World Happiness Report

Professor: Alberto Sanz

Statistics and Data Analysis

IE University

December, 2019

Nicole Caballero: ncaballero.ieu2018@student.ie.edu

Mario Bevilacqua: mbevilacqua.ieu2018@student.ie.edu

Pierre De St Vincent: pdesaintvinc.ieu2018@student.ie.edu

Ghali Chraibi: gchraibi.ieu2018@student.ie.edu

Abstract: Our report, World Happiness Report, was made to discover: What makes countries happier than others? To prove our research question we used a Kaggle dataset, World Happiness 2017, and merged it with an Excel file to include government type and continent into the dataset. This dataset includes information about the World Happiness statistics in 2017, influenced by factors such as GDP, life expectancy, government type, freedom, generosity, government trust and continents. During the study, we defined our unit of analysis, countries, and carefully selected the factors that will affect the happiness score of each country. After performing our study, we found out that the major factors affecting the happiness score were GDP per capita and life expectancy, but other factors such as generosity, freedom and government type would also affect the score.

KeyWords: Happiness Score, Correlation, Countries, Population, and Confidence.
Word Count: 2382 words

TABLE OF CONTENTS

INTRODUCTION

The world is a rapidly changing place. Generally speaking, while some countries are constantly improving their living conditions, others fall into despair. Political tensions, economic incidents and civil rights violations are a few of the numerous reasons countries can be classified as having a large discontented population. On the other hand, other countries with distinct situations have enjoyed new opportunities to grow both economically and socially. However, the disparities between these two recurring and evolving situations raises the question as to how these differences affect countries overall happiness. How happy are people around the world and how our living conditions affect us has always been something that has come to our minds. Therefore, the research question we will be analyzing in this report is “What makes some countries happier than others?”, based on 2017 statistics. This will allow us to measure how different aspects of people's lives around the world affect their happiness and satisfaction levels, and to discover what factors lead to a happier and fulfilling life.

THEORY AND HYPOTHESIS

The phenomena to be studied is how external factors surrounding individuals influence their state of happiness, measured by the happiness score that is our dependent variable for this report, and compared with our unit of analysis, countries. Thus, a number of hypotheses have to be tested and are stated as:

1. **Continent**

   H0: The mean happiness score of the continents are all equal.
   H1: Depending on the continent, people are differently happy (At least one the continents has different happiness score mean than the other continents).

2. **Government**

   H0: The mean happiness score is the same for all types of government.
   H1: People feel happier or not depending under which government type they are faced (At least one mean happiness score is different within each government type).

3. **GDP per Capita**

   H0: The 25% poorest countries (within Q1 of GDP per capita) have a higher or equal happiness score mean than the rest.
   H1: Poorer the country, less happier their inhabitants are (Countries within Q1 of GDP have a smaller happiness mean score than the rest).

4. **Life Expectancy**

   H0: Life expectancy score is correlated with the mean happiness score
   H1: Life expectancy score is not associated with the mean happiness score.

5. **Freedom**

   H0: Countries within the 25% of lower degree of freedom score (the first quartile Q1) have a bigger than or equal happiness mean score than the rest.
   H1: Less degree of freedom leads to a lower happiness score (Countries within the first quartile (Q1) of freedom have a smaller happiness mean score than the rest).

6. **Generosity**

H0: Countries within the 25% of the lower level generosity scores (Q1) have a bigger than or equal happiness mean score than the rest.

H1: Countries within Q1 of generosity have a smaller happiness mean score than the rest.

