**Dissecting the World Happiness Report**

**Introduction**

Happiness as a concept is hard to define and has been the subject of much philosophical discussion. In principle, happiness is an intangible goal that all humans aspire to, but ironically there is no consensus about its definition or how to achieve it. Is happiness life satisfaction? Is it life quality? Is it a positive inner experience?

The World Happiness Report has been ranking countries through annual reports intending to quantify the happiness level of each country around the world. The methodology aggregates survey data from the *Gallup World Poll* (GWP), providing the self-reported happiness level of a sample of individuals from each country and linking this with indicators of economic production, social support, life expectancy, freedom, absence of corruption and generosity.

Due to the complex nature of happiness, we decided to investigate this methodology further, aiming to contextualise its meaning, and spot missing links or abnormalities which could highlight this subjectivity. Some areas for further exploration are as follows:

* Economics: There has been criticism that current classifications are too dependent on economic factors and biased towards our modern-day definition of well-being.
* Genetics: Science has been finding that happiness may be pre-determined by our genes at some level. Our study will not dive into this factor, but if genetics play a part, then a combination of external factors alone may not be enough to predict happiness.
* Culture & Language: Is happiness different from a Hindu or Christian perspective? Science has proven that our native tongue moulds our brain and the way we think. What does it mean in terms of happiness? Are some languages more predisposed to experience happiness?

**Methodology**

**Compiling Data**

We decided to merge all the countries classified in the World Happiness Report into a single dataset, with a wide range of factors that may relate to happiness. Some countries were excluded at the end (e.g. Palestine, Taiwan) for lack of consistent data and for having contested or partial sovereignty.

Data was organised using parameters similar to the report (Economy, Health, Trust in the Government, Social) but data relating to cultural aspects was also collected (Religion, Language, Region of the World, Cultural Values).

**Data Exploration**

Tableau was used for initial data exploration, to quickly capture which parameters looked directly correlated with happiness and to identify patterns that could be of later interest. For example, it was seen that Latin American countries were generally classified as happier when compared with Asian countries at similar economic levels.

* Show graphs xx

**Country Breakdown**

An initial exploration shows that certain factors, for example economic factors, must be present for base levels of happiness. However, after a certain point, increased prosperity does not always equate to increased happiness. There are studies indicating that wages are not as important for happiness levels after a certain bracket, so poorer countries may be more dependent on economic factors.

We decided to split the countries into three groups and separate the analysis: developed countries, developing countries and countries with severe humanitarian issues.

Kmeans was used to cluster these countries into three groups, where the GDP and child mortality were used as evaluation criteria.

(show picture xx)

**Machine learning analysis**

Once these three groups are defined, it is possible to see how the different parameters (Health, Economics, Social) will interact with the happiness score. Where there are clear differences between groups, this may provide clues as to the importance of each of the parameters.

Random Forest was useful to perceive the contribution of each happiness parameter. The results are as follows:

(show pictures xx)

**Conclusion**

This analysis is not a critique of the existence of Happiness ratings. Accurate or not, they are useful to raise debate about the human quest for happiness, with the intent of moving towards a better world.

The risk of the World Happiness Report may be to detract from intangible factors, specific to a country, that may generate overall happiness or satisfaction for its citizens.

The cultural elements highlighted in this analysis suggest that quantifying happiness with an integer value may always be an impossible task. Certainly, further data and analysis will be brought to this issue in the next years and in this scenario, one could suggest that quantifying happiness with a defined value may not be the most efficient evaluation. Should we change the definition itself for such ratings? Should we call it relative happiness? Or happiness potential?