

The Search for Happiness

Happiness scored according to economic production and social support

Term Project

Data Science Team 1

Alina Aliakhnovich

Mustafa Avci

Mei Hui Chen

Chetan Khare

Buse Tuncel

Table of Content

Table of Contents

| | |
|--|-----------|
| Table of Content..... | 2 |
| Background | 3 |
| Overview..... | 4 |
| Definitions | 4 |
| Data Sources..... | 4 |
| Approach | 5 |
| Cleaning: Missing Value..... | 5 |
| Cleaning: Variable Check/ Exclusions | 6 |
| Statistics | 7 |
| Skewness and Kurtosis Control:..... | 7 |
| Data Analysis: | 8 |
| Data Analysis by Variables | 8 |
| Correlation Plot..... | 8 |
| During this analysis correlation between happiness score and other variables is calculated. | 8 |
| Correlation in Pair Graphs..... | 9 |
| Impact of Variables on Happiness Score | 11 |
| Data Analysis By Continent..... | 14 |
| Happiness Score by Region | 14 |
| Bubble Chart (3D) | 17 |
| Pie Chart..... | 19 |
| Data Analysis by Country | 19 |
| Economy Analysis based on Countries | 20 |
| Family Analysis based on Countries | 21 |
| Health Analysis based on Countries | 22 |
| Further Analysis | 22 |
| Impact of each Variable & Top 5 countries:..... | 24 |
| Conclusion..... | 26 |
| References | 27 |

Background

The World Happiness Report is an annual publication of the United Nations Sustainable Development Solutions Network. It is a survey of the state of global happiness that ranks 158 countries by how happy their citizens perceive themselves to be. This report was first published in 2012, and since then, this report demonstrated that happiness and wellbeing are crucial indicators of nations' economic and social status.

Since 2012, this report has been published annually on March 20, the date that the United Nations has declared to be the International Day of Happiness. The report ranks countries on six key variables that support well-being: income, freedom, trust, healthy life expectancy, social support, and generosity.

The World Happiness Report 2015 focused on the changes in happiness levels in 158 countries and investigated the reasons behind the statistics. It also allows world leaders the opportunity to realize the current situation and reshape the global agenda. It will continue to attract global recognition as governments, organizations, and society to inform policymakers.

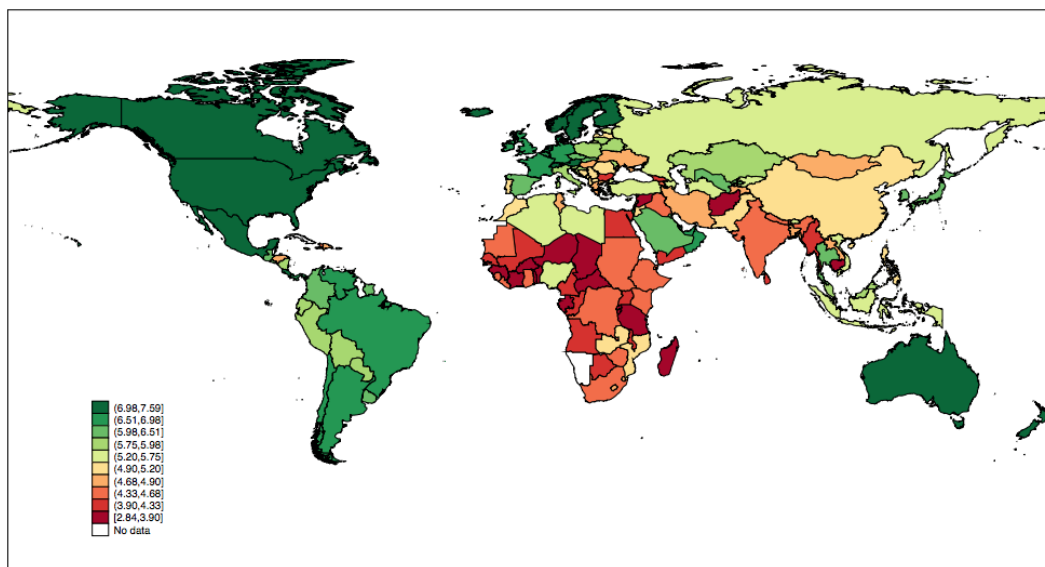


Figure: The Geography of Happiness in 2015

Figure 1 shows the geographic distribution of 2012 to 2014 average national values for the Cantril ladder survey. Cantril ladder survey asks its respondents to value their lives today on a 0 to 10 scale, with the worst possible life as a 0 and the best possible life as a 10. In this map, countries are divided into 10 groups with the darkest green for the highest averages and darkest red for the lowest.

In these reports, experts in various fields including economics, psychology, survey analysis, and national statistics, report how measurements of well-being can be used effectively to assess the progress of nations. Each report is organized by chapters that dive deeper into issues relating to happiness, including mental illness, the objective benefits of happiness, the importance of ethics, policy implications, and links with the Organisation for Economic Co-operation and Development's (OECD) approach to measuring subjective well-being and other international and national efforts.

Overview

Definitions

- ❑ Dystopia is an imaginary country that has the world's least-happy people. The purpose in establishing Dystopia is to have a benchmark against which all countries can be favorably compared (no country performs more poorly than Dystopia) in terms of each of the six key variables, thus allowing each sub-bar to be of positive width. The lowest scores observed for the six key variables, therefore, characterize Dystopia. Since life would be very unpleasant in a country with the world's lowest incomes, lowest life expectancy, lowest generosity, most corruption, least freedom and least social support, it is referred to as "Dystopia," in contrast to Utopia.
- ❑ The residuals, or unexplained components, differ for each country, reflecting the extent to which the six variables either over- or under-explain average 2014-2016 life evaluations. These residuals have an average value of approximately zero over the whole set of countries.
- ❑ The columns: GDP per Capita, Family, Life Expectancy, Freedom, Generosity, Trust Government Corruption describes the extent to which these factors contribute to evaluating happiness in each country. If all these factors with residual are added up, the happiness score can be calculated.

Data Sources

For this analysis 2015.csv (with 12 columns) is used. Data set can be approached from here: <https://www.kaggle.com/unsdsn/world-happiness#2015.csv>

Metadata

Metadata of 2015.csv (16.17 KB) Open Dataset – World Happiness Report

| # | Variables | Type | Scale | Description |
|----|-------------------------------|------------------|------------|--|
| 1 | Country | Char | Nominal | Name of the country. |
| 2 | Region | Char | Nominal | Region the country belongs to. |
| 3 | Happiness Rank | Numerical-Int | Continuous | Rank of the country based on the Happiness Score. |
| 4 | Happiness Score | Numerical-Float | Continuous | A metric measured in 2015 by asking the sampled people the question: "How would you rate your happiness on a scale of 0 to 10 where 10 is the happiest." |
| 5 | Standard Error | Numerical- Float | Continuous | The standard error of the happiness score. |
| 6 | Economy (GDP per Capita) | Numerical-Float | Continuous | The extent to which GDP contributes to the calculation of the Happiness Score. |
| 7 | Family | Numerical-Float | Continuous | The extent to which Family contributes to the calculation of the Happiness Score |
| 8 | Health (Life Expectancy) | Numerical-Float | Continuous | The extent to which Life expectancy contributed to the calculation of the Happiness Score |
| 9 | Freedom | Numerical-Float | Continuous | The extent to which Freedom contributed to the calculation of the Happiness Score. |
| 10 | Trust (Government Corruption) | Numerical-Float | Continuous | The extent to which Perception of Corruption contributes to Happiness Score. |
| 11 | Generosity | Numerical-Float | Continuous | The extent to which Generosity contributed to the calculation of the Happiness Score. |
| 12 | Dystopia Residual | Numerical-Float | Continuous | The extent to which Dystopia Residual contributed to the calculation of the Happiness Score. |

Table : Metadata: variable, type, scale and description.

