

Measuring and Visualising Global Happiness

Fiona Marshall
Ulster University
Shore Road, Newtownabbey
Co. Antrim, UK. BT370QB
Marshall-F1@ulster.ac.uk

Raymond Bond
Ulster University
Shore Road, Newtownabbey
Co. Antrim, UK. BT370QB
rb.bond@ulster.ac.uk

Shuai Zhang
Ulster University
Shore Road, Newtownabbey
Co. Antrim, UK. BT370QB
s.zhang@ulster.ac.uk

Happiness scores from the World Happiness Reports have been examined alongside data from the World Health Organisation and the World Bank to show that happiness is an effective measure of well-being. There is a strong correlation between happiness and GDP and other factors such as years of schooling and life expectancy. Using data from the World Health Organisation and World Bank, the happiness of a nation can be predicted with a RMSE of 0.3. Perhaps, predicting the happiness of a community, nation or industry can help inform interactive digital interventions for happiness building. Possible, such measures could also feature as predictors in affective computing contexts or indeed via social media.

Happiness, Well-Being, Visualisation

1. INTRODUCTION

Happiness as “*the state of being happy*” with *happy* defined as “*feeling or showing pleasure or contentment*” [1-2]. As happiness is subjective, it can appear difficult to measure and most studies rely on simply asking participants how happy they are. Although this may appear too subjective to be of any use in scientific comparisons, research claims that the subjective feelings expressed can be backed up by objective data [3].

This paper is based on the findings of the World Happiness Reports for 2015-2017. The first report, published in 2012, opens: “*We live in an age of stark contradictions. The world enjoys technologies of unimaginable sophistication; yet has at least one billion people without enough to eat each day.*” [4]. It points out that whilst enjoying economic success, wealthy countries seem to be suffering from previously unknown crises such as obesity and mental health. The introduction concludes: “*It makes sense... to pursue policies to raise the public’s happiness as much as it does to raise the public’s national income*” [4].

Gross Domestic Product (GDP) represents the total value of a countries goods and services over a period of time. GDP rapidly became the tool of choice for gauging the health of a countries economy. However, GDP has long had its critics. Jigme Singye Wangchuck declared that his aim was not to increase his country’s GDP, rather its GNH –

Gross National Happiness. Richard Easterlin observed that whilst the GDP of the USA had been rising consistently, happiness level had remained static – the Easterlin Paradox [5]. David Cameron, as UK prime minister led the way when he proposed to parliament (in the middle of recession) the creation of a new measure – the general wellbeing index stating that “GDP is an incomplete way of measuring a country’s progress” [6]. Since 2010 many other countries have followed suit.

The quality of life in the UK is measured using 43 different indicators covering areas such as personal well-being, health, relationships and where we live amongst others. Happiness is included within personal well-being. The chief statistician comments, “*There is growing recognition that how we are doing as a nation is at least as much about people’s well-being as it is about the country’s economic health*” [7].

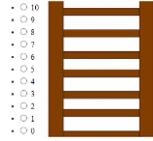
The Organisation for Economic Cooperation and Development (OECD) has encouraged member states to report on 12 different indices as part of its Better Life Initiative which was launched in 2011. Indices include housing, income, jobs, communities, education, environment, civic engagement, life satisfaction, safety and work-life balance [8].

The World Happiness Report is produced annually since 2012. The 2017 report states “*Increasingly, happiness is considered to be the proper measure of social progress and the goal of public policy*” [9]. A

total of 1,000 people in 150 countries each year are asked to evaluate their happiness by placing themselves on the Candrill Ladder:

Rate Your Life As A Whole

The ladder depicts life satisfaction. How good is your life when you step back and think about it? The 10 at the top represents the best possible life for you, with lower numbers indicating lesser degrees of fulfillment. On which step of the ladder do you feel you stand now?



The purpose of this paper is to discover if there is any relationship between wealth and happiness and to identify other factors that influence the perceived happiness of individuals and societies. Specifically, it will aim to answer the following questions:

- Is happiness changing over time?
- How is happiness distributed throughout the world?
- What are the factors that affect our happiness?
- Can we predict happiness?

