Final Report GLBL 550
Nithya Srinivasan and Sam Maniscalco
Prof. Casey King

# Happiness and Inequality

## I. Project Overview, Motivation, and Brief Literature Review

Our project aims to understand how inequality levels and happiness levels have evolved in different countries over the last few years and see whether there are any similarities in patterns that emerge in changes in inequality and happiness levels over time. Our initial hypothesis was that happiness may depend on relative rather than absolute levels of deprivation and we sought to use data visualization and descriptive tools from Python libraries (such as Seaborn, Matplotlib, and GeoPandas) to be able to observe time trends.

We approached the problem in two ways; first, we created several choropleth maps of the world in order to visualize the distribution of different measures of inequality across the globe, in addition to visualizing how the distribution of happiness has changed over time. Second, we focused on 7 countries— US, UK, China, India, Ethiopia, Bhutan, and Norway—from different levels of development and different regions of the world in order to see how measures of inequality and happiness have varied over time within countries, and to observe any trends within certain countries regarding happiness and levels of inequality.

A brief literature review on these topics has indicated that this is an important and developing topic within international development. The theme of the 2019 UN Human Development Report, for example, was inequality.[1] The role of inequality in determining the well-being of individuals across the globe is gaining increasing attention, particularly as the globalized economy allows for more extreme disparity between the ultra-rich and the ultra-poor in societies across the globe.[2] Further, while the World Happiness Report has only been published since 2015, development economists and researchers alike are placing increasing weight towards individual's own conception of happiness when determining and targeting their development programs.[3] That is to say, there is increasing research at the nexus between inequality and happiness, and thus it will be illuminating to explore the descriptive similarities in distribution and time trends of happiness and inequality.

## II. Questions

*Examining patterns in happiness and inequality across countries and over time*

---

[1] "Human Development Report 2019." UNDP. Accessed May 1, 2020. http://report.hdr.undp.org/.

[2] Kelley, Jonathan, and M. D. R. Evans. "Societal Inequality and Individual Subjective Well-Being: Results from 68 Societies and over 200,000 Individuals, 1981–2008." *Social Science Research* 62 (2017/02/01/ 2017): 1-23. https://doi.org/https://doi.org/10.1016/j.ssresearch.2016.04.020.

[3] Oishi, Shigehiro, Selin Kesebir, and Ed Diener. "Income Inequality and Happiness." *Psychological Science* 22, no. 9 (2011): 1095-100. Accessed May 7, 2020. www.jstor.org/stable/41319994.

In particular, we explored the following sub-questions:

- How have happiness scores of the world changed over time? Which are the regions of the world with the highest and lowest happiness in the period 2015-18?
- Which were the regions with the highest inequality in 2018? How does this differ for different types of inequality-human inequality, gender inequality, education inequality, income inequality etc?
- How has happiness and different types of inequality changed over time for the 7 selected countries? Do any patterns emerge?
- What are the sub-components of the happiness score and which have the strongest correlations with the overall score? Are these relationships ratified by data collected from a different source?

## III. Data Source and Cleaning

In order to test our hypotheses regarding the relationship between happiness and inequality, we decided to turn to the UN's open source data repositories. We found our primary dataset from the United Nations Development Program (UNDP) Human Development Report. Each year, the UNDP publishes a Human Development Report "as independent, analytically and empirically grounded discussions of major development issues, trends and policies" in human development. The UN publishes all the figures and data collected in making the report online at http://hdr.undp.org, where one can search specific statistics, then download a csv file that contains the values for each country in that specific statistic over time, generally beginning in 1990, when the first HDR report was published. Coming from the UN, we figured the data would be reasonably reliable, given the UN's interest to track statistics regarding inequality and development over time empirically and independently.

We utilized a supplementary dataset, similarly published by the UN, in order to approximate the measures of happiness over time. Annually since 2015, the UN Sustainable Development Solutions Network has published a World Happiness Report, which ranks 156 countries by how happy their citizens perceive themselves to be. The report is researched and written by independent experts, separate from the UN, and utilizes self-reported surveys in partnership with Gallup in order to calculate happiness indices, which are published along with index scores in measures like GDP per capita, social support, and other factors. We downloaded these datasets from Kaggle,[4] where csv files were located for each year that we could then import into Python.

One of the initial problems in working with the data came when trying to download the GeoPandas package. For reasons related to dependencies, we ran into a bunch of errors when trying to import GeoPandas, and once it finally downloaded, it was missing some of the functionalities it was supposed to have. As a result, we learned about conda environments, and we created a new conda environment, and uploaded all necessary packages, including GeoPandas to the new environment.

[4] https://www.kaggle.com/unsdsn/world-happiness#2019.csv

In terms of data cleaning, utilizing the read_csv command efficiently, merging datasets, and type casting variables correctly were the most common problems we ran into. When downloaded, the UN csv files had lots of inconsistencies and were set up in a way difficult for Python to load into a dataframe. Utilizing the na_values argument to replace incomplete data with NA values proved useful and made the data much easier to work with.