Approach

Cleaning: Missing Value

There are 158 observations and no missing values in the dataset.

```

RangeIndex: 158 entries, 0 to 157
Data columns (total 12 columns):
Country                158 non-null object    Country                0
Region                158 non-null object    Region                0
Happiness Rank        158 non-null int64      Happiness Rank        0
Happiness Score       158 non-null float64    Happiness Score       0
Standard Error        158 non-null float64    Standard Error        0
Economy (GDP per Capita) 158 non-null float64    Economy (GDP per Capita) 0
Family                158 non-null float64    Family                0
Health (Life Expectancy) 158 non-null float64    Health (Life Expectancy) 0
Freedom               158 non-null float64    Freedom               0
Trust (Government Corruption) 158 non-null float64    Trust (Government Corruption) 0
Generosity            158 non-null float64    Generosity            0
Dystopia Residual      158 non-null float64    Dystopia Residual     0
dtypes: float64(9), int64(1), object(2)

```

Cleaning: Variable Check/ Exclusions

| | Country | Region | Happiness Rank | Happiness Score | Standard Error | Economy (GDP per Capita) | Family | Health (Life Expectancy) | Freedom | Trust (Government Corruption) | Generosity | Dystopia Residual |
|---|-------------|----------------|----------------|-----------------|----------------|--------------------------|---------|--------------------------|---------|-------------------------------|------------|-------------------|
| 0 | Switzerland | Western Europe | 1 | 7.587 | 0.03411 | 1.39651 | 1.34951 | 0.94143 | 0.66557 | 0.41978 | 0.29678 | 2.51738 |
| 1 | Iceland | Western Europe | 2 | 7.561 | 0.04884 | 1.30232 | 1.40223 | 0.94784 | 0.62877 | 0.14145 | 0.43630 | 2.70201 |
| 2 | Denmark | Western Europe | 3 | 7.527 | 0.03328 | 1.32548 | 1.36058 | 0.87464 | 0.64938 | 0.48357 | 0.34139 | 2.49204 |
| 3 | Norway | Western Europe | 4 | 7.522 | 0.03880 | 1.45900 | 1.33095 | 0.88521 | 0.66973 | 0.36503 | 0.34699 | 2.46531 |
| 4 | Canada | North America | 5 | 7.427 | 0.03553 | 1.32629 | 1.32261 | 0.90563 | 0.63297 | 0.32957 | 0.45811 | 2.45176 |

| | Happiness Score | Economy (GDP per Capita) | Family | Health (Life Expectancy) | Freedom | Trust (Government Corruption) | Generosity | Dystopia Residual |
|-------|-----------------|--------------------------|------------|--------------------------|------------|-------------------------------|------------|-------------------|
| count | 158.000000 | 158.000000 | 158.000000 | 158.000000 | 158.000000 | 158.000000 | 158.000000 | 158.000000 |
| mean | 5.375734 | 0.846137 | 0.991046 | 0.630259 | 0.428615 | 0.143422 | 0.237296 | 2.098977 |
| std | 1.145010 | 0.403121 | 0.272369 | 0.247078 | 0.150693 | 0.120034 | 0.126685 | 0.553550 |
| min | 2.839000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.328580 |
| 25% | 4.526000 | 0.545808 | 0.856823 | 0.439185 | 0.328330 | 0.061675 | 0.150553 | 1.759410 |
| 50% | 5.232500 | 0.910245 | 1.029510 | 0.696705 | 0.435515 | 0.107220 | 0.216130 | 2.095415 |
| 75% | 6.243750 | 1.158448 | 1.214405 | 0.811013 | 0.549092 | 0.180255 | 0.309883 | 2.462415 |
| max | 7.587000 | 1.690420 | 1.402230 | 1.025250 | 0.669730 | 0.551910 | 0.795880 | 3.602140 |

Happiness rank and Standard error variables haven't been used in the analysis of the dataset.

There is an inverse correlation between "Happiness Rank" and all the other numerical variables. In other words, the lower the happiness rank, the higher the happiness score, and the higher the other seven factors that contribute to happiness. Therefore the happiness rank is removed from the analysis.

Note that “Standard Error” is also excluded from this analysis as it is not used and does not contribute to this analysis.

Statistics

Variables’ descriptive statistics are shown below:

| | Country | Region | Happiness Rank | Happiness Score | Standard Error | Economy (GDP per Capita) | Family | Health (Life Expectancy) | Freedom | Trust (Government Corruption) | Generosity | Dystopia Residual |
|--------|---------|--------------------|----------------|-----------------|----------------|--------------------------|------------|--------------------------|------------|-------------------------------|------------|-------------------|
| count | 158 | 158 | 158.000000 | 158.000000 | 158.000000 | 158.000000 | 158.000000 | 158.000000 | 158.000000 | 158.000000 | 158.000000 | 158.000000 |
| unique | 158 | 10 | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN |
| top | Bahrain | Sub-Saharan Africa | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN |
| freq | 1 | 40 | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN | NaN |
| mean | NaN | NaN | 79.493671 | 5.375734 | 0.047885 | 0.846137 | 0.991046 | 0.630259 | 0.428615 | 0.143422 | 0.237296 | 2.098977 |
| std | NaN | NaN | 45.754363 | 1.145010 | 0.017146 | 0.403121 | 0.272369 | 0.247078 | 0.150693 | 0.120034 | 0.126685 | 0.553550 |
| min | NaN | NaN | 1.000000 | 2.839000 | 0.018480 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.328580 |
| 25% | NaN | NaN | 40.250000 | 4.526000 | 0.037268 | 0.545808 | 0.856823 | 0.439185 | 0.328330 | 0.061675 | 0.150553 | 1.759410 |
| 50% | NaN | NaN | 79.500000 | 5.232500 | 0.043940 | 0.910245 | 1.029510 | 0.696705 | 0.435515 | 0.107220 | 0.216130 | 2.095415 |
| 75% | NaN | NaN | 118.750000 | 6.243750 | 0.052300 | 1.158448 | 1.214405 | 0.811013 | 0.549092 | 0.180255 | 0.309883 | 2.462415 |
| max | NaN | NaN | 158.000000 | 7.587000 | 0.136930 | 1.690420 | 1.402230 | 1.025250 | 0.669730 | 0.551910 | 0.795880 | 3.602140 |

Skewness and Kurtosis Control:

Skewness Control

```
Happiness Score      0.097769
Economy (GDP per Capita) -0.317575
Family               -1.006893
Health (Life Expectancy) -0.705328
Freedom              -0.413462
Trust (Government Corruption) 1.385463
Generosity           1.001961
Dystopia Residual    -0.238911
dtype: float64
```

Kurtosis Control

```
Happiness Score      -0.776075
Economy (GDP per Capita) -0.866986
Family               0.918819
Health (Life Expectancy) -0.393935
Freedom              -0.460778
Trust (Government Corruption) 1.384787
Generosity           1.746528
Dystopia Residual     0.534121
dtype: float64
```

There is no Kurtosis and Skewness problem in the dataset. Note that there are no tails and long hills in the variable distributions.