2. METHODS

Three datasets were used containing 155-158 observations and 14 features. The datasets included, 1) data from the *World Happiness Report 2015-2017*, 2) *World Bank GDP Per Capita data (GDP for 264 countries from 1959 to 2016)* and 3) *the World Health Organisation Life Expectancy (2900 observations and 23 features)*.

The final 9 features described in the World Happiness Report are the residuals and are the authors' attempt to explain the extent to which each feature contributes towards the happiness score for each country [10].

- Country
- Region
- Happiness Rank
- Happiness Score
- Standard Error
- Economy
- GDP per Capita
- Family
- Health
- Life Expectancy
- Freedom
- Trust Government Corruption
- Generosity
- Dystopia Residual

The World Health Organisation Life Expectancy data comprised of the following features:

- Country
- Year
- Status
- Life expectancy
- Adult Mortality
- infant deaths
- Alcohol
- Hepatitis B
- Measles
- BMI
- Schooling
- Income
- Under-five deaths
- Polio
- Total expenditure
- Diphtheria
- HIV/AIDS
- GDP
- Population
- thinness 1-19 years
- thinness 5-9 years
- percentage expenditure
- composition of resources

RStudio was used for all data analytics. Data preprocessing involved imputation of missing values and ensuring that identical names were used for each country. In the *World Happiness Reports*, there were no missing values, however there were slight differences in the Countries for whom data was available each year.

For the World Bank data, there were 46 missing values from 2015 and 2016 that were imputed using mice and data from year 2000 onwards.

For the WHO data, there were 521 missing values (mainly expenditure) that were imputed using kNN due to high level of correlation between features.

3. RESULTS

To study the spread of happiness throughout the world, countries are coloured according to the mean of their happiness, with yellow being the most happy and red the unhappiest. The same map was coloured according to the rank order of each country's GDP (mean of 2015 and 2016) per person. See Figure 1 which illustrate the similarity between maps.

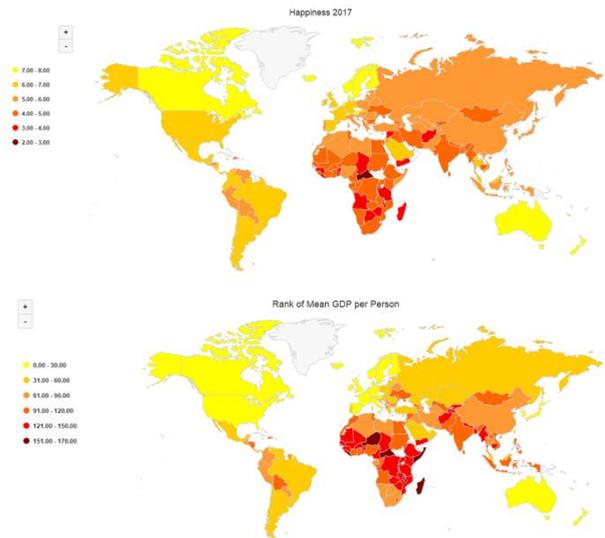


Figure 1. World happiness and rank of mean GDP.

Most of the data from the world happiness residuals were not normally distributed, hence a Kendall's coefficient correlation matrix was used to identify relationships. The residuals of the world happiness report show that there is a strong correlation between happiness and GDP, family and health and life expectancy (see Figure 2). The remainder of this report will be to explore whether these correlations hold when the other data sets are used.

Correlation Between Happiness and its Residuals 2017

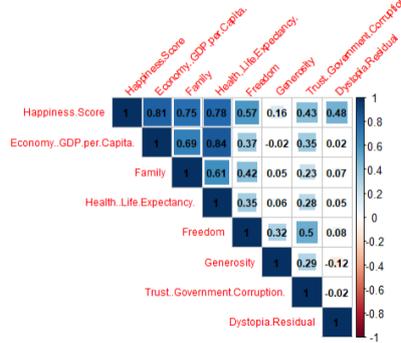


Figure 2. Correlation analysis matrix.