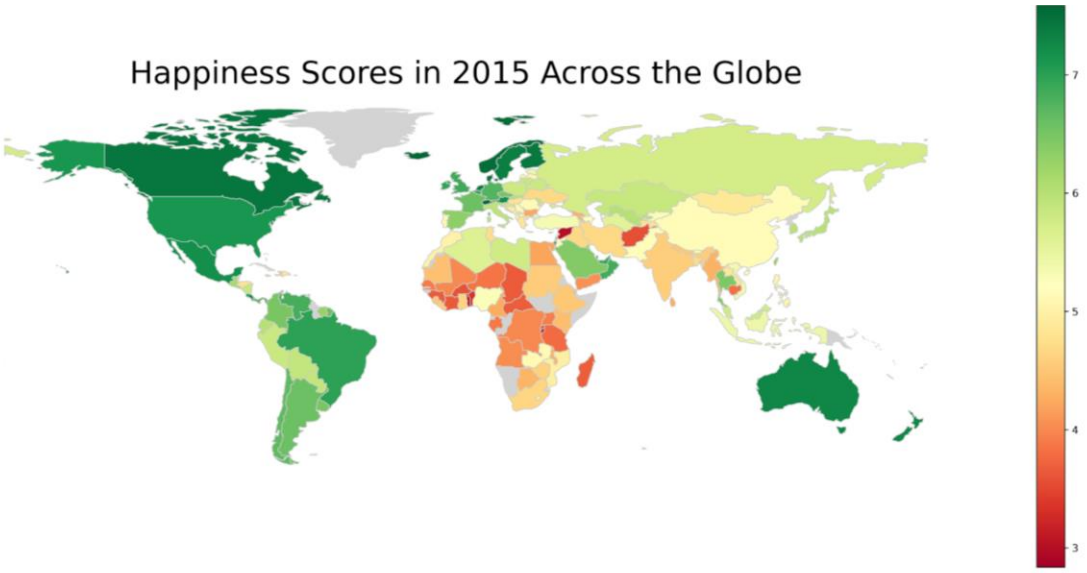
When merging datasets, we ran into several problems with country labels not matching. For example, certain files included the full name of every country, while others utilized non-uniform abbreviations so that a country like the Republic of Korea was labeled as "Rep. of Korea", "South Korea" and "S. Korea" in addition to "Republic of Korea" in the datasets we were working with. This required renaming certain countries such that all indices would match when merging datasets and creating the data frame that we would map, which needed to contain the coordinates of the countries as well as the data that would be displayed. Finally, when visualizing the data, we realized many of the quantitative variables were interpreted by Python as strings and not as floats or integers, and thus we had to typecast many of the variables into floats, so that Python would display them on a quantitative scale.

## IV. Figures and Discussion

CROSS COUNTRY ANALYSIS

### 1. Happiness

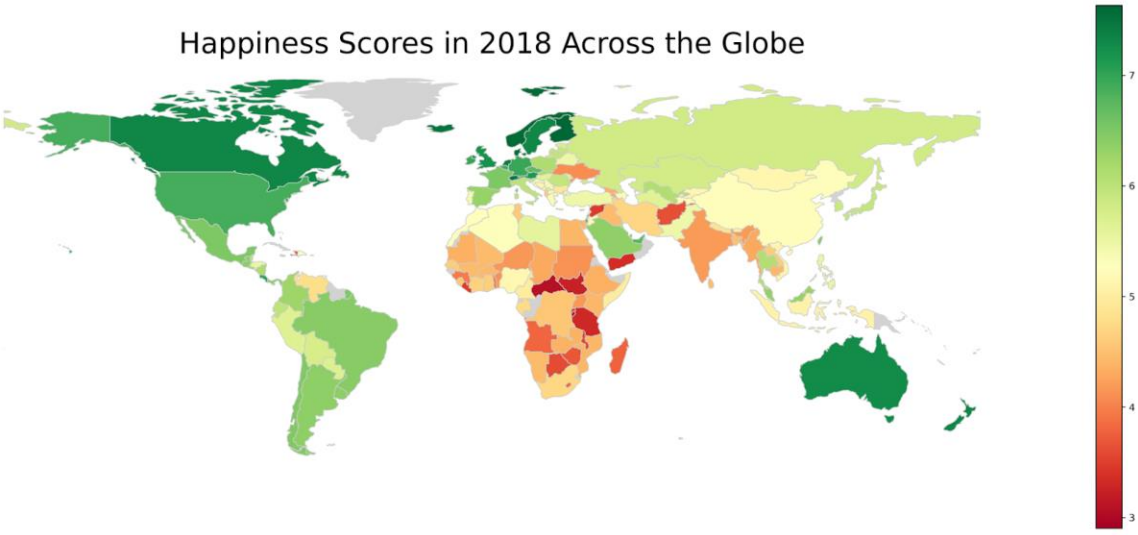
In order to get an understanding of the relationship between happiness and inequality we started by observing the distribution of happiness across the globe in the years 2015-2018, although only years 2015 and 2018 are shown here in the interest of brevity. Visualizing the data in this way serves two purposes, first, by displaying the distribution of happiness cross-sectionally, it is easy to visualize which countries of the world have the highest levels of happiness and which have the lowest. By showing the changes in the distribution over time, it is interesting to try and contextualize the changes visible over the four years to get some sense of what contributes to self-reported measures of happiness.




Note: Countries Represented in light grey do not have data.

a. 2015

Looking at the 2015 map, the most obvious observation is that the distribution of happiness closely mirrors the distribution of wealth. The OECD countries in particular have respondents who perceive themselves to be happier than citizens of the poorer countries in the world. The most unhappy countries are distributed heavily within sub-Saharan Africa, which is expected as sub-Saharan Africa is currently the epicenter of the world with regard to poverty. Moreover, Syria, Afghanistan, and Myanmar all display abnormally high levels of happiness relative to their neighboring countries, which is likely a result of the armed conflicts each country was suffering through in 2015.




Note: Countries Represented in light grey do not have data.

