



World Happiness & Internet Usage

Ellen Leib, Kadi Sammul, Robert Leht

Introduction

Everyone wants to be happy. There are so many factors that have an effect on people's happiness, like their health, wealth and family. We will attempt to predict the results of this year's happiness scores and will also try to find the biggest contributor to a nation's people's happiness. With the number of daily Internet users growing each year, we also tried to see if there might be a correlation between happiness and the amount of Internet they consume.

Data

We are using two datasets from Kaggle. The first one is a happiness dataset, which contains the Happiness Score for 153 countries along with the factors used to explain the score. The Happiness Score is a national average of the responses to the main life evaluation question asked in the Gallup World Poll (GWP), which uses the Cantril Ladder. The second dataset is an internet usage dataset, which consists of the list of countries by the number of internet users in 2018. The data was collected from Wikipedia.

Methods

We worked as a team in Google Colab using Python. We cleaned and organised both of the datasets by changing column names, getting rid of NaN values and dropping features, where they were not useful. To analyse the data, we used a lot of data visualization implementing *matplotlib* bar, line and scatterplots, also *seaborn* heatmaps. To predict, we trained *RandomTree Regressors*, *DecisionTree Regressors* and *Support Vector Machines*. We then chose the method we thought seemed to make the most sense, so we used the *DecisionTree Regressor* to predict on 2020's data. We did not expect the machines to predict too well, as data was not extensive in either of the datasets.