Data Analysis:

The data analysis section is divided into four sections:

- 1) **By variable:** A demonstration of the impact of variables on the happiness score
- 2) **By continent:** An analysis based on continent level
- 3) **By country:** An analysis based on country level
- 4) **Further analysis:** An analysis of newly created features

Data Analysis by Variables

Correlation Plot

During this analysis correlation between the happiness score and other variables is calculated.

| | Happiness Score | Economy (GDP per Capita) | Family | Health (Life Expectancy) | Freedom | Trust (Government Corruption) | Generosity | Dystopia Residual |
|-------------------------------|-----------------|--------------------------|----------|--------------------------|----------|-------------------------------|------------|-------------------|
| Happiness Score | 1.000000 | 0.780966 | 0.740605 | 0.724200 | 0.568211 | 0.395199 | 0.180319 | 0.530474 |
| Economy (GDP per Capita) | 0.780966 | 1.000000 | 0.645299 | 0.816478 | 0.370300 | 0.307885 | -0.010465 | 0.040059 |
| Family | 0.740605 | 0.645299 | 1.000000 | 0.531104 | 0.441518 | 0.205605 | 0.087513 | 0.148117 |
| Health (Life Expectancy) | 0.724200 | 0.816478 | 0.531104 | 1.000000 | 0.360477 | 0.248335 | 0.108335 | 0.018979 |
| Freedom | 0.568211 | 0.370300 | 0.441518 | 0.360477 | 1.000000 | 0.493524 | 0.373916 | 0.062783 |
| Trust (Government Corruption) | 0.395199 | 0.307885 | 0.205605 | 0.248335 | 0.493524 | 1.000000 | 0.276123 | -0.033105 |
| Generosity | 0.180319 | -0.010465 | 0.087513 | 0.108335 | 0.373916 | 0.276123 | 1.000000 | -0.101301 |
| Dystopia Residual | 0.530474 | 0.040059 | 0.148117 | 0.018979 | 0.062783 | -0.033105 | -0.101301 | 1.000000 |

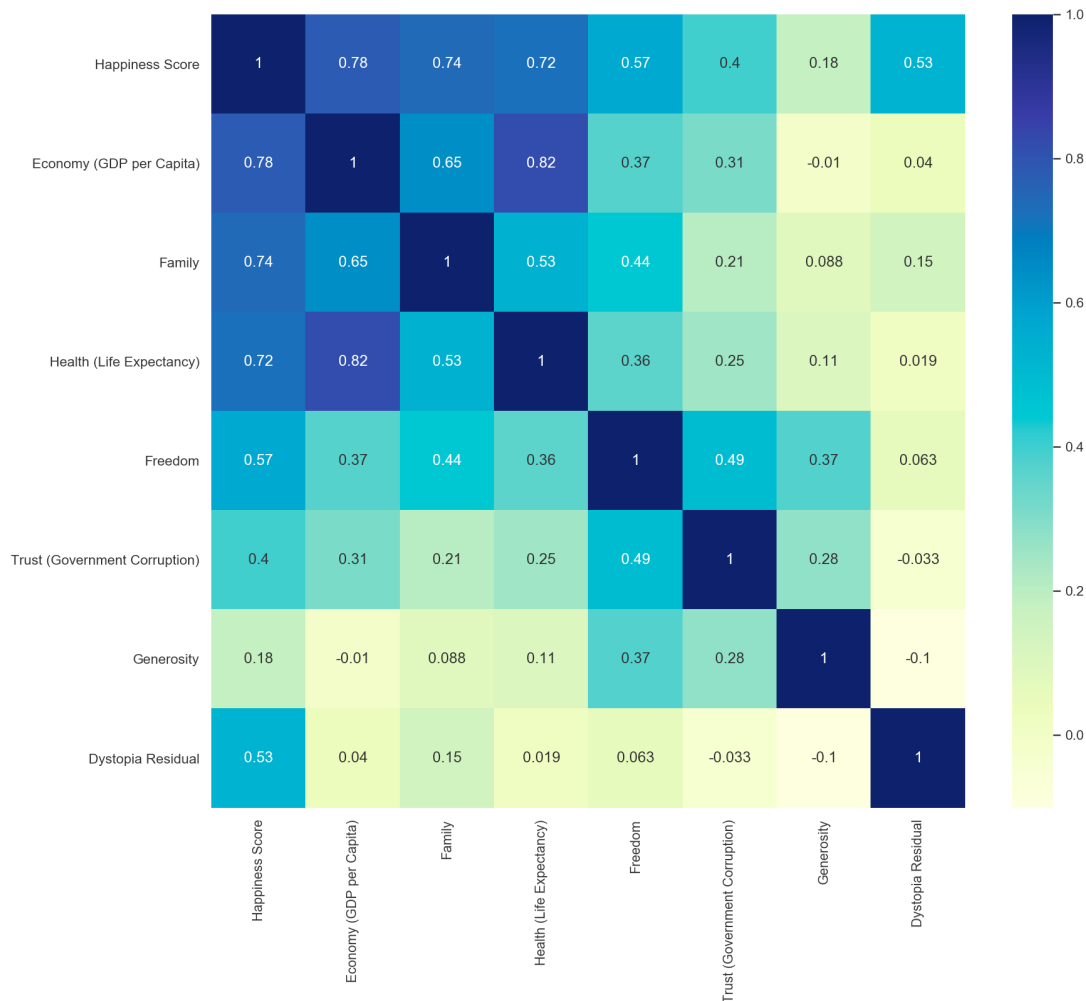


Figure: Correlation Plot

From the figure and the table above it can be concluded that all variables are positively correlated with the Happiness Score and it is primarily correlated with Economy (0.78). It is also highly correlated with Family (0.74) and Health (0.72). It is important to note that Generosity (0.18) is the least correlated with Happiness score; therefore, it does not impact the Happiness Score as much as the other variables. Freedom (0.57), Trust (0.5) and Dystopia (0.53) are moderately correlated with Happiness Score.

Correlation in Pair Graphs

During this analysis correlations are investigated using pair graphs.

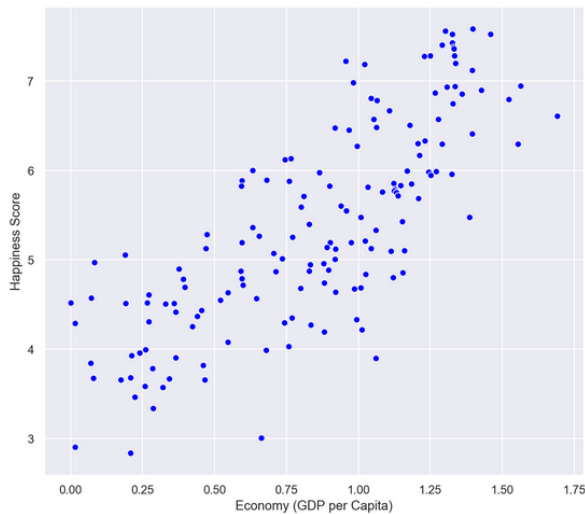


Figure: Correlations in pair graphs

Impact of Variables on Happiness Score

Below please find further investigation of variables' impact on happiness score.

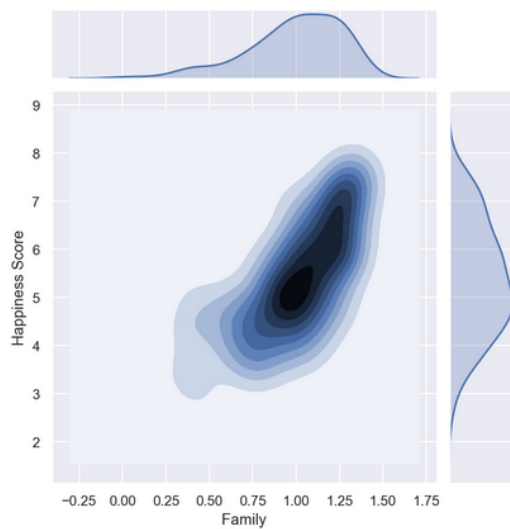
a) Happiness Score vs. Economy



Happiness Score is positively correlated with the economy of nations.