Happiness scores were plotted against GDP (per person) for each country and a smoothed regression line was added. Shading shows the 95% confidence interval of the line. Colour denotes the region.

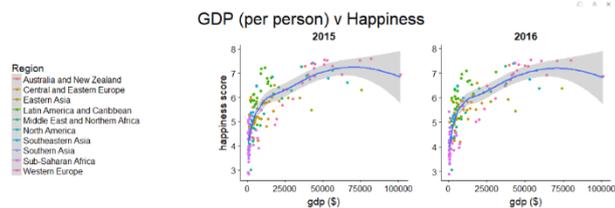


Figure 3. Happiness scores were plotted against GDP.

We can see a strong, although changing relationship between GDP and happiness. Where GDP is low (less than 20,000 per person), happiness increases sharply with any increase in the national wealth. Between \$20,000 and \$50,000, as GDP increases, happiness continues to increase, albeit at a slower rate. Above \$50,000, the increase in happiness with GDP is less assured and the regression line suggests that too much wealth may, possibly, not be in a country's best interest where happiness is concerned. These correlations are shown even more clearly when a power transformation of both variables is taken. Spearman's correlation coefficient (used because data is not normal and has outliers) is 0.63 for 2015 and 0.64 for 2016.

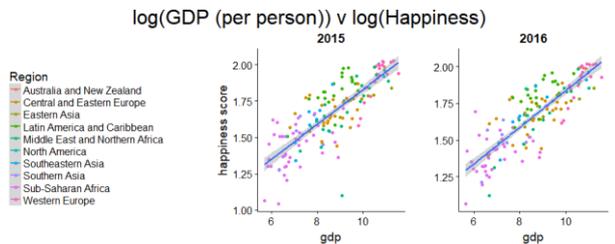


Figure 4. With a power transformation, happiness scores were plotted against GDP (per person).

Life expectancy and schooling data (Figure 5) from the WHO was plotted against the mean happiness score. A correlation between the 2 features are

seen when life expectancy is above 60 years old, and school leaving age is above 10 years old.

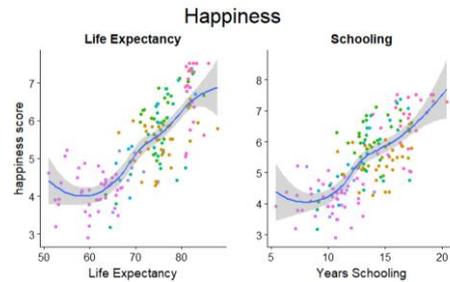


Figure 5. Association between life expectancy and schooling.

Kendall's correlation was used as the data was not normally distributed and did not contain outliers. The correlation between happiness and life expectancy was 0.56 and between happiness and schooling was 0.52. (Kendall's tau usually gives smaller values than Spearman's rho [11]).

We wanted to test whether happiness has changed over the 3 years of the report. The Shapiro-Wilk's test failed to suggest normality, however, Shapiro-Wilk's is sensitive to outliers and large sample sizes [12]. Hence, QQ-plots were used to compare the distribution of happiness to the normal distribution [13]. Here the critical value of r was 0.991, 0.991 and 0.933 for the 3 years – clearly suggesting normality. The p-value of a paired ANOVA test was 0.9779, suggesting that happiness levels have not changed over the 3 years. This result is reflected in the density plot of the 3 years happiness scores.

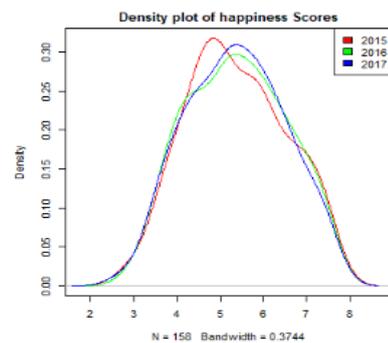


Figure 6: Density plot of happiness for each year.