7. **Government Trust**

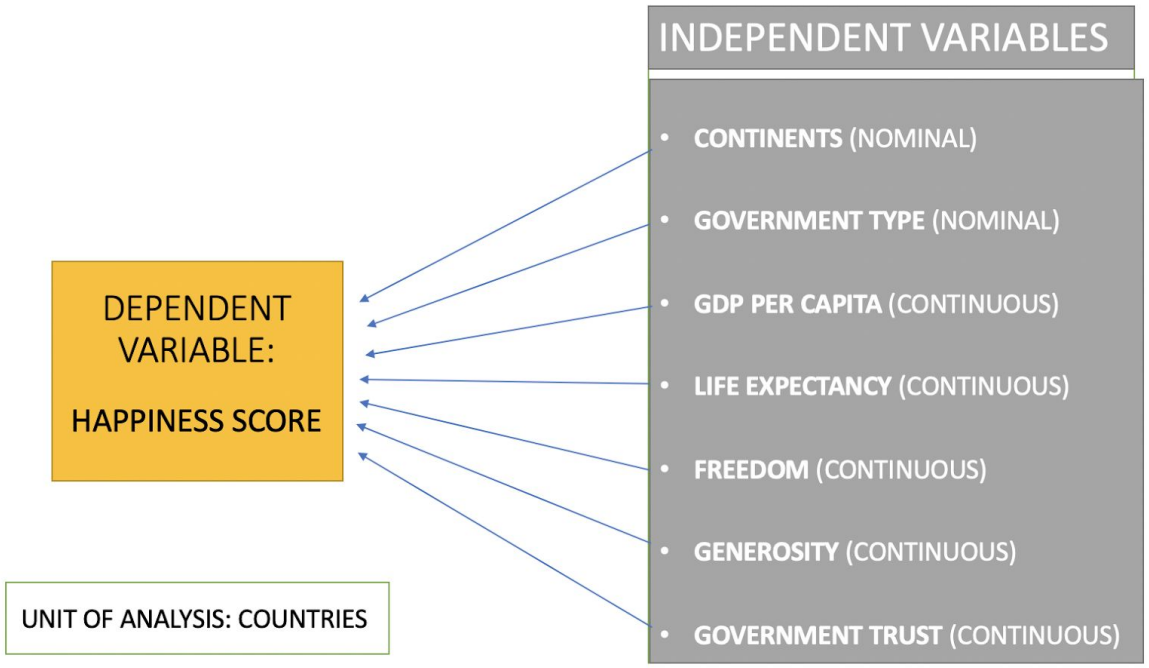
H0: Countries within the Q4 of government trust have a smaller or equal happiness mean score than the rest.

H1: The higher the government trust, happier are the people (countries within the Q4 of government trust have a higher happiness mean score than the rest.)

INDEPENDENT AND DEPENDENT VARIABLES

From the statement of the hypothesis clearly explained in the previous section, the independent variables that play the role of “measurements” instruments of the factors in the “explanation” or prediction of the dependent variable are shown in *Figure 16.*

Figure 16. Dependent and Independent Variables and Unit of Analysis.



DATA

To test our hypothesis, and analyze the dependent and independent variables, we used a Kaggle dataset made by “Sustainable Development Solutions Network”, and named World Happiness Report based on the study of 153 countries, but more specifically we used World Happiness Report 2017-csv, we also used an Excel file in which we added two more columns to study, that where the Government Type and the Continent of each of the countries in the dataset. The excel file and the dataset were merged to better analyze it. To be sure about the type of government each country was experiencing, we used “Wikipedia Democracy Index”, which was based on 4 main government types: Full Democracy, Flawed Democracy, Hybrid Regime and Authoritarian Regime.

For better data manipulation we create identifiers for each of the government types and continents, and were assigned to the countries studied as shown in *Chart 1* and *Chart 2*.

Chart 1. Government Type and Identifier.

GOVERNMENT TYPE	IDENTIFIER
Full Democracy	1
Flawed Democracy	2
Hybrid Regime	3
Authoritarian Regime	4

Chart 2. Continent and Identifier.

GOVERNMENT TYPE	IDENTIFIER
America	1
Europe	2
Asia	3
Africa	4

RESULTS

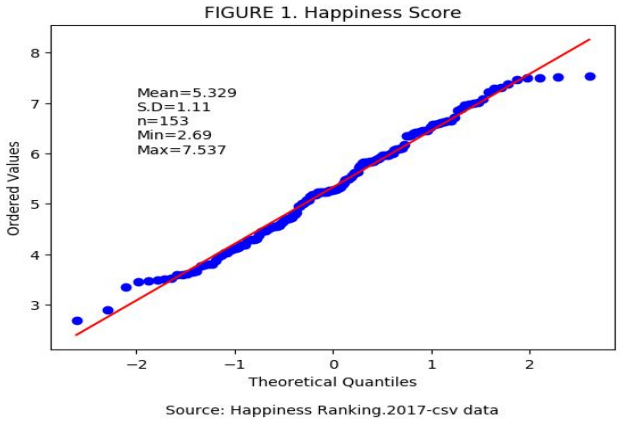
Before testing our previously stated hypothesis, we described each variable with their respective calculation and we found the following results:

Dependent Variable Description:

- **Happiness Score**

As we can see in Figure 1, the population is normal since the majority of dots are at the line, we also got those numbers after describing it.

