

# **Project proposal Data Mining and Machine Learning**

## **Where you found the dataset?**

The World Happiness Report is an annual publication of the United Nations. The report is a measurement of life quality in the different countries all around the world.

This happiness measurement is based on different features that correlates with life factors. The different features of the dataset are Health information, the efficiency of social support, freedom level, democratic quality and others.

There are also information about the resource distribution between the residents of the countries. The dataset<sup>1</sup> can be found on the official website of the world happiness report<sup>2</sup>.

The dimensionality of the data set is 1704 and the cardinality is 26.

## **What data mining machine learning algorithm you will use?**

We are going to use different methods and tools that we learned during this class in order to extract and analyse the data.

- Linear and logistic regression : correlation between features
- Classification : gathering same characteristics among the countries
- Clustering : identifying populations with similar behaviours
- Comparison : comparison between different countries and continents
- Analyse : predictions about future results

## **Any other related background about the usefulness and impact of this approach?**

We will try to link our dataset with others to see if we can detect correlations with political and/or economic decisions from the different governments of the countries.

With the gathered informations we ll be able to predict with some accuracy whether the development slopes would increase, decrease or stay the same.

The objective is to predict a development slope but also to analyse how and why this development occurs.

## **Team Members :**

Haque Anika  
Boudry Alexandra  
Ahmed Shahzeb  
Bouzerda Samy

---

<sup>1</sup> <https://s3.amazonaws.com/happiness-report/2019/Chapter2OnlineData.xls>

<sup>2</sup> <https://worldhappiness.report/ed/2019/>