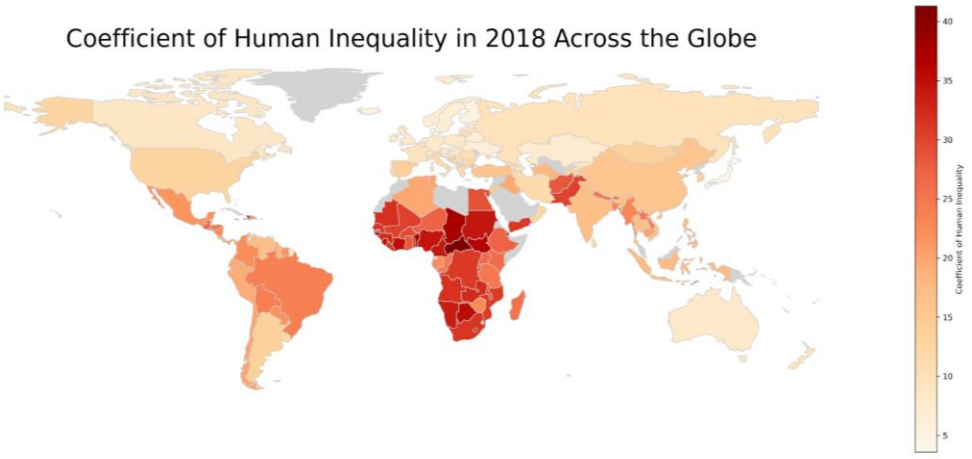
b. 2018

When comparing the distribution of happiness in 2015 to 2018, some interesting trends emerge. While the overall distribution of happiness has remained relatively constant, with wealthier countries appearing happier and poorer countries appearing less happy, large countries like the US, Brazil, Mexico, and India all have decreasing levels of happiness relative to their 2015 levels. In the three years between 2015 and 2018, all four countries had populist leaders come to power, leaders who have inspired massive internal controversy within each respective country. Perhaps the increasing political animosity in each country had some effect on individuals' own interpretations of their happiness, as the news cycle may have been more focused on internal divisions and vitriol within each country's political environment.

2. **Inequality**

With the distribution of happiness across the globe in mind, we decided to next turn our attention towards the distribution of inequality across the globe in 2018. In creating these maps, we found it insightful to compare the distribution of varying measures of inequality with the distribution of happiness in 2018 in order to get a sense of how well each statistic mirrors the distribution of happiness, supposing it might provide insight into certain measures of inequality weighing more heavily on an individual's happiness. To that end, it is interesting to focus on two countries, Mexico and Brazil, throughout this analysis to get a sense of the relationship between happiness and inequality. Both countries displayed higher than average levels of happiness in each year from 2015 to 2018.

It is important to note, though, that all comparisons we make are purely comparisons. We have not performed inferential tests and do not assume to have identified any semblance of causation in our analyses. Instead, we are purely identifying trends and providing visualizations of the distribution of inequality and happiness in order to illuminate some descriptive similarities, similarities which could be a product of hundreds of unobservable effects and requiring much more in-depth research and design.

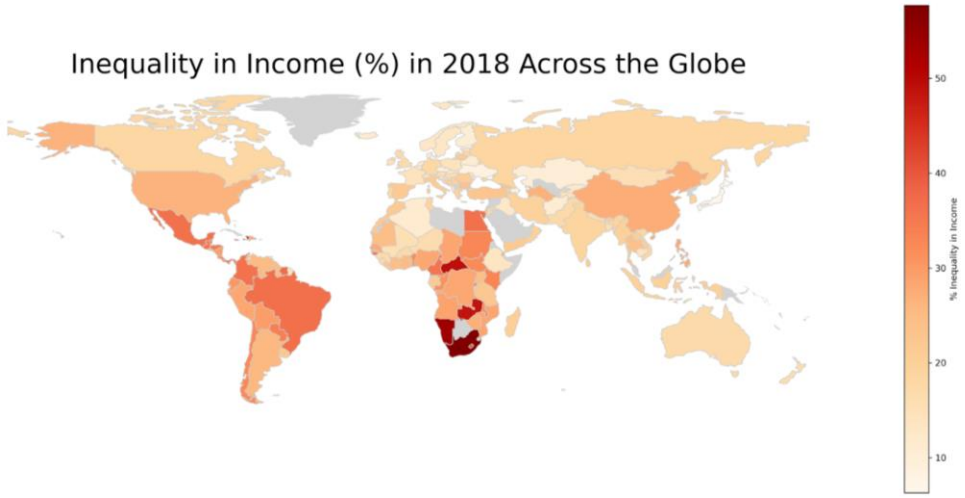



Note: Countries Represented in light grey do not have data. Source: UNDP

**c. Coefficient of Human Inequality**

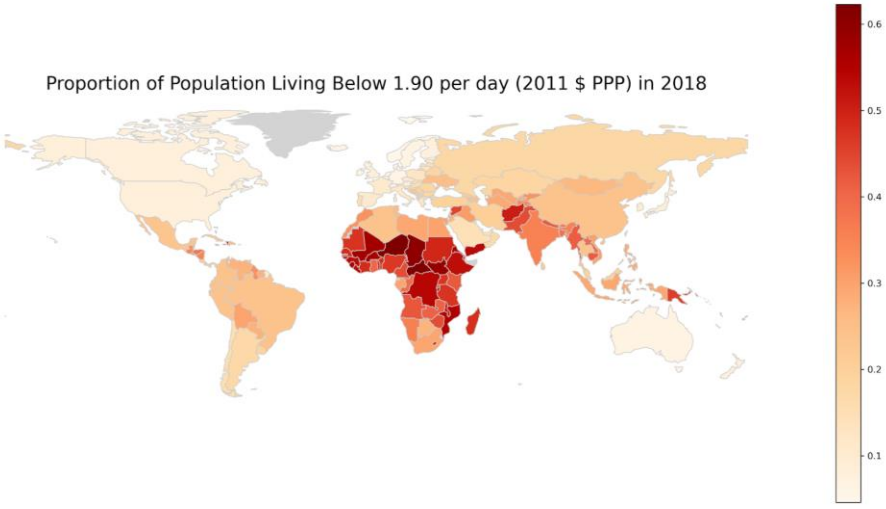
We first looked at the Coefficient of Human Inequality, a statistic which is a simple average of inequalities in health, education, and income calculated by more advanced indices in the Human

Development Report. As measured by the CHI, inequality appears most heavily distributed in Africa, with particularly high levels in Pakistan and Afghanistan, as well. The wealthier OECD countries have relatively low levels of inequality, which is in line with our expectations of happier countries having lower levels of inequality. Interestingly, though, both Brazil and Mexico have very high levels of inequality, yet both countries had relatively high levels of happiness, as well. This suggests that happiness may not be solely dependent on the broad levels of inequality within a country, and inspired us to look next at the component measures of this coefficient in order to see if any one form of inequality was more closely correlated with happiness.