Figure 1. Happiness Score Probability Plot



Independent Variables Description:

Continent

As we can see in Chart 3, we got the following percentages in respect of the distribution of each country into their respective continent.

Chart 3. Continent Distribution

CONTINENT	AMOUNT (%)
America	15.69
Europe	26.14
Asia	29.41
Africa	28.76

Government Type

As seen in Chart 4, we got the following results based on the distribution of the government type of each country.

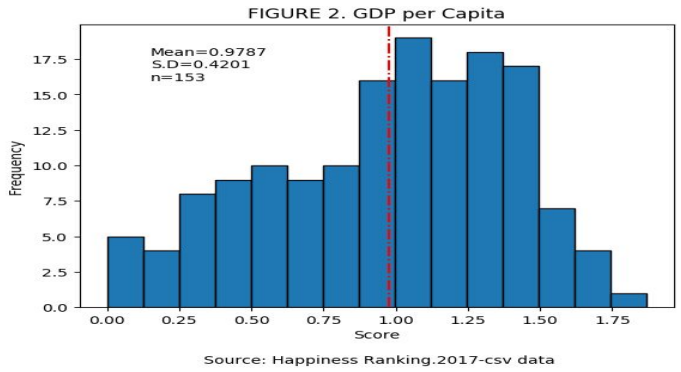
Chart 4. Government Type Distribution

GOVERNMENT TYPE	AMOUNT (%)
Full Democracy	11.77
Flawed Democracy	33.33
Hybrid Regime	24.84
Authoritarian Regime	30.06

GDP per Capita

As we can see in Figure 2, the population is normally distributed, and we also got those numbers after describing it.

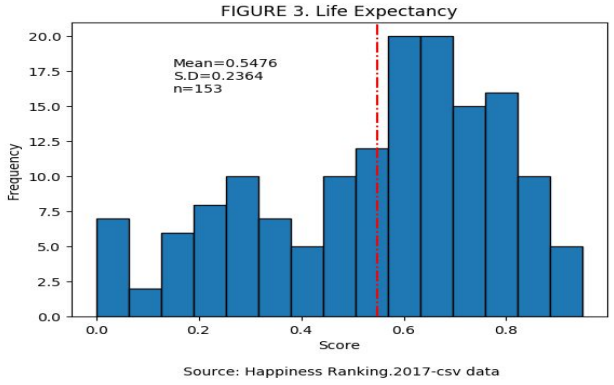
Figure 2. GDP per Capita Histogram



Life Expectancy

As we can see in Figure 3, the population is not normally distributed but we assume normality for the sake of the report, and we also got those numbers after describing it.

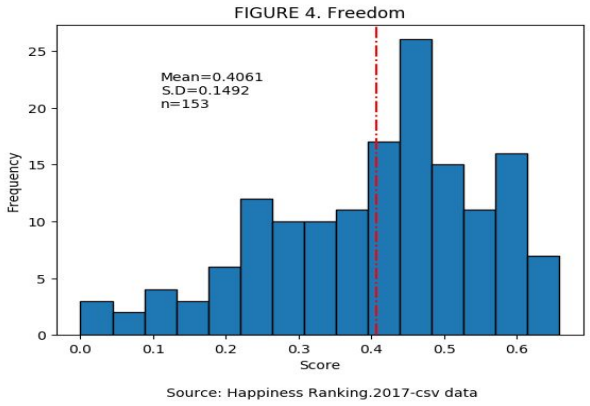
Figure 3. Life Expectancy Histogram



Freedom

As we can see in Figure 4, the population is not normally distributed, but we assume normality for the sake of the report, and we also got those numbers after describing it.

Figure 4. Freedom Histogram



Generosity

As we can see in Figure 5, the population is normally distributed, but in Figure 6 we see there are some outliers in generosity scores, and we also got those numbers after describing it.

Figure 5. Generosity Histogram.