Figure: Happiness Score vs. Economy

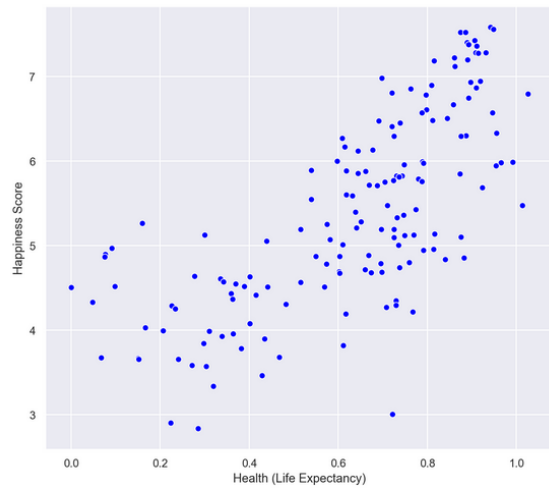
b) Happiness Score vs. Family



Family and Happiness Score are positively correlated

Figure: Happiness score vs. Family

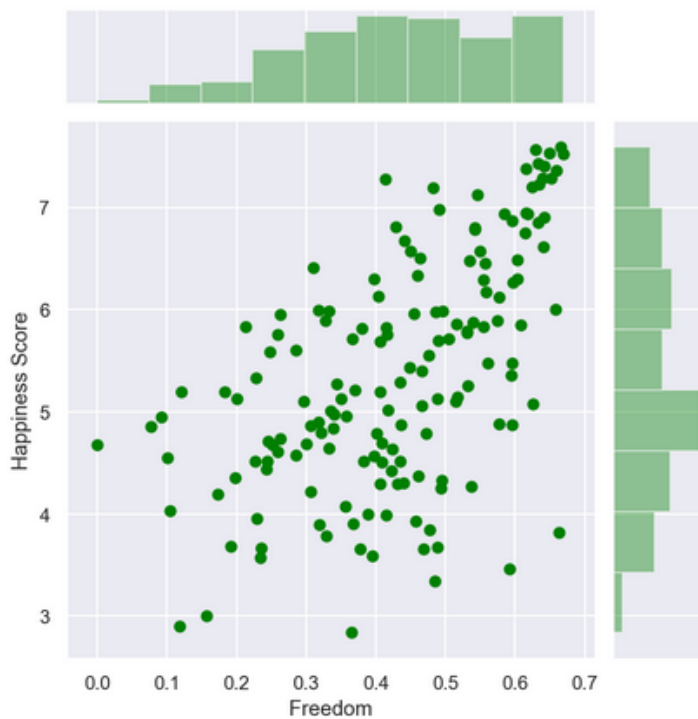
c) Happiness Score vs. Health (Life Expectancy)



Happiness Score is correlated with health.

Figure: Happiness Score vs. Health

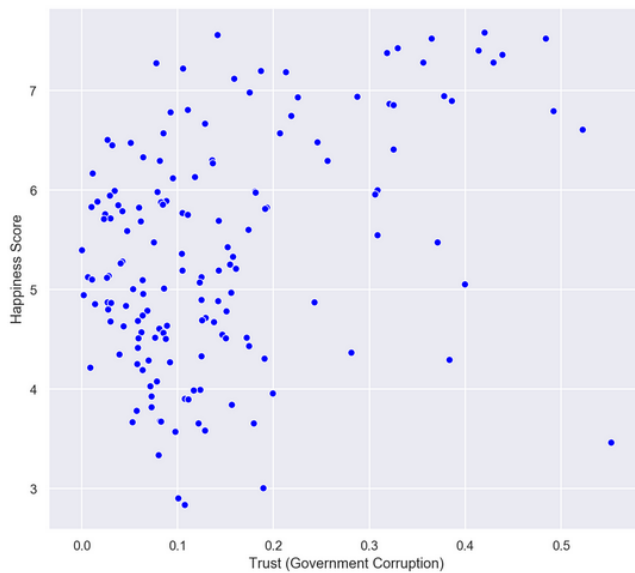
d) Happiness Score vs Health (Life Expectancy)



Happiness score and freedom are slightly positive correlated.

Figure: Happiness Score vs. Happiness

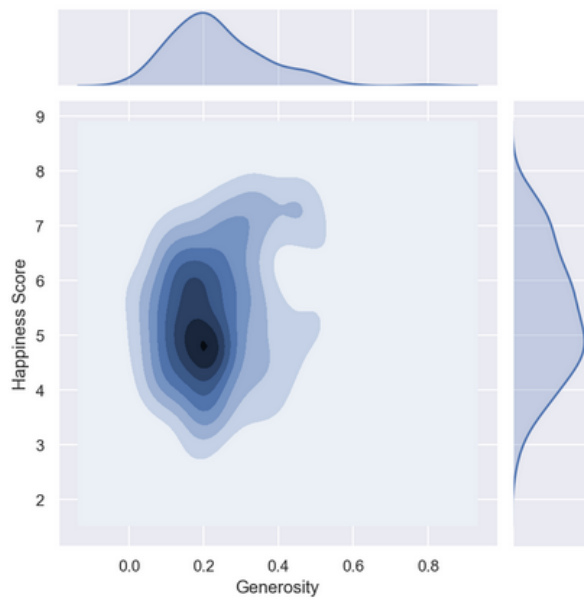
e) Happiness Score vs. Trust (Government Corruption)



Happiness score and Trust are slightly positive correlated.

Figure: Happiness Score vs. Trust

f) Happiness Score vs. Generosity



Happiness score and Generosity are positively correlated.

Figure: Happiness Score vs Generosity

g) Happiness Score vs. Dystopia

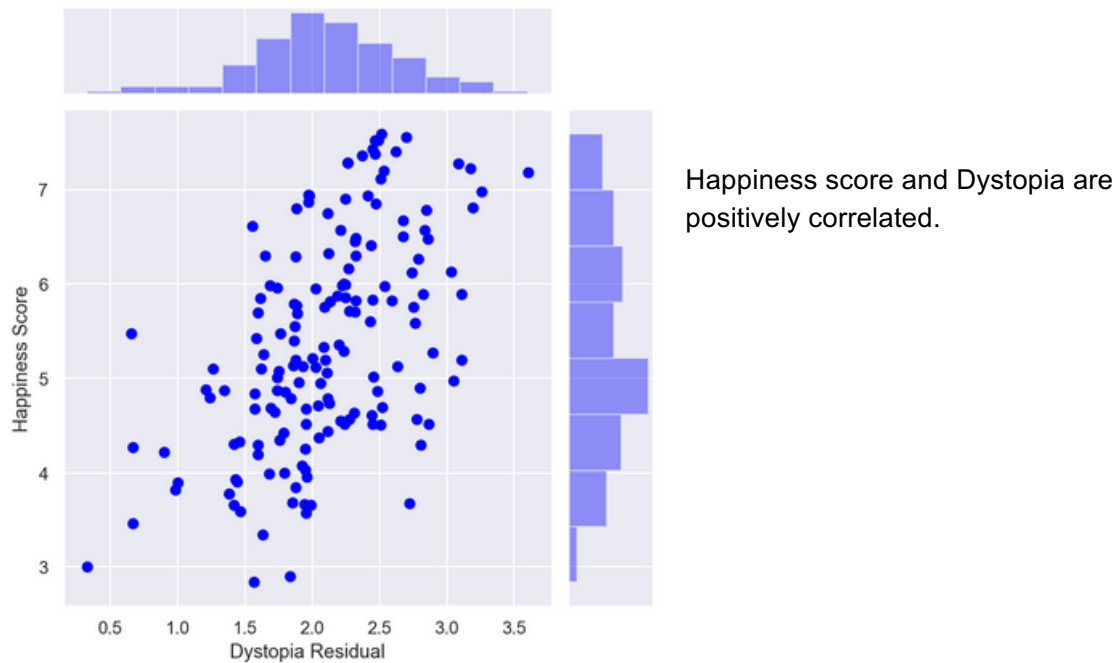


Figure: Happiness Score vs. Dystopia Residual

Data Analysis By Continent

Happiness Score by Region

The “world happiness” data set was analyzed by the region. The main idea of this section of our research is to identify the correlation of the dependent variable, independent variables by the region and verify top continents by happiness score.