**d. Inequality in Income (%)**

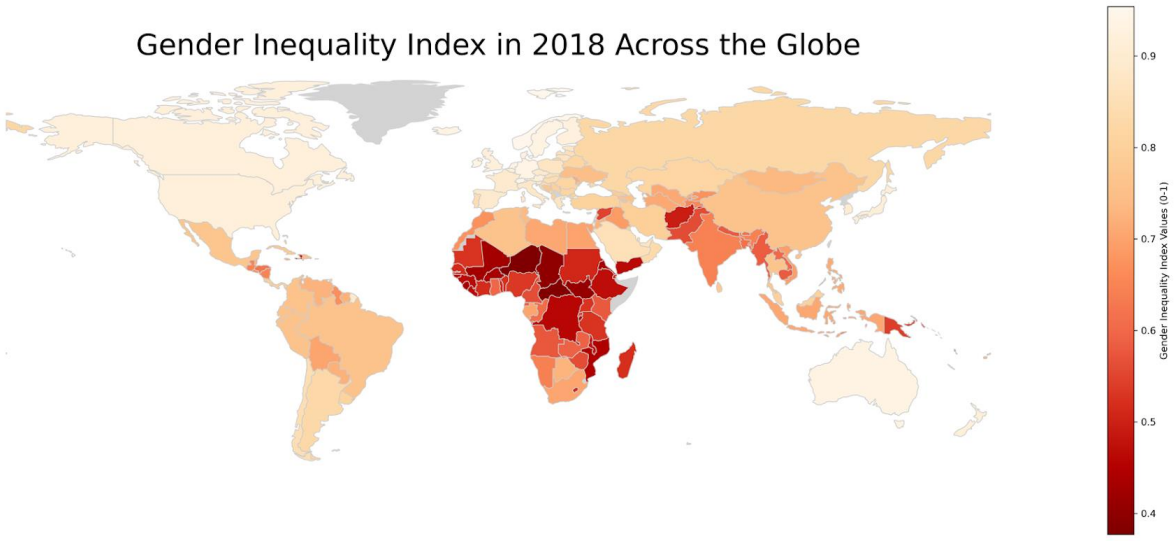
This map displays the distribution in income inequality across the globe in 2018. Curiously, the countries with the highest inequality, South Africa, Namibia, among others, are not the countries with the highest levels of unhappiness. Moreover, both the United States and China have higher levels of inequality relative to their peer countries, yet both countries reported relatively high levels of happiness. Turning to our focus countries of Brazil and Mexico, both countries have relatively high levels of income inequality, but as noted, they had high levels of happiness compared to their peer countries. Taken together, this map suggests that our hypothesis might be incorrect; income inequality might not be the best predictor of happiness. This inspired us to look next at the distribution of poverty across the globe in order to see if absolute measures of deprivation might be a better predictor of happiness levels across the globe.




Note: Countries Represented in light grey do not have data.

### e. Population Living Below International Poverty Line (%)

When looking at the distribution of absolute poverty, the proportion of people living below the World Bank's international ultra-poor poverty line, a map more closely resembling the original happiness maps we observed emerges. Both Mexico and Brazil have low levels of their population living in extreme poverty, as do most other OECD countries. In contrast, the unhappiest countries, that is sub-Saharan Africa, Pakistan, and Afghanistan, all have very high levels of their population living below the poverty line. This suggests that, contrary to our original hypothesis, measures of absolute deprivation may have a higher correlation with the distribution of happiness than relative measures of deprivation as captured by the income inequality graph. However, we have primarily looked at income as yet, and will thus turn to the distribution of other measures of inequality to see their distribution across the globe as well as their similarities to the happiness maps.




Note: Countries Represented in light grey do not have data.