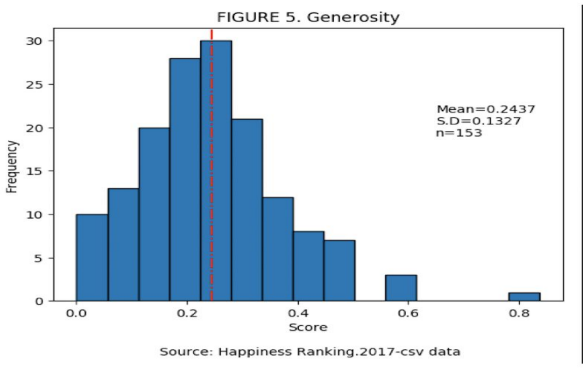
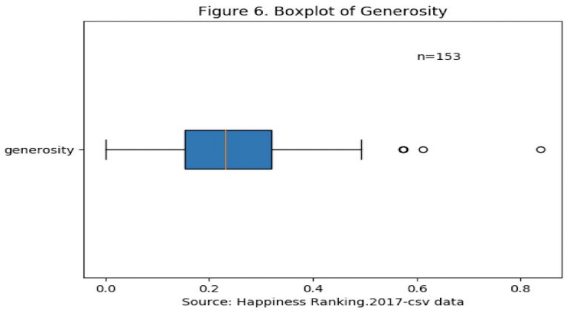


Figure 6. Generosity Boxplot.



Government Trust

As we can see in Figure 7, the population is not normally distributed and we assume normality for the sake of the report, we also see in Figure 8 multiple outliers in government scores, and we also got those numbers after describing it.

Figure 7. Government Trust Histogram

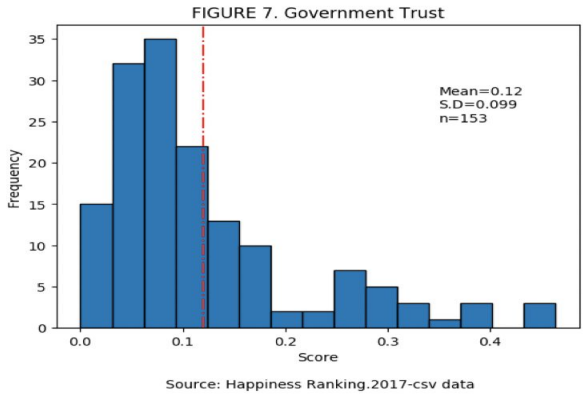
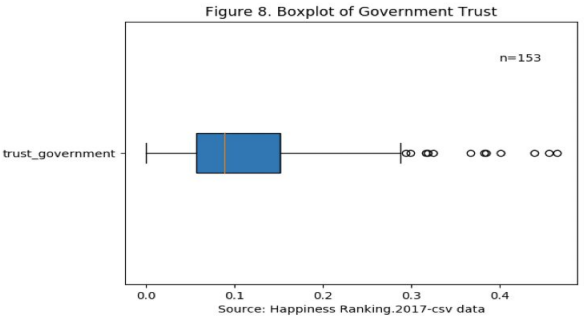


Figure 8. Government Trust Boxplot



After the previously stated description of variables, we are going to show the results from the hypothesis explained and stated in the Theory and Hypothesis section.

Results Hypothesis Testing:

1. **Continent**

The population variances are similar, so homoscedasticity is met, Africa and Asia has a normal population and we assume normality in Europe and America to continue testing ANOVA.

We reject the null hypothesis with a 95% confidence level, since the majority of the mean happiness score of the continents are different from each other. The only continents that have the similar mean happiness scores are America (1) and Europe (2) under the Post Hoc comparison with 95% confidence.

Figure 9. Point Plot Continent and Happ Score

Chart 5. Post Hoc comparison Continent and Happ Score.

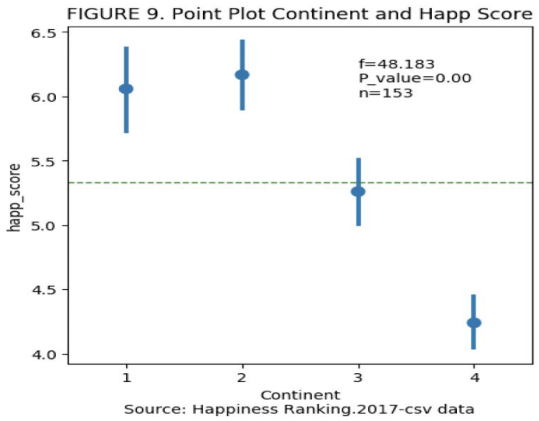


Chart 5. Post Hoc comparison Continent and Happ Score

group1	group2	stat	pval	pval_corr	reject
1	2	-0.5135	0.6094	1.0	False
1	3	3.7113	0.0004	0.0025	True
1	4	9.9076	0.0	0.0	True
2	3	4.8457	0.0	0.0	True
2	4	11.5728	0.0	0.0	True
3	4	6.1094	0.0	0.0	True

2. Government Trust

The population variances are different, so we assume Homoscedasticity, we also assume normality in the government type distribution in order to continue testing ANOVA.

