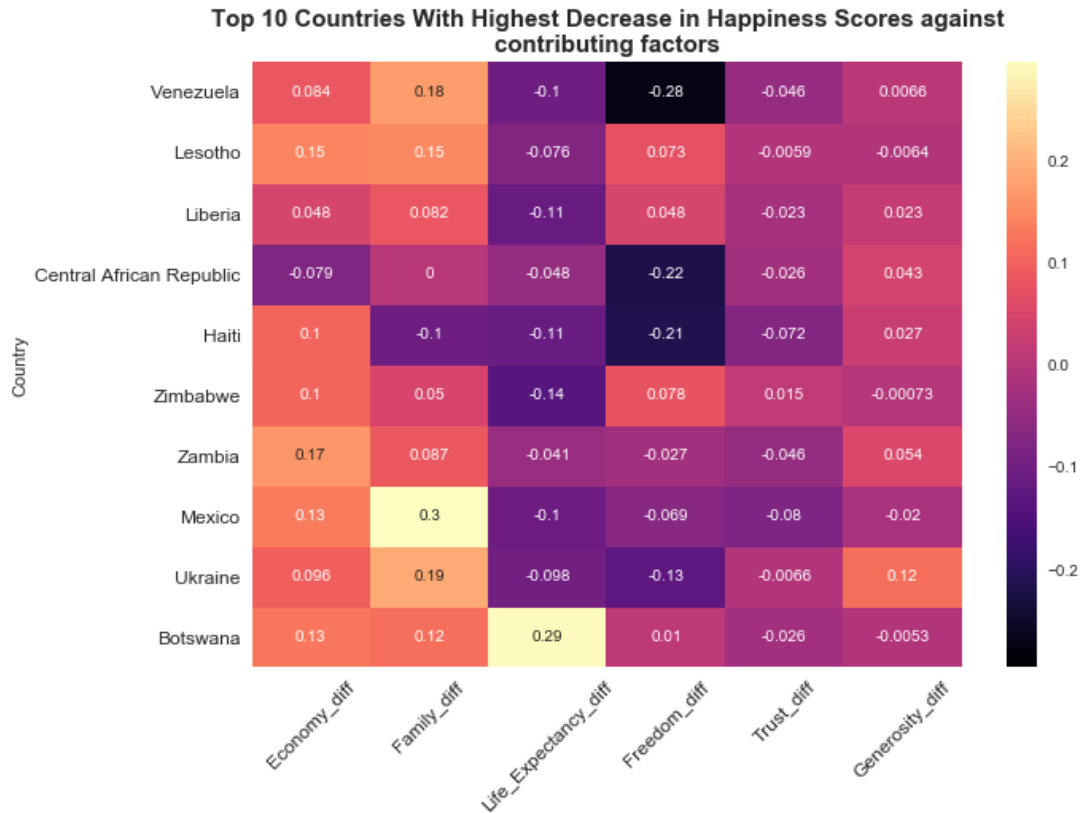


Comparing the number of countries which had an increase in happiness score with the number of countries that had a decrease in happiness score from 2015 to 2017, In the above scatter plots, they seem to have almost the same number of points. In fact, the total number of countries who saw a decrease in happiness scores were 75, while the number of countries who saw an increase in happiness scores were 74. It is surprising how there are almost equal number of countries with increase and decrease in happiness. Is the world somehow keeping its balance? It is a question that is outside the scope of this project, but is still quite interesting. The plots above seem to have a similar distribution to those for the countries who saw a decrease in happiness scores. Most countries do not see major increase in happiness scores, except for a few like Latvia, Romania and Togo, who saw an increase in happiness scores by over 0.65 points.

Having seen how the happiness score has varied for all the countries from 2015 to 2017, we need to further see how our analysis can help uncover factors which lead to increase or decrease in happiness scores. To do this, we focus our attention on the top ten countries who saw the most amount of increase to see what factors led to this increase. Similarly, we pick the top ten countries who saw the most amount of decrease in happiness scores to uncover the factors which saw the most amount of downgrade. To explore these factors, we have drawn heatmaps for these countries to see if we can see a pattern in there.



The heatmap above expands on the contributing factors of the 10 countries with highest increase in happiness scores. From the graph above we can see which factors had the highest contribution in increasing the happiness scores for those countries. From this heatmap, it can be clearly inferred that difference in economy and family indices seem to have increased substantially for all the countries in this chart. Other factors show a very minimal amount of negative value, which can be ignored without a loss of information.