Results

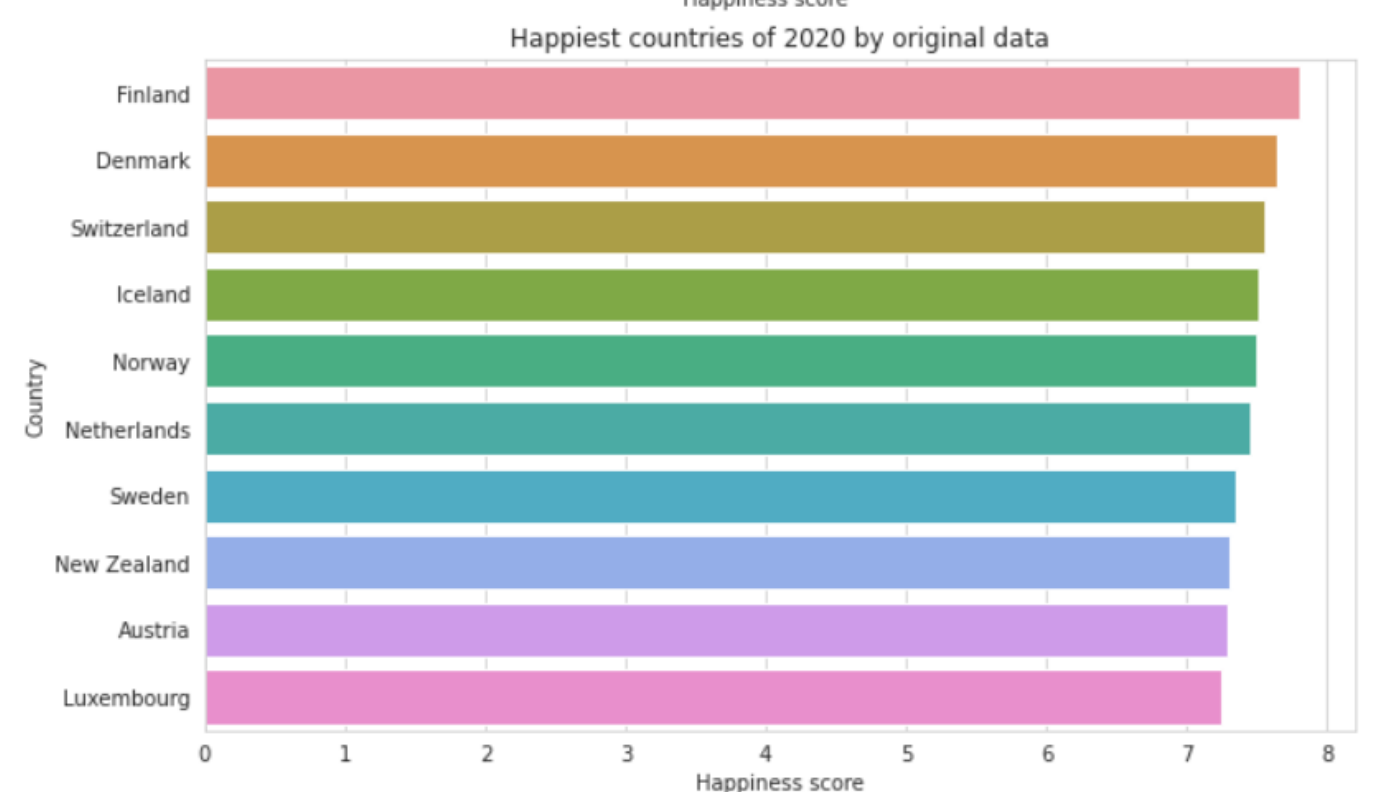
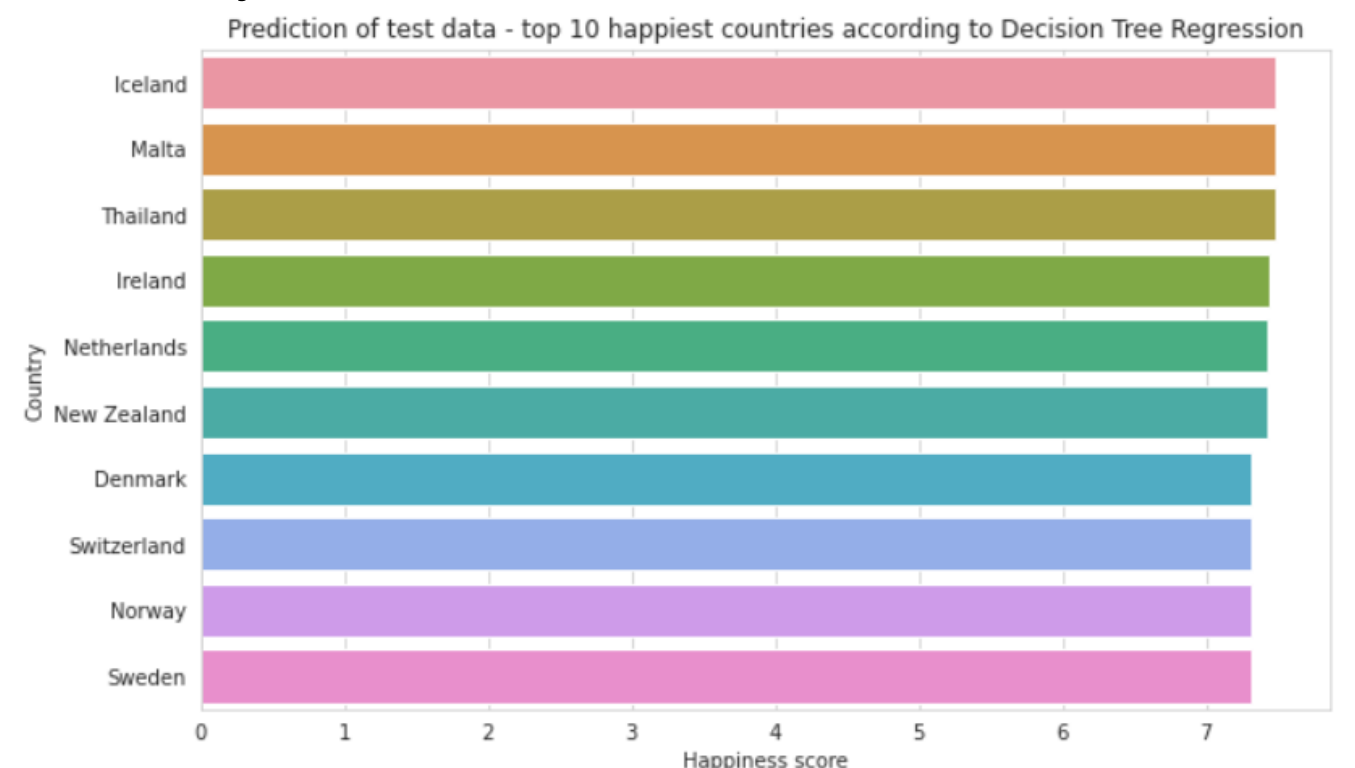
- Happiness score has shown a growing trend over the past 3 years