Having shown that happiness has remained static over the 3 years, we then carry on to see whether the distribution of wealth has remained static over time. The data was clearly not normal. The p-value for the paired t-test for GDP is greater than 0.05, and so we are confident that the distribution has not changed. However, the unpaired t-test gives a higher p-value, suggesting that whilst we are extremely confident in the spread of GDP being unchanged over the two years, we are less confident in each individual country's GDP being unchanged.

Density plots (Figure 6) show clearly the difference in happiness levels in different regions. Formal tests were not used to show differences in happiness levels by region or change in regional happiness over time due to the small sample size of some of the regions (Australia & New Zealand and North America both only contain 2 countries). Box plots (Figure 7) show how happiness levels have remained mostly static over the 3 years in each region. Interestingly, it appears that North America is the only region that looks as if its happiness levels may have changed - showing a slight decrease.

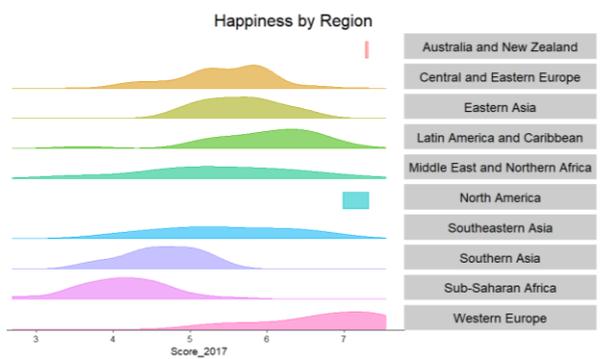


Figure 6. Happiness distribution in different regions.

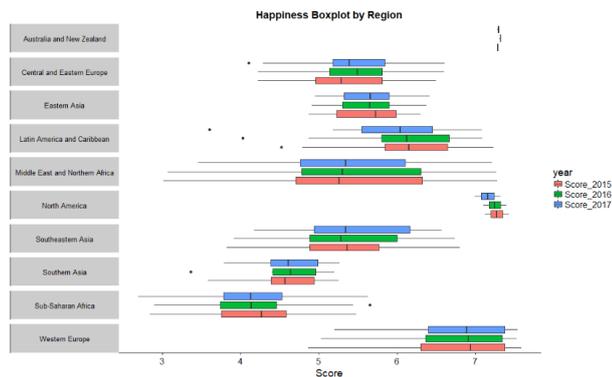


Figure 7. Boxplots of happiness in different regions.

Machine learning was used with all the features from the WHO dataset and the GDP for 2015 and 2016 to see whether happiness could be predicted. Linear regression, neural networks and random forest algorithms were used to represent the spectrum of different models available. All 3 models were trained using 10-fold cross validation (CV) on 80% of the data, and then tested on the remaining 20%. The root mean square error (RMSE) of the results were compared (Figure 8 and 9 showing results).

Random Forest performed best in both training and testing, most likely due to the fact that it uses an ensemble of algorithms and does not require feature selection. However, with a RMSE of 0.3 in predicting a score between 1 and 10, it has been proven that it is possible to predict a nations happiness with known latent features.

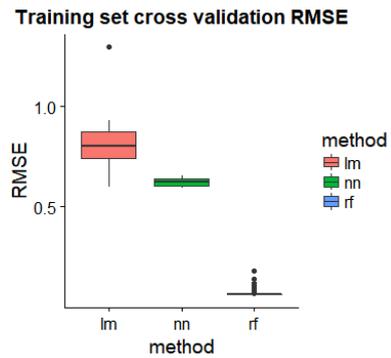


Figure 8. RMSE of each method during cross validation.

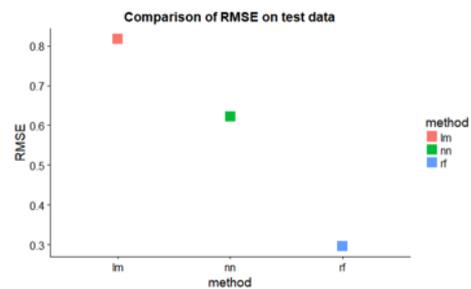


Figure 8. RMSE of each model during testing.