We reject the null hypothesis with a 95% confidence level, since the majority of the mean happiness score of the continents are different from each other. The only continents that have the similar mean happiness scores are Hybrid Regime (3) and Authoritarian Regime (4) under the Post Hoc comparison with 95% confidence level.

Figure 10. Point Plot Government Type and Happ Score

Chart 6. Post Hoc comparison Government Type and Happ Score.

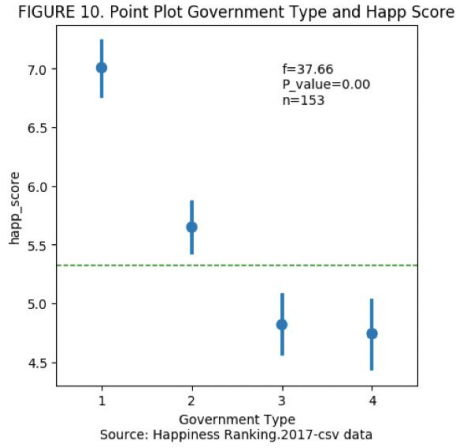


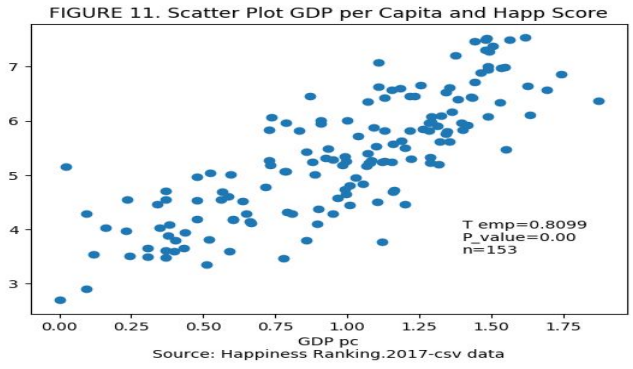
Chart 6. Post Hoc comparison Government Type and Happ Score

group1	group2	stat	pval	pval_corr	reject
1	2	6.5495	0.0	0.0	True
1	3	10.638	0.0	0.0	True
1	4	9.0357	0.0	0.0	True
2	3	4.7786	0.0	0.0	True
2	4	4.8698	0.0	0.0	True
3	4	0.379	0.7056	1.0	False

3. GDP per Capita

After performing a hypothesis testing of two populations we obtained that we reject the null hypothesis with a minimum alpha of approximately 0.00. On the other hand, we got that the GDP per capita and the happiness score are highly correlated as shown in Figure 11 with a correlation of 0.81.

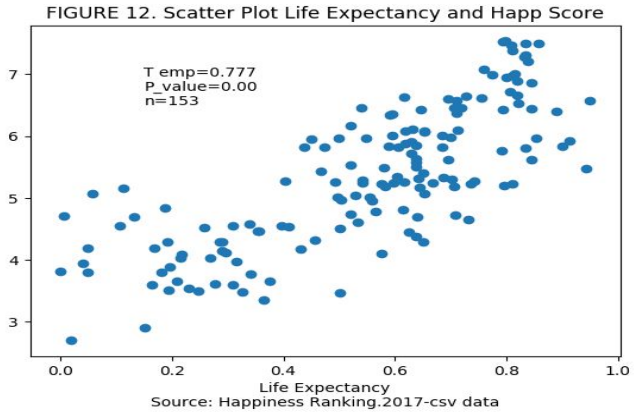
Figure 11. Scatter Plot GDP per Capita and Happ Score



4. Life Expectancy

After performing a correlation study, we were not able to reject the null hypothesis. As shown in Figure 12, we can see that the happiness score and the life expectancy is highly correlated with T empirical of 0.77 being 1 the maximum correlation score.

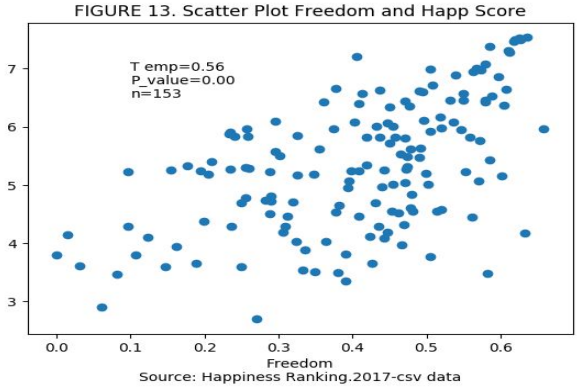
Figure 12. Scatter Plot Life Expectancy and Happ Score



5. Freedom

After performing a hypothesis testing of two populations we obtained that we reject the null hypothesis with a minimum alpha of approximately 0.00. On the other hand, we found that the freedom and happiness scores have a moderate correlation as shown in Figure 13 with a correlation score of 0.56.