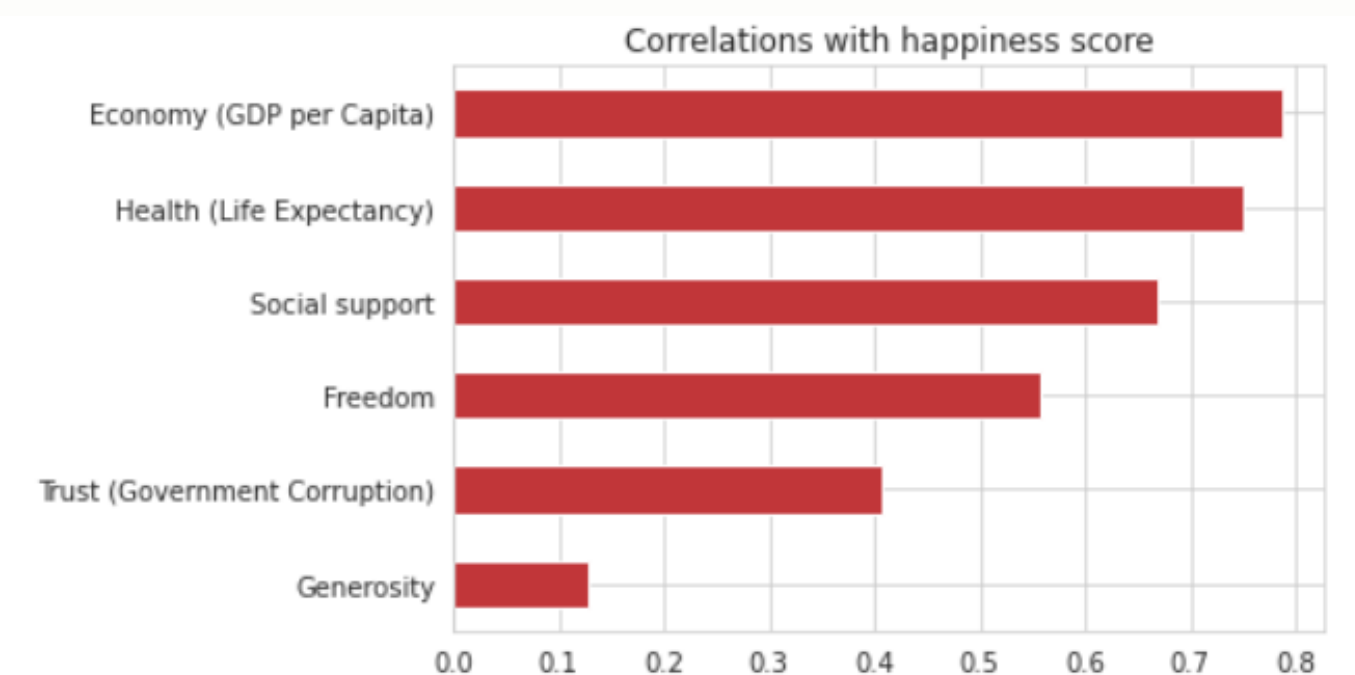
- From 2015 to 2020, Benin's happiness score increased the most, while Venezuela's decreased the most.

Results

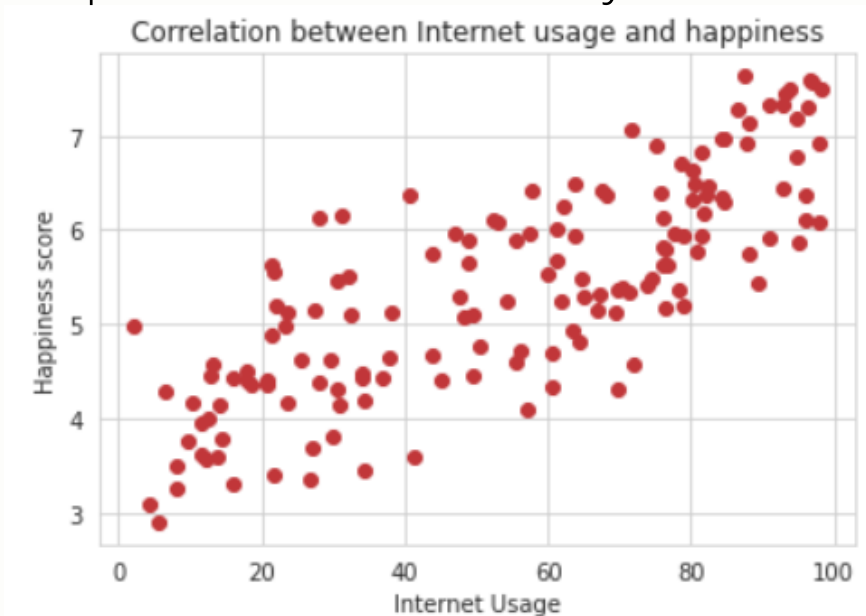
- Our machine did not rank any of the countries correctly, however 7 countries were in both tables.



- Happiness is most strongly correlated to a country's wealth, health and social support systems.



- Happiness is strongly correlated to Internet usage, however this is not conclusive as Internet usage is also strongly dependent on a country's economic state.



Necessary links

- <https://github.com/Sabasik/IDS-2020>
- <https://www.kaggle.com/mathurinache/world-happiness-report>
- <https://www.kaggle.com/tanuprabhu/list-of-countries-by-number-of-internet-users>