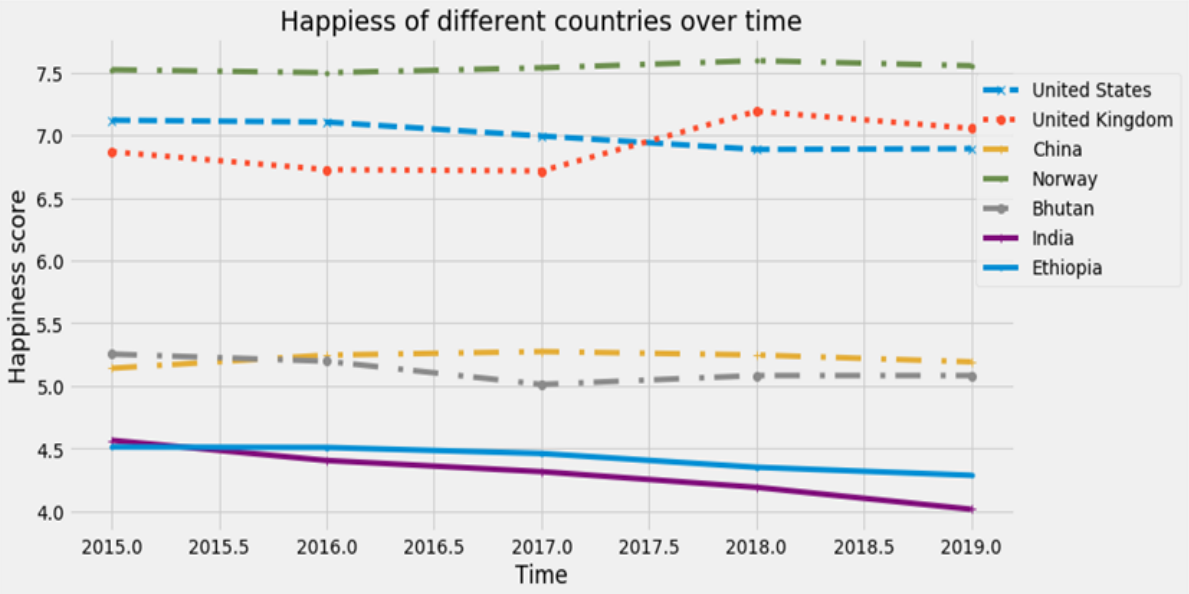
f. Gender Inequality Index

Gender inequality was not considered in the Coefficient of Human Inequality figure displayed in section II, part a. of this report. Thus, we thought it would be interesting to compare gender inequality, a measure many economists and thinkers alike surmise has a considerable effect on happiness, to our graphs measuring the distribution of happiness. This index in particular factors in labor force participation, empowerment, and reproductive health to determine its measure of gender inequality. Interestingly enough, the distribution of gender inequality closely mirrors that of the happiness graphs. Both Mexico and Brazil have relatively normal levels of gender inequality relative to their Latin American peers, and most OECD countries have particularly low levels of gender inequality. The areas with the highest levels of dissatisfaction, sub-Saharan Africa, Pakistan, and Afghanistan, all display extremely high levels of gender inequality, which suggests that there may be some correlation with levels of gender inequality and happiness. We have theorized that this correlation could be a result of the higher impact gender inequality would have in the day-to-day life of an individual than other measures of inequality explored like income and education.

TIME TREND ANALYSIS

To examine time trends in happiness and inequality we focussed on 7 countries with diverse geographies and levels of economic development. These were: USA, UK, Bhutan, India, Ethiopia, China and Norway

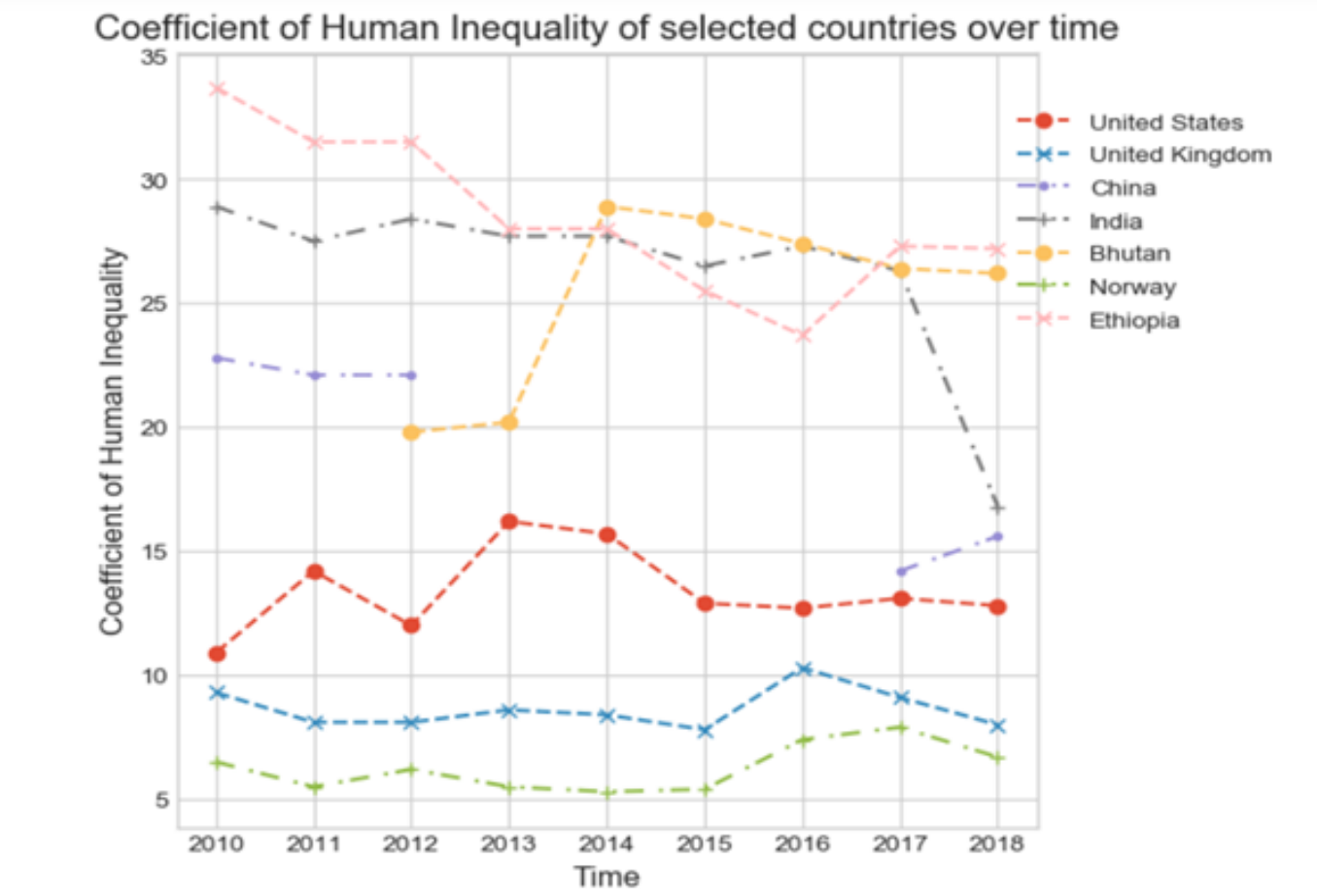
1. **Trends in Happiness over time**




Norway has much higher happiness scores (as do many other Scandinavian countries) and it has more or less remained constant over time. In the US, it is surprising to see that happiness declined in 2017 and 2018 compared to earlier levels. Similarly, happiness has been continuously declining in India and Ethiopia since 2016 despite high growth rates in GDP per capita, falling poverty and increased life expectancy.