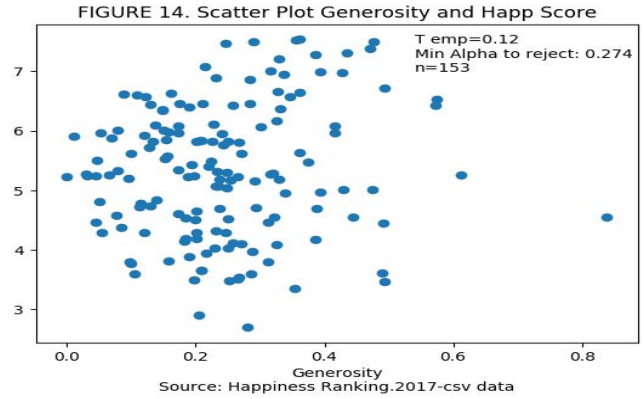
Figure 13. Scatter Plot Freedom and Happ Score



6. Generosity

After performing a hypothesis testing of two populations we obtained that we don't reject the null hypothesis with a confidence level of 95%, for rejecting the null hypothesis, the minimum alpha is 0.274. On the other hand, we found that the generosity and happiness scores have almost no correlation, but it still has a weak correlation as shown in Figure 14 with a score of 0.12.

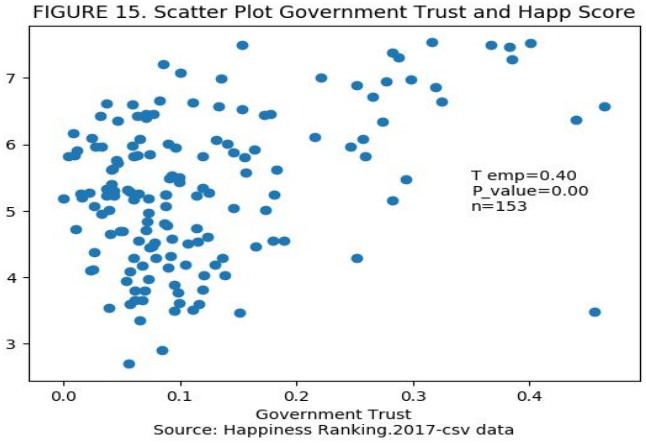
Figure 14. Scatter Plot Generosity and Happ Score



7. Government Trust

After performing a hypothesis testing of two populations we obtained that we reject the null hypothesis with a minimum alpha of approximately 0.00. On the other hand, we found that the government trust and happiness scores have a moderate correlation as shown in Figure 14 with a correlation score of 0.40.

Figure 15. Scatter Plot Government Trust and Happ Score



CONCLUSION

From the analysis of the dataset, several conclusions can be drawn regarding the importance and the effect of external factors on the state of happiness of inhabitants of a given country; what translates the link between our independent variables and the dependent variable. By combining hypothesis testing, ANOVA techniques and correlation, these statements can be affirmed with a 95% level of confidence :

- The level of happiness differs between continents
- The level of happiness is different depending on which government type people are faced
- The richer the country (GDP per capita), the happier its inhabitants are
- Higher the life expectancy, higher the happiness state
- Higher the government trust, higher the happiness state

Moreover the correlation study leads us to draw more insights by highlighting what is preponderent in making people happier; what could help leaders to change some aspects of their countries to turn the inhabitants happier and thus more productive:

Chart 7. Correlation ranking between IV and DV

Independent variable	Correlation with DV
GDP per capita	0.81
Life Expectancy	0.77
Degree of Freedom	0.56
Government trust	0.4
Generosity	0.12 (Makes sense with H0 non-rejection)

To conclude, we can rank the factors that make us happier with Chart 7. The higher the correlation, the higher the effect it has on our happiness. The two major factors affecting our life happiness are the money you earn and the longer you live.

REFERENCES

Democracy Index. (2019, December 12). Retrieved from: https://en.wikipedia.org/wiki/Democracy_Index

Rumsey, D. J. (n.d.). How to Interpret a Correlation Coefficient r. Retrieved from https://www.dummies.com/education/math/statistics/how-to-interpret-a-correlation-coefficient-r/.