4. SUMMARY

GDP is no longer the sole measure by which a countries progress is measured. It is clear however that GDP and happiness are strongly linked, especially for countries who do not find themselves in the top quartile of GDP. Life expectancy and years of schooling are also show a correlation to happiness, although as both of these can be the result of a country's wealth, this may prove to be as a result of GDP being a confounding factor. Publishing with the 2018 World Happiness Report was, for the first time, one database containing all the happiness scores since 2012. This opens up the possibility to compare changes in happiness over greater periods of time and to look for reasons to explain changes. Another avenue of research is to look at data for individual countries and examine whether the trend towards happiness with greater wealth, life expectancy and schooling is reflected within regions of the same country.

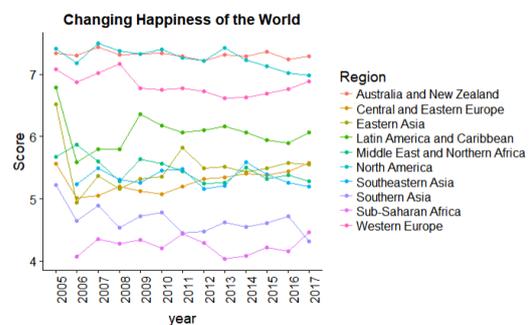


Figure 9. Time series of happiness over regions of the world.

5. REFERENCES

- [1] (2018) English Oxford Living Dictionaries, <https://en.oxforddictionaries.com/definition/happiness> (retrieved May 2018)
- [2] Khoddam, R, What's Your Definition of Happiness (2015) <https://www.psychologytoday.com/intl/blog/the-addiction-connection/201506/whats-your-definition-happiness> (retrieved May 2018)
- [3] Oswald, A.J, Stephen, Wn. (2010) Objective Confirmation of Subjective Measures of Human Well-Being: Evidence from the USA. *Science* Vol. 327, pp. 576-579.
- [4] World Happiness Report 2012 (2012) <http://worldhappiness.report/ed/2012/> (retrieved May 2018)
- [5] Easterlin, R (1974). Does Economic Growth Improve the Human Lot? Some Empirical Evidence In Paul A. David; Melvin W. Reder. *Nations and Households in Economic Growth: Essays in Honor of Moses Abramovitz.* Academic Press, Inc. New York.
- [6] Cameron, D. (2010) PM Speech on Well Being. <https://www.gov.uk/government/speeches/pm-speech-on-wellbeing> (retrieved May 2018)
- [7] (2018) Measures of National Well Being Dashboard. <https://www.ons.gov.uk/peoplepopulationandcommunity/wellbeing/articles/measuresofnationalwellbeingdashboard/2018-04-25> (retrieved May 2018)
- [8] Manclossi, S. (2018) Measuring National Well-being: Quality of Life in the UK, <https://www.ons.gov.uk/peoplepopulationandcommunity/wellbeing/articles/measuringnationalwellbeing/qualityoflifeintheuk2018> (retrieved May 2018)
- [9] World Happiness Report 2017 (2017) <http://worldhappiness.report/ed/2017/> (retrieved May 2018)
- [10] Understanding How Gallup Uses the Cantril Scale, <http://news.gallup.com/poll/122453/understanding-gallup-uses-cantril-scale.aspx> (retrieved May 2018)
- [11] Xu weichao , yunhe hou , hung y. s. & yue Xian zou (2010) Comparison of Spearman's rho and Kendall's tau in normal and contaminated normal models. Manuscript submitted to *IEEE Transactions on Information Theory* <https://www.degruyter.com/downloadpdf/j/quageo.2011.30.issue-2/v10117-011-0021-1/v10117-011-0021-1.pdf> (retrieved May 2018)
- [12] Gentleman, JF, Wilk MB, Detecting Outliers in a Two-Way Table: Statistical Behaviour of Residuals, *Technometrics* , Volume 17 (1975)
- [13] Ghamesi, A, Zahediasl, S, Normality Test for Statistical Analysis, *A Guide for Non-Statisticians*, (2012) *Int J Endocrinol Metab*, Vol 10, p486-489.