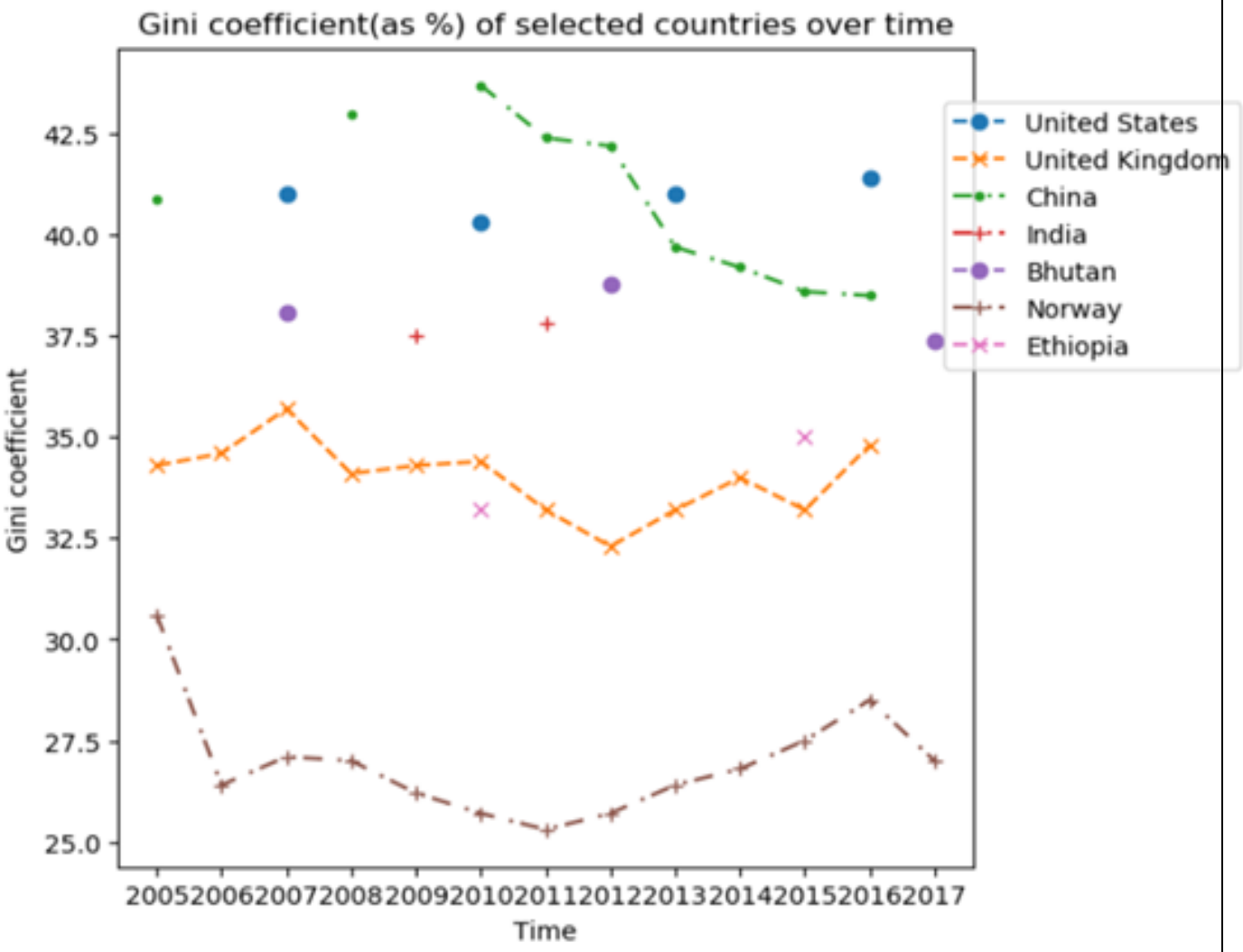
## 2. Trends in Inequality over time

### a. Coefficient of Human Inequality



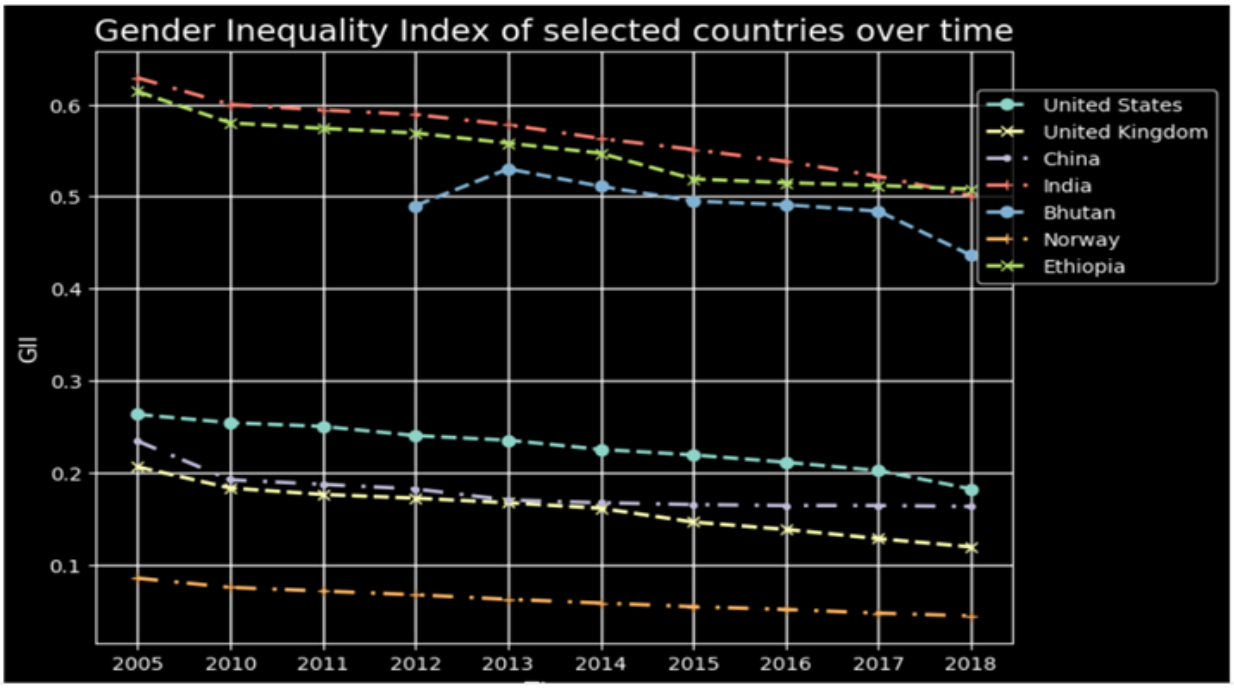

Ethiopia and India both have very high coefficients of human inequality compared to the rest of the countries studied. In particular, Ethiopia seems to have made huge improvements on this front, with massive decreases in inequality per year. Bhutan has increasing levels of human inequality over time though as the previous figure indicates their happiness levels remain constant which is an interesting divergence. UK and USA seem to have a lot of variation over time and there was a rise in human inequality between 2012-15 in the US and between 2015-17 in the UK though the values declined after these periods in both the countries.