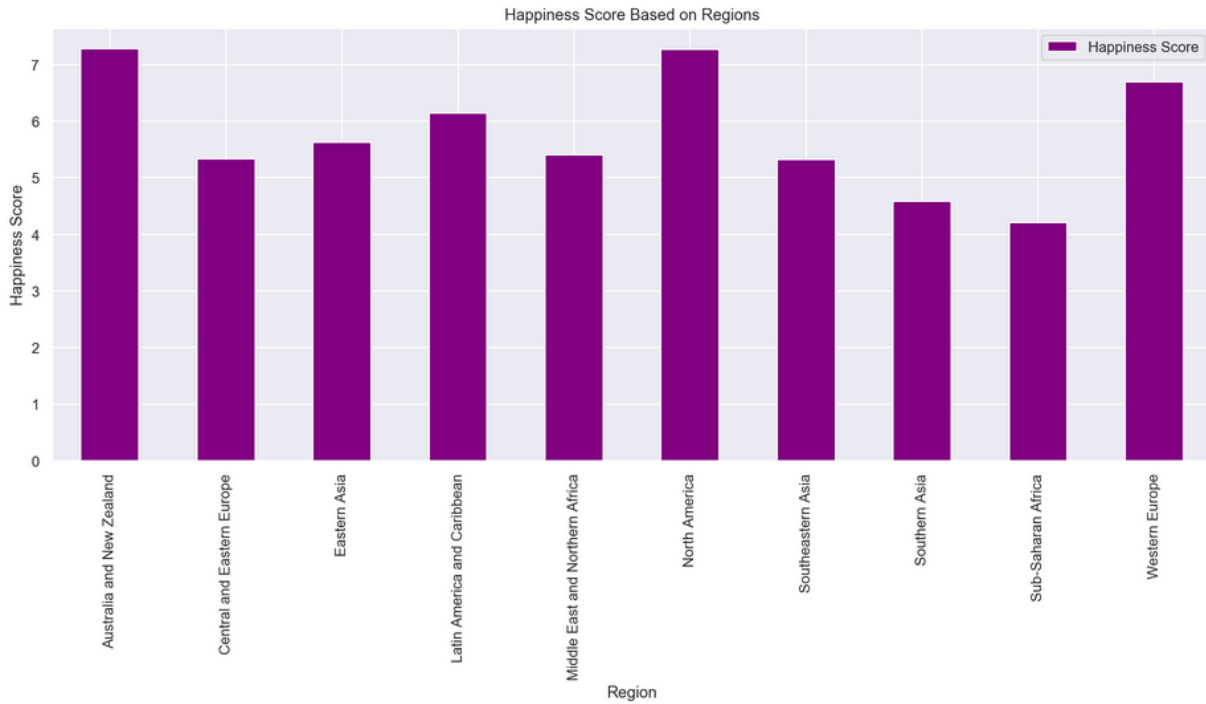


Figure: Happiness Score based on the Region.

| Happiness Score | |
|---------------------------------|----------|
| Region | |
| Australia and New Zealand | 7.285000 |
| Central and Eastern Europe | 5.332931 |
| Eastern Asia | 5.626167 |
| Latin America and Caribbean | 6.144682 |
| Middle East and Northern Africa | 5.406900 |
| North America | 7.273000 |
| Southeastern Asia | 5.317444 |
| Southern Asia | 4.580857 |
| Sub-Saharan Africa | 4.202800 |
| Western Europe | 6.689619 |

The graph and table show that the happiest region is Australia and New Zealand with 7.285 happiness score and Sub-Saharan Africa has the lowest happiness score with 4.202.

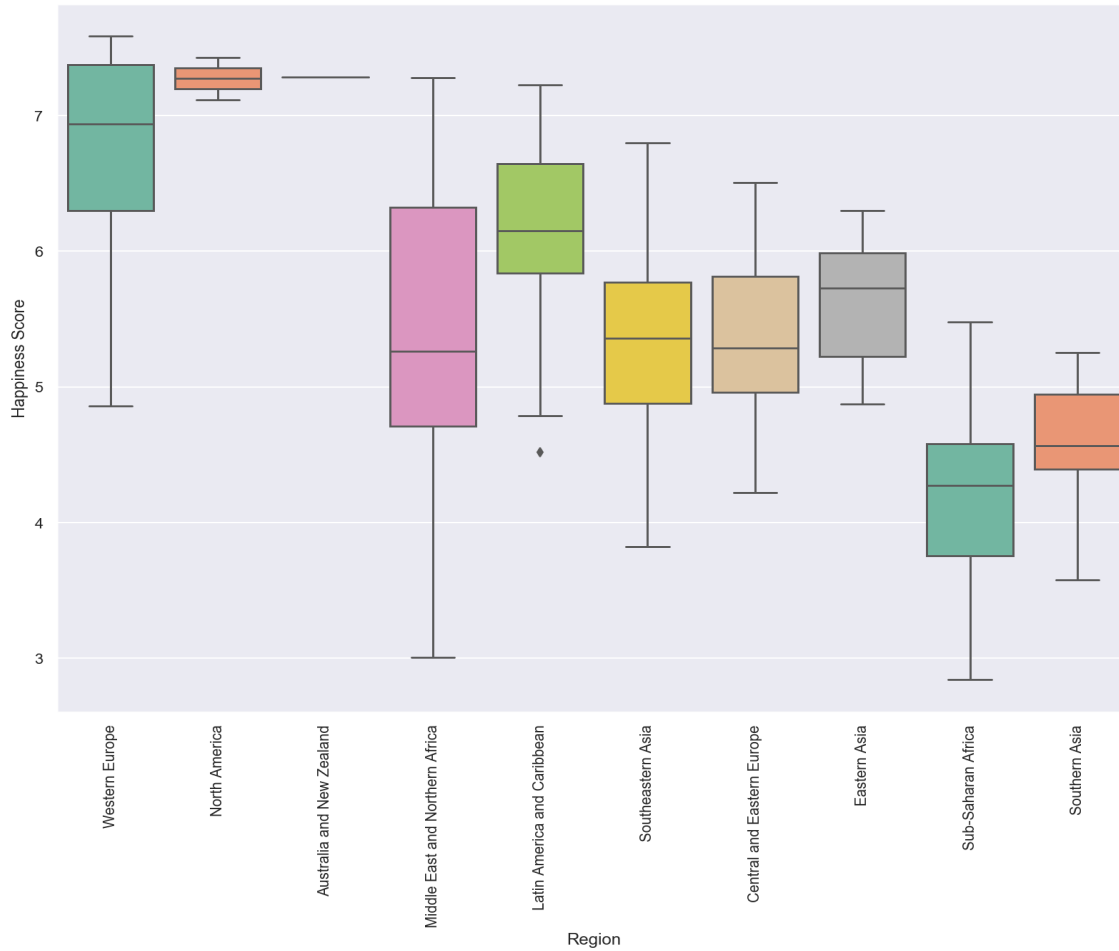


Figure: Happiness Score by Region

The boxplot shows that European Countries and Australia are the top two Happiest Regions. America is another “happy region” that shows high a Happiness Index. The happiness score of Asian and African countries is much lower, that is why Asian and African countries don’t stand out at the top positions in terms of the Happiness Index.

Bubble Chart (3D)

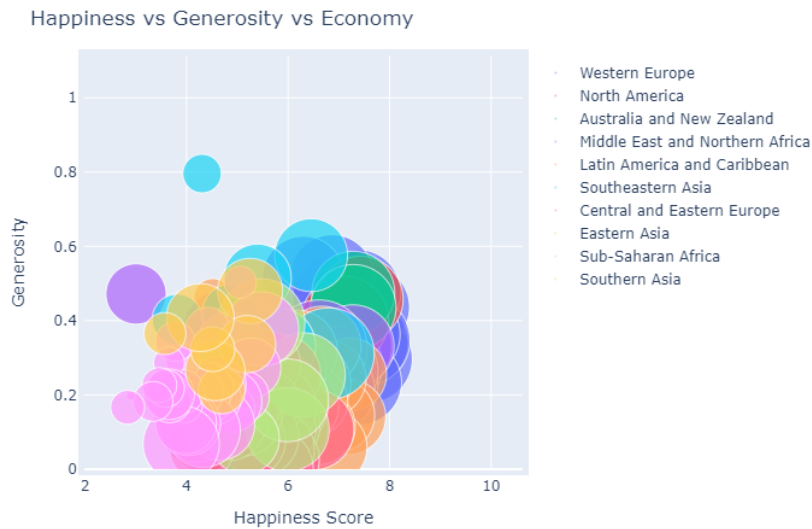


Figure: Bubble Chart for Happiness, Generosity and Economy

The bubble chart describes the relationships between happiness, generosity and economy (size of the bubble). Happiness score is directly proportional to the economy; however, when the economy score is high, the generosity variable is not as proportional.

