What Makes You Happy?

Using Statistical Analysis To Measure Significance Of Happiness Scores

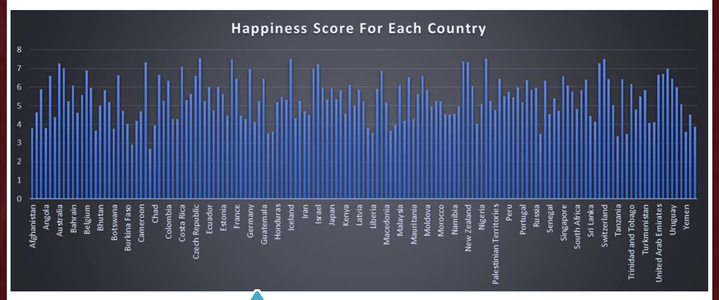
Emmett & Miriam McCoy College of Business Administration

Presented by: Brittney Lopanec, Katlyn Batterton, and Luis Maldonado.

QMST 4373C Data Analytics, McCoy College of Business, Texas State University

INTRODUCTION

Happiness is a state of mind that relates to life satisfaction and positive experience of emotions. Happiness is recognized as a measure for social progress. It depicts the effectiveness of a country. We have used clustering and regression as analytical methods for our analysis. Our project will give insights as to why each country holds their ranking in the World Happiness Report.



Top 5 Happiest Countries:

Norway 7.537

Denmark 7.522

Iceland 7.504

Switzerland 7.494

Finland 7.469

7.537 Norway

Range:

2.693 Central African Republic

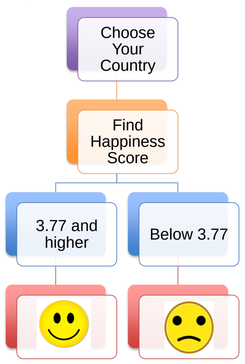
The USA is ranked at 14 with a score of 6.993

MOTIVATION

Keeping the happiness score high means the country is being effective and the people who live in that country are for the most part happy.

We were motivated to analyze this topic because of its relevance to our lives and many others. Being happy is something that everyone strives for, but what can we do to improve our happiness? Why aren't others happy? Can this dilemma be measured, bettered, and applied?

The procedure is simple…

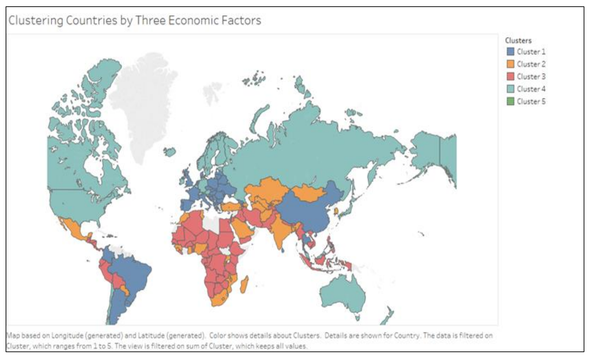


Do you consider yourself happy? What gives you that feeling of happiness?

DATA AND METHODOLOGY

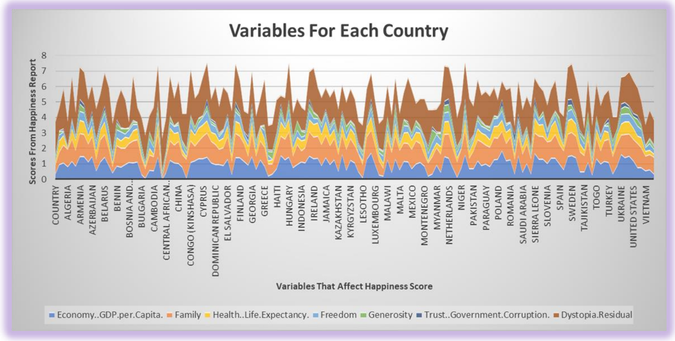
World Data Bank Variables Used For Clustering

1.GDP/capita: Gross Domestic Product per capita in ppp
2.Agriculture: percent of land that is sustainable for farming.
3.Population growth annual percent



World Happiness Report Variables

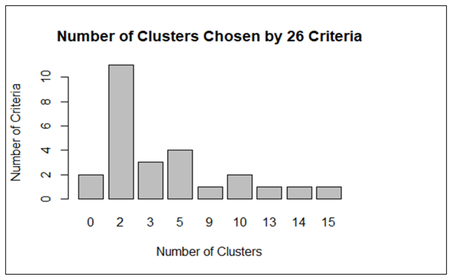
1. Economy (GDP per capita/production)
2. Family (Social Support)
3. Health (Life Expectancy)
4. Freedom
5. Generosity
6. Trust in Government (Absence of Corruption)
7. Dystopia: an imaginary country that has the world's least happy people. Dystopia Happiness Score(1.85) + the residual value or the unexplained value for each country.
8. Clusters-New variable we added using World Bank Data.



Process

We used "R" to determine the amount of clusters to use. After this, we accessed World Bank Data for Linear Regression and clustering to help us to find our Beta's significant scores. This led us to determine our variables for each cluster.

RESULTS



Determining Clusters

By using Nbclust package in "R" we were able to choose how many clusters to use for our regression based on 26 criteria.