The above heatmap shows the decrease in contributing factors for the top 10 countries with the highest decrease in happiness scores from 2015 to 2017. From the graph, we can infer that life expectancy has seen a decrease for all the countries here, except Botswana which saw quite a substantial increase in the life expectancy index. Even the freedom quotient has decreased substantially for Venezuela, Central African Republic and Haiti. A very striking inference from this graph is that even though the economic and family related indices seem to have shown an increase for almost all the countries in the above graph, they still seem to have the most amount of decrease in happiness scores from 2015 to 2017. This seems to validate the existing work in (*Happiness and Public Policy, Layard R, 2006*) which implies that increase in economic conditions alone is not sufficient in increasing happiness scores. We further plan to expand our analysis to see if these trends differ from region to region. We think that the trends would be different because of different perception of life and importance of various factors across different regions of the world.

Conclusion and Future Work

Based on the analysis done so far, we can say that economy is the most dominant influencer in happiness Score. Although, it is not the sole contributor to an increase in economy. The next most significant factor is health. Health index is also a prominent contributor to Happiness score. Factors like trust in government, and generosity indicators have a lesser impact on happiness score, and do not tend to contribute much to an improvement in happiness score.

For our future work, we plan to analyze the factors further to identify if there is any specific factor that causes decrease in happiness score. We plan to do so by grouping on regions, or a cluster

similar to region, so that we can group factors based on geographic, socio-economic and political factors. On getting such factors for various countries we plan on linking the data with world news from 2015 to 2017 about major changes in public and health policies, and also major news events. This way we can connect the factors causing the decrease in Happiness score for a country and can also represent the major events supporting the decreased value in a contributing factor. Thereby, we will be able to make visualizations showing the factor causing decrease in Happiness score and also the event which caused the decrease in the value of that factor.

Apart from the factors present in the data (Economy, Family, Health, Freedom, Trust and Generosity) we believe that there are more socio-economic factors which also have an influence in explaining the happiness score. We plan on finding more such data that can be factors influencing the happiness score of a country. Some factors that could possibly be explanatory factors for happiness scores in the western countries are luxury convenience of a country and contentment with lifestyle. These two factors should be major influencers in the happiness score of developed western country. The aim is to research further on these factors, and add new socio-economic factors that would help us in our evaluations.

References

- [1] C. Graham, "happiness, economics of," in *The New Palgrave Dictionary of Economics*, S. N. Durlauf and L. E. Blume, Eds. Basingstoke: Palgrave Macmillan, 2008.
- [2] "World Happiness Report | Kaggle." [Online]. Available: <https://www.kaggle.com/unsdsn/world-happiness>. [Accessed: 19-Sep-2017].
- [3] R. Layard, "Happiness and Public Policy: a Challenge to the Profession*," *Econ. J.*, vol. 116, no. 510, pp. C24–C33, Mar. 2006.
- [4] C. Graham, "Happiness and health: lessons--and questions--for public policy.," *Health Aff. (Millwood)*, vol. 27, no. 1, pp. 72–87, Jan. 2008.
- [5] J. Helliwell, R. Layard, and J. Sachs, "WORLD HAPPINESS REPORT 2017."
- [6] E. Ortiz-Ospina and M. Roser, "Happiness and Life Satisfaction - Our World in Data." [Online]. Available: <https://ourworldindata.org/happiness-and-life-satisfaction/>. [Accessed: 19-Sep-2017].
- [7] "Data visualization uncovers world's happiest countries | Yellowfin Business Intelligence." [Online]. Available: <https://www.yellowfinbi.com/blog/2015/08/yfcommunitynews-data-visualization-uncovers-world-s-happiest-countries-203209>. [Accessed: 19-Sep-2017].
- [8] V. Krishna, "Happiness Index and Terrorism! | Kaggle." [Online]. Available: <https://www.kaggle.com/vamsikrishna/happiness-index-and-terrorism>. [Accessed: 19-Sep-2017].
- [9] E. Diener, W. Ng, J. Harter, and R. Arora, "Wealth and happiness across the world: Material prosperity predicts life evaluation, whereas psychosocial prosperity predicts positive feeling.," *J. Pers. Soc. Psychol.*, vol. 99, no. 1, pp. 52–61, Jul. 2010.
- [10] W. Glatzer and J. Gulyas, "Cantril Self-Anchoring Striving Scale," in *Encyclopedia of Quality of Life and Well-Being Research*, A. C. Michalos, Ed. Dordrecht: Springer Netherlands, 2014, pp. 509–511.