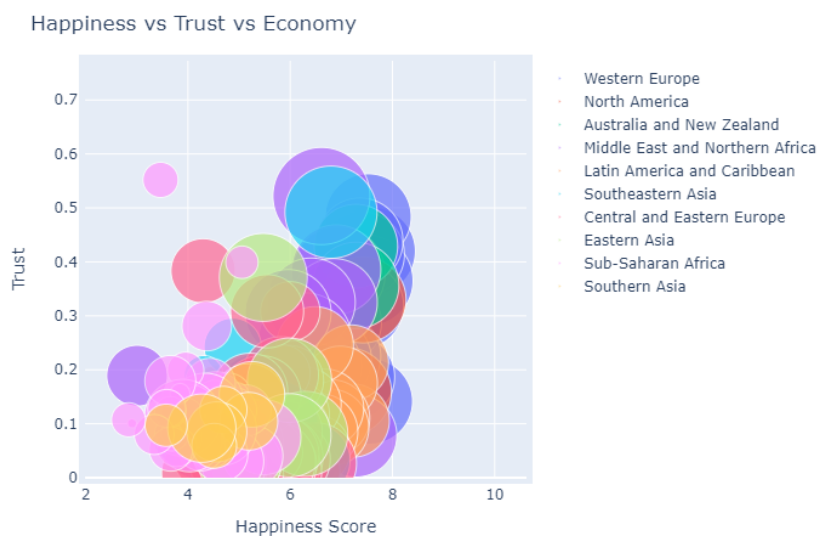


Figure: Bubble Chart for Happiness, Trust and Economy

The bubble chart describes the relationships between happiness, trust and economy (size of the bubble). Economy is the major contributor to high happiness index, whereas the trust variable is not as proportional to the happiness score.

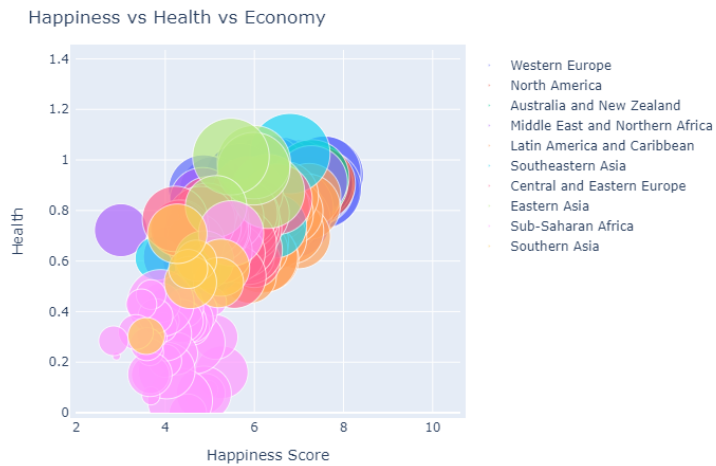


Figure: Bubble Chart for Happiness, Trust and Economy

The bubble chart describes the relationships between happiness, health and economy (size of the bubble). Happiness, health and economy are directly proportional. The life expectancy increases when the economy has a higher score.

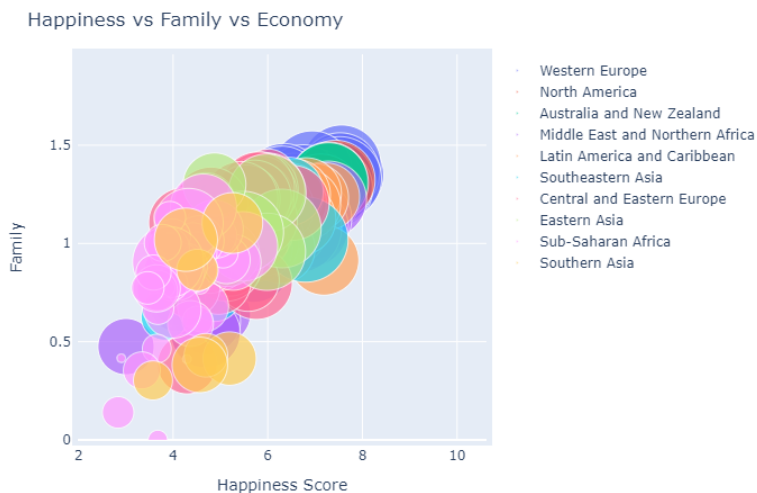


Figure: Bubble Chart for Happiness, Family and Economy

The bubble chart describes the relationships between happiness, family and economy (size of the bubble). Family Satisfaction ratings increase when the Happiness Score increases. It means that those two variables have a direct relationship with each other.

Pie Chart

Regions

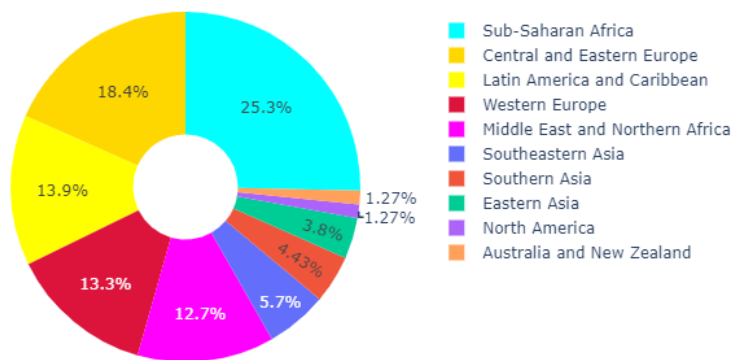


Figure: The Percentage of Countries in Each Region

There are only two countries from North America (the USA and Canada), and Australia (Australia, and New Zealand) Regions. The highest number of countries are from Sub-Saharan, and Central and Eastern Europe Regions with 40 and 29 countries, respectively.

Data Analysis by Country

We have analyzed the happiness score by different regions. This part of our analysis will showcase the data more granularly: happiness score by country. We have identified the top 10 happiest countries from the represented dataset. Also, as the analysis shows that the economy (0.78), family (0.74) and health (0.72) are the main variables that have a high correlation with the happiness score, we will analyze the relationship of the variables in each country.

Top 10 Happiest Countries

| Country | Region | Happiness Rank | Happiness Score | Standard Error | Economy (GDP per Capita) | Family | Health (Life Expectancy) | Freedom | Trust (Government Corruption) | Generosity | Dystopia Residual |
|-------------|---------------------------|----------------|-----------------|----------------|--------------------------|---------|--------------------------|---------|-------------------------------|------------|-------------------|
| Switzerland | Western Europe | 1 | 7.587 | 0.03411 | 1.39651 | 1.34951 | 0.94143 | 0.66557 | 0.41978 | 0.29678 | 2.51738 |
| Iceland | Western Europe | 2 | 7.561 | 0.04884 | 1.30232 | 1.40223 | 0.94784 | 0.62877 | 0.14145 | 0.43630 | 2.70201 |
| Denmark | Western Europe | 3 | 7.527 | 0.03328 | 1.32548 | 1.36058 | 0.87464 | 0.64938 | 0.48357 | 0.34139 | 2.49204 |
| Norway | Western Europe | 4 | 7.522 | 0.03880 | 1.45900 | 1.33095 | 0.88521 | 0.66973 | 0.36503 | 0.34699 | 2.46531 |
| Canada | North America | 5 | 7.427 | 0.03553 | 1.32629 | 1.32261 | 0.90563 | 0.63297 | 0.32957 | 0.45811 | 2.45176 |
| Finland | Western Europe | 6 | 7.406 | 0.03140 | 1.29025 | 1.31826 | 0.88911 | 0.64169 | 0.41372 | 0.23351 | 2.61955 |
| Netherlands | Western Europe | 7 | 7.378 | 0.02799 | 1.32944 | 1.28017 | 0.89284 | 0.61576 | 0.31814 | 0.47610 | 2.46570 |
| Sweden | Western Europe | 8 | 7.364 | 0.03157 | 1.33171 | 1.28907 | 0.91087 | 0.65980 | 0.43844 | 0.36262 | 2.37119 |
| New Zealand | Australia and New Zealand | 9 | 7.286 | 0.03371 | 1.25018 | 1.31967 | 0.90837 | 0.63938 | 0.42922 | 0.47501 | 2.26425 |
| Australia | Australia and New Zealand | 10 | 7.284 | 0.04083 | 1.33358 | 1.30923 | 0.93156 | 0.65124 | 0.35637 | 0.43562 | 2.26646 |