**b. Trends in Gini coefficients over time**




Income inequality measured by the Gini is one of the most common (and often only) measures of inequality in development economics. Data for some of the countries were missing over certain periods of time for this particular variable for the US and Bhutan. Nevertheless, China stands out for its huge decreases in income equality (they also witnessed huge decreases in human inequality). UK and Norway, on the other hand, have had an increase in income inequality since 2012(with small intermediary decreases) which is quite contrary to their performance on some of the other inequality measures.

### c. Trends in Gender inequality over time




Gender inequality is often not included in some of the broader measures of inequality and is not a component of happiness scores. However, it is encouraging to see that most countries in the sample have had significant decreases in gender inequality over time.

### d. Trends in life expectancy over time

Life expectancy is not a commonly considered component in the study of inequality. However it is a component of the happiness scores measured by the UN-SDSN and on plotting a heat map on happiness scores and sub-components we see it is highly correlated with the level of happiness of countries. A scatter plot of the life expectancy inequality and happiness using different data from the UNDP report reconfirms the negative (and high) correlation between the two. On plotting time trends, it is again encouraging to see that most counties have made great progress on this over the last decade particularly middle-income countries such as Ethiopia, India and Bhutan which have had high inequality in life expectancies.

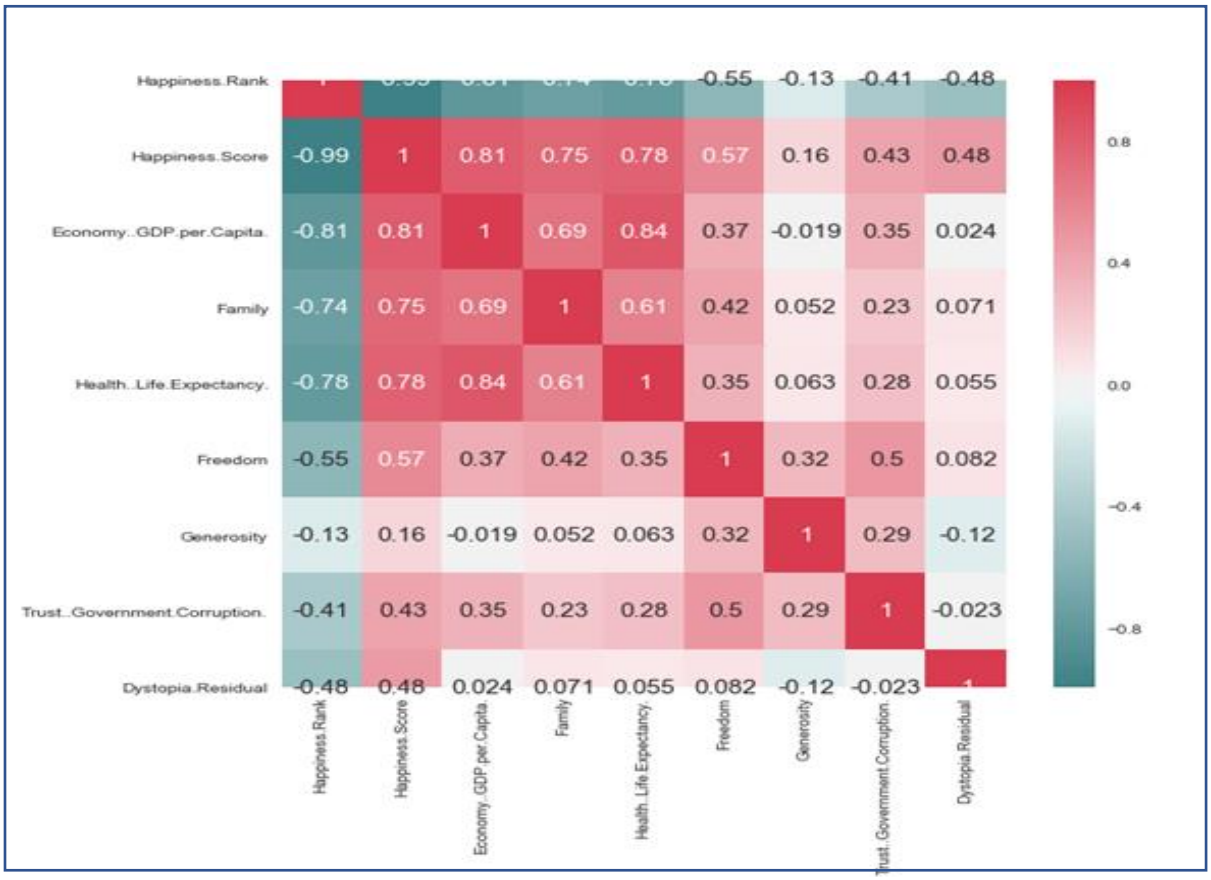



*Figure: Heatmap of happiness scores components*

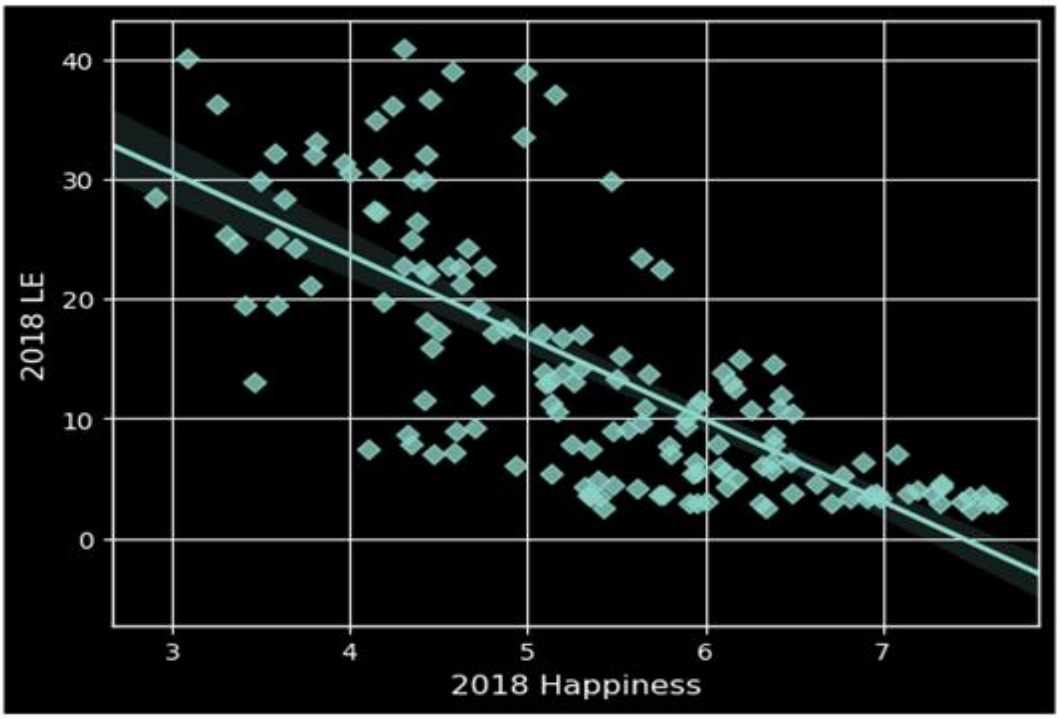
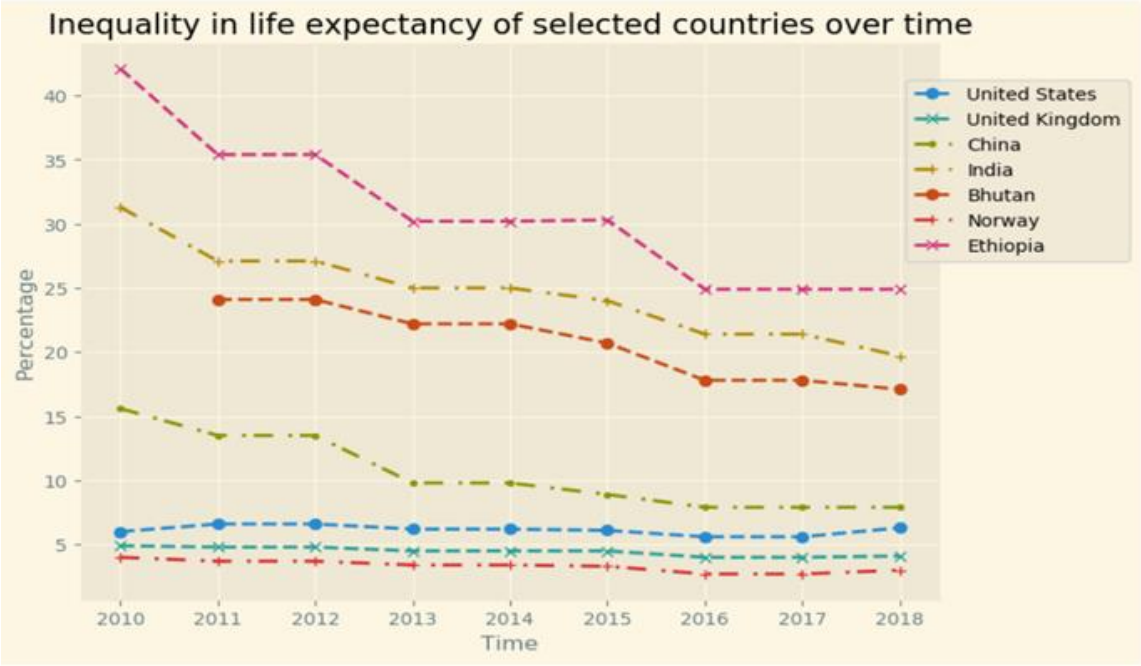



*Figure: Correlation of Happiness score and Inequality in life expectancy*