Clustering vector:
[1] 4 5 2 2 5 5 3 3 4 3 4 5 3 2 2 2 2 5 2 5 5 5 2 4 2 2 3 2 2 3 5 5 5 2 2 2 5 3 5 5 5 2 2 2
[43] 3 2 3 5 2 5 3 4 5 2 4 4 2 3 5 3 4 2 2 2 3 3 5 5 3 2 4 2 1 4 5 4 4 2 5 1 5 4 4 3 2 3
[85] 2 5 4 5 4 3 4 4 2 2 2 5 3 2 2 4 3 2 2 2 4 4 2 2 5 5 1 5 3 4 4 2 5 4 1 5 3 4 4 5 5 2 3
[127] 3 2 2 2 5 4 3 4 4 4 4 5 3 5 3 5 4 2 2 2 2
Within cluster sum of squares by cluster:
[1] 12.06823 43.66892 35.37754 26.73255 31.15664
(between_SS / total_SS = 66.0 %)

Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
• The top four richest countries in the world. • All physically very small countries. • They are big Oil and Petroleum processors.	• Most all of them are countries in Africa. • Huge counts of famine loom over them. • There is a lot of conflict in relationships between these countries.	• Five of the Seven largest countries (land wise) are in this grouping. • Eight out of the ten most democratic countries in the world are in this cluster.	• Twelve out of twenty-five of these countries are the poorest in the world. • Thirteen of these countries are on the list for my most undeveloped countries in the world.	• Nine out of seventeen of these countries are considered the "most fun" in the world. • Seven out of fifteen of these countries provide the most culture packed experience.

Linear Regression Summary Results for Each Cluster

Coefficients:

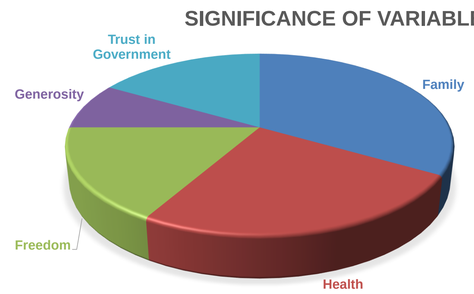
	Estimated	Standard Error	T value	Pr (> | t |)
Intercept	1.8642, 2.3791, 1.499, 2.3829, 1.0278	N/A, .3931, 1.4547, .4406, .9154	N/A, 6.052, 1.029, 5.409, 1.123	N/A, 3.07e-07, .3151, 1.3e-05, .270420
Family	2.4123, 0.8571, 2.7574, 1.5010, 1.9057	N/A, .3584, 1.1194, .3591, .4629	N/A, 2.392, 2.463, 4.180, 4.117	N/A, .0212, .0225, .000788, .000277
Health	0.4429, 3.1570, .1820, 1.8015, 2.0638	N/A, .4939, 1.0604, .4717, 1.1845	N/A, 6.392, .172, 3.819, 1.742	N/A, 9.85e-08, .8653, .000788, .091701
Freedom	1.8576, 0.8131, .2805, .8593, 1.5204	N/A, .6060, 1.5887, .8602, .7663	N/A, 1.342, .177, .999, 1.984	N/A, .1869, .8615, .327378, .056461
Generosity	N/A, -0.3893, 1.8404, -1.0411, -0.1421	N/A, .5593, .8462, .9289, .7461	N/A, -.696, 2.175, -1.121, -.190	N/A, .4901, .0412, .273019, .850236
Trust in Govt	N/A, -0.4465, 1.5318, -.3248, 2.2426	N/A, 2.0968, 1.1110, 1.2053, 1.1856	N/A, -.213, 1.379, -.269, 1.891	N/A, .8324, .1825, .789795, .068254

Color Key:
Cluster One
Cluster Two
Cluster Three
Cluster Four
Cluster Five

	Cluster One	Cluster Two	Cluster Three	Cluster Four	Cluster Five
Residual Std Error	NaN 0 DF	.5193 on 43 DF	.47 on 21 DF	.5354 on 25 DF	.54422 on 30 DF
R squared	1: ADJ NaN	.7246: ADJ .6925	.659: ADJ .5778	.7351: ADJ .6821	.6421 ADJ .5824
F-statistic	NaN on 3 & 0 DF	22.62 on 5 & 43 DF	8.116 on 5 & 21 DF	13.88 on 5 & 25 DF	10.76 on 5 & 30 DF
P-Value	N/A	4.715e-11	.0002151	1.545e-06	5.407e-06

CONCLUSION

After running linear regression for each of the five clusters, we were able to find which variables are significant towards their happiness score:



So yes, happiness is measurable, can be bettered, and can be applied to your country once one realizes the factors that matter most. Happiness is that attainable for any country and for any person!

RECOMMENDATIONS

We should all move to Norway!
Depending on what cluster you live in, you might find certain variables that lead to certain happiness scores being more significant than others. Our recommendation is for you to determine which variables make you happy and based on that, you should move to or visit that country. We can also analyze the significant variables in the happiest countries and work to improve those in the less happy countries.

REFERENCES

1. Risdal, Megan, et al. "World Happiness Report." *World Happiness Report | Kaggle*, CCO: Public Domain, 14 June 2017, www.kaggle.com/unsdsn/world-happiness.

2. Yong Kim, Jim. "World Bank Open Data." *World Bank Open Data*, 1 Jan. 2018, data.worldbank.org/.

ACKNOWLEDGEMENTS

This presentation is done as part of a project for QMST 4373C Data Analytics class, which is supervised by Dr. Tahir Ekin. We also acknowledge the support of Brandon Dee Roberts Professorship.

TEXAS STATE UNIVERSITY
The rising STAR of Texas