Economy Analysis based on Countries

Economy (GDP per Capita)

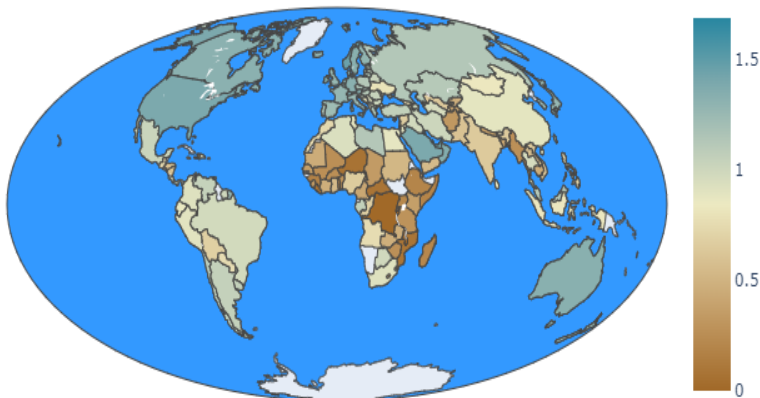


Figure: Economy Index by Country

The figure shows that African countries have the lowest economy index, whereas Australia, Europe and North American countries have the highest score.

Family Analysis based on Countries

Family

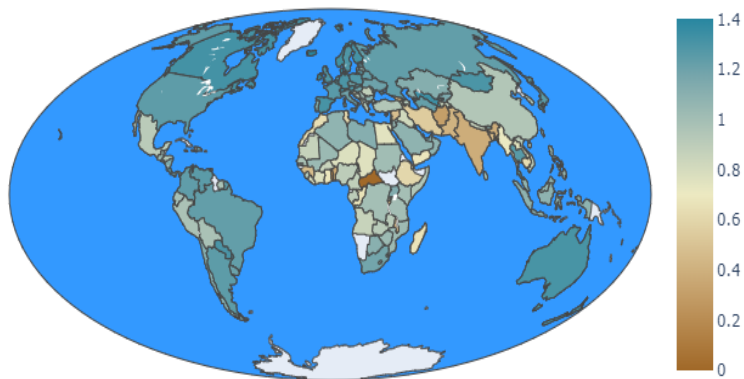


Figure: Family Index by Country

The figure shows that some of the African countries have a low family index; however, the number of countries with a low index is not as significant as at the economy figure. Australia, Europe and North American have a high family index.

Health Analysis based on Countries

Health (Life Expectancy)

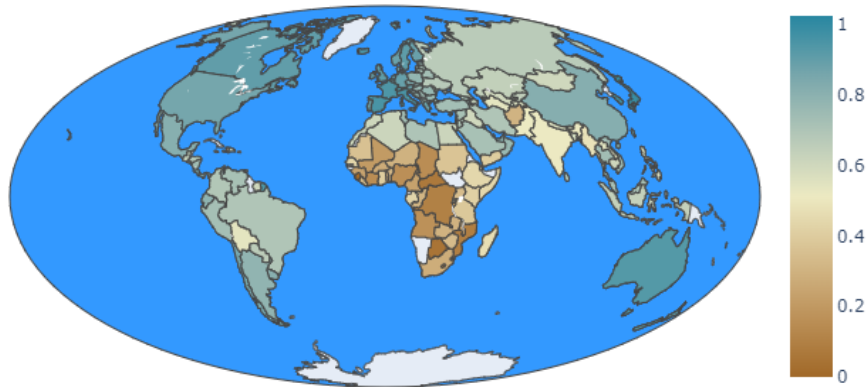


Figure: Healthy Index by Country

The health index is really low in African countries as well as in some of the South Asia countries. Australia, Europe and North American have high health ratings.

Further Analysis

The conducted research shows that the “Happiness Score” is the sum of Economy, Family, Health, Freedom, Trust, Generosity and Dystopia Residual variables. Therefore, for further analysis, we have calculated the percentage of each variable to identify the percentage impact of 6 factors to the happiness score.

To calculate the percentage of the impact of all the factors we have created the additional features.

Created Features:

- Impact Percentage of Generosity on happiness score - impact_gen_happ_percent
- Impact Percentage of Freedom on happiness score - impact_freedom_happ_percent,
- Impact Percentage of Health on happiness score - impact_health_happ_percent,
- Impact Percentage of Family on happiness score - impact_family_happ_percent,

- Impact Percentage of Economy on happiness score -
impact_economy_happ_percent,
- Impact Percentage of Trust on happiness score -impact_trust_happ_percent