## V. Conclusion

We initially began this study with an aim to study the relationship between inequality and happiness with the hypothesis that they could have a strong relationship since people often think about their levels of personal satisfaction in relation to that of others. While statistical tools for causation etc. were not used, descriptive statistics gave us various interesting insights indicating the importance of this topic and the potential for more careful analysis on this issue.

Cross-country analysis indicated that one of the most striking observations, the poorest countries were also the unhappiest. While there may be a chance that when comparing two countries with different income levels, the richer country may not always be the happier country, this does seem to hold below a threshold level of income. For many countries in the world, particularly in Sub-Saharan Africa, the absolute level of deprivation is so low hence that becomes more important than relative levels of deprivation. And since many countries in the world have high absolute deprivation this could be a possible reason for why we observe per-capita income at the world level to be highly correlated with happiness. Another important factor was life-expectancy in the country and we observed strong negative correlations between happiness and inequality in life-expectancy ratified by different data sources. For some countries such as India and the US even though GDP per capita has been rising over time, happiness scores have been declining. China does extremely well on dimensions of income inequality (values have been consistently falling) and Ethiopia shows great improvements in human equality (falling coefficient of human inequality). Gender inequality has also improved in most countries over time and this is usually not covered in broader measures of inequality. These analyses show us that there are possible relationships between these two variables and a careful in-depth analysis with good data could prove to be really useful from a policy standpoint.