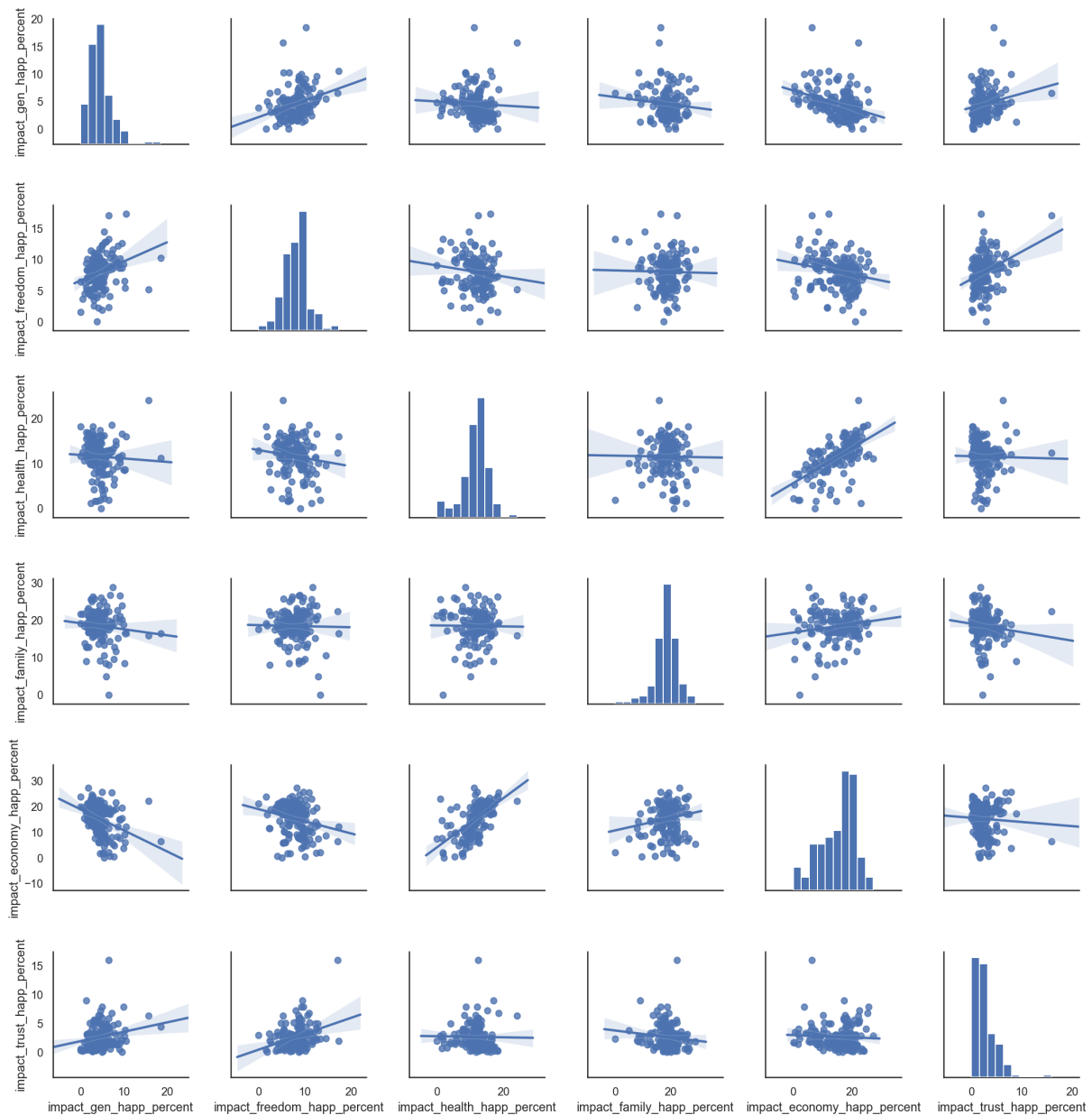


Figure: Relationship of the New Features

1. Impact of Generosity and impact of Economy on Happiness score are inversely related. Countries with strong economy don't always have generous people and we can see proof of this fact in our analysis.

2. The high economy rating shows that freedom isn't the main factor and contributor to the happiness score. In countries with a high economy index, people are happy regardless of the freedom level.
3. The impact of trust or impact of family has very little correlation with other factors.

Impact of each Variable & Top 5 countries:

Out[192]:

| | Country | Happiness Score | Economy (GDP per Capita) | Generosity | Impact_gen_happ_percent | Impact_economy_happ_percent |
|-----|-------------|-----------------|--------------------------|------------|-------------------------|-----------------------------|
| 128 | Myanmar | 4.307 | 0.27108 | 0.79588 | 18.478756 | 6.293940 |
| 155 | Syria | 3.006 | 0.66320 | 0.47179 | 15.694943 | 22.062542 |
| 144 | Cambodia | 3.819 | 0.46038 | 0.40359 | 10.567950 | 12.054988 |
| 118 | Haiti | 4.518 | 0.26673 | 0.46187 | 10.222886 | 5.903718 |
| 152 | Afghanistan | 3.575 | 0.31982 | 0.36510 | 10.212587 | 8.946014 |

These are the Top 5 countries which have high impact of Generosity on their happiness. and it shows a surprising factor that even their GDP per capita is low, happiness score is low .. still they have high importance for generosity

Table: Top countries with the highest impact of a Generosity score on a happiness rank

Out[193]:

| | Country | Happiness Score | Impact_health_happ_percent | Health (Life Expectancy) |
|-----|-----------|-----------------|----------------------------|--------------------------|
| 155 | Syria | 3.006 | 24.016301 | 0.72193 |
| 71 | Hong Kong | 5.474 | 18.510778 | 1.01328 |
| 133 | Bulgaria | 4.218 | 18.171882 | 0.76649 |
| 101 | Greece | 4.857 | 18.162034 | 0.88213 |
| 102 | Lebanon | 4.839 | 17.348006 | 0.83947 |

Top 5 Countries having most impact of health (life expectancy) on happiness.

Table: Top countries with the highest impact of a Health score on a happiness rank

Out[194]:

| | Country | Happiness Score | impact_freedom_happ_percent | Freedom |
|-----|--------------------------|-----------------|-----------------------------|---------|
| 144 | Cambodia | 3.819 | 17.346426 | 0.66246 |
| 153 | Rwanda | 3.465 | 17.085426 | 0.59201 |
| 154 | Benin | 3.340 | 14.505988 | 0.48450 |
| 147 | Central African Republic | 3.678 | 13.289560 | 0.48879 |
| 157 | Togo | 2.839 | 12.840085 | 0.36453 |

Top 5 Countries having most impact of freedom on happiness.

Table: Top countries with the highest impact of a Freedom score on a happiness rank

Out[195]:

| | Country | Happiness Score | impact_trust_happ_percent | Trust (Government Corruption) |
|-----|-------------------|-----------------|---------------------------|-------------------------------|
| 153 | Rwanda | 3.465 | 15.928139 | 0.55191 |
| 129 | Georgia | 4.297 | 8.920410 | 0.38331 |
| 27 | Qatar | 6.611 | 7.897141 | 0.52208 |
| 90 | Somaliland region | 5.057 | 7.895590 | 0.39928 |
| 23 | Singapore | 6.798 | 7.238894 | 0.49210 |

Top 5 Countries having most impact of Trust on Government on happiness.

Table: Top countries with the highest impact of a Trust score on a happiness rank

Conclusion

We have analyzed the happiness score in 158 countries. First of all, we have compared the variables of the data set and identify whether the factor are correlated or not. The hit map of the variables display that the least correlated variables are Freedom (0,57), Trust (0,4) and Generosity (0,18).

Also, the hit map showcases that there are a couple of factors that are highly correlated with happiness score, such as Economy (0,78), Health (0,72) and Family (0,74). Also, those three factors have significantly higher ratings in Western Europe, North America and Australia. That supports the fact that these regions are included in the top 3 continents by happiness score. Central and Eastern Europe, Latin America and Caribbean regions are following the top three regions with a relatively high index of happiness.

The 10 top countries are represented by small or medium-sized countries. Most of the countries from the list located in Western Europe. Beyond the first ten, the geography immediately becomes more varied, with the second 10 including countries from four of the nine global regions. In the top 10 countries, we see the following factors: Economy, Family and Health have the highest score, whereas generosity or trust ratings not necessarily high. For example, Iceland is the second happiest country; however, the level of trust is the lowest out of 10 represented countries.

The interesting fact that was identified by conducting the data analysis. The happiness score in the data set is the sum of Economy, Family, Health, Freedom, Trust, Generosity and Dystopia Residual variables. As a result, we have identified the contribution of the 6 factors to the happiness score.

The conducted research shows that the "Happiness Score" is the sum of Economy, Family, Health, Freedom, Trust, Generosity and Dystopia Residual variables. Therefore, for further analysis, we have calculated the percentage of each variable to identify the percentage impact of 6 factors to the happiness score. Here are the main observations that we wanted to highlight while analyzing the dataset and taking into account the contribution of each factor:

- Impact of Generosity and impact of Economy on Happiness score are inversely related. Countries that have significantly high economy don't necessarily mean that people are more generous. For instance, countries such as Haiti, Syria or Myanmar does not have a strong economy, however generosity is the core factor that contributed to sustaining the happiness score.
- High Economy or Health rating shows that freedom isn't the main factor and contributor to the happiness score. In countries with a high economy index, people are happy regardless of freedom level. Freedom factor is the main contributor to

the score of happiness for the following countries: Cambodia, Rwanda, Benin, Central African Republic and Togo.

The impact of trust or impact of family has very little correlation with other factors. However, trust has a slight significance as a contributor to the overall happiness score. Since trust is the main contributor to the happiness score of the countries with one of the highest happiness ranks, such as Qatar and Singapore.

References

- 1) https://s3.amazonaws.com/happiness-report/2015/WHR15_Sep15.pdf
- 2) <https://www.kaggle.com/unsdsn/world-happiness>
- 3) <https://www.kaggle.com/unsdsn/world-happiness#2015.csv>
- 4) <https://worldhappiness.report/ed/2019/>
- 5) <https://www.ctvnews.ca/lifestyle/canada-falls-to-ninth-in-world-happiness-report-finland-first-again-1.4343774>
- 6) <https://www.kaggle.com/imrandude/python-cluster-analysis-world-happiness-report>