**World Happiness Dataset (2015)**

Number of observations: 142  
Number of variables: 23

Description of the Data:

The dataset used for the multivariate analysis of “World Happiness 2015” is mainly obtained from a world happiness report that was released by the United Nations on the International Day of Happiness. The report reviews the state of happiness in the world due to some common measurements of well-being and ranks them accordingly. The goal of this report was to assess not only the progress of nations, but also to inform governments, organizations and civil society about their policy-making decisions.

The so called “Happiness Score” basically assigns a ranking to 142 countries worldwide. The scores are based on answers to the main life evaluation of an individual performed by poll survey of Gallup World Poll. The features following the happiness score estimate the extent to which each factor contributes to a high life evaluation. These features are about

* Economy production (Economy GDP per Capita, ,
* Social support,
* Life expectancy,
* Freedom,
* Absence of corruption and
* Generosity.

To enlarge the dataset with variables that are related to happiness, the final dataset used was extended by some publicly available external features that are mentioned below.

* Quality of life (Clean Water, Alcohol Consumption, Air pollution, Legal Rights)
* Mental health (Suicide, Mental and Substance Disorder,
* Physical health (Child mortality, Infant Immunization Measles, HIV Disease, Obesity)
* Population (Birth Rate, Total Population, Population Growth,
* Infrastructure (Cellular Subscriber, Internet Use, Electricity Access)
* Environment (Forest Area Land Percentage, Agricultural Land Percentage, Urban Population Percentage)
* Education (Compulsory Education)
* Economy (Entrepreneurship Cost)
* Safety (Automotive Mortality)

With the described that available, we are answering the following two main questions:

* Which are the factors linked to happiness?
* Which factors lead to a low, mid or high happiness score?

The work includes the following multivariate techniques:

* Data preparation
* Dimension Reduction (PCA, FA)
* Clustering
* Regression

Add world map which shows the happiness (Tableau graph) and explain maybe the 3 happiest countries with highest scores and maybe the least happy one and why.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **#Nr** | **Variable Name** | **Type** | **Description** | **Source** |
| 1 | Country | String | Name of the country (baseline) | <https://www.kaggle.com/unsdsn/world-happiness> |
| 2 | Clean Water | Numeric | Share of the population with access to improved drinking water | <https://ourworldindata.org/water-access-resources-sanitation> |
| 3 | Suicide | Numeric | Suicide rates are number of deaths per 100,000 people; that is, the number of suicides in a country in a year, divided by the population and multiplied by 100,000. | <https://ourworldindata.org/suicide> |
| 4 | Mental and Substance Disorder | Numeric | People experienced an anxiety disorder | <https://ourworldindata.org/mental-health> |
| 5 | Birth Rate | Numeric | Children born per woman | <https://ourworldindata.org/fertility-rate> |
| 6 | Alcohol Consumption | Numeric | Average consumption of different beverage types per person | <https://ourworldindata.org/alcohol-consumption> |
| 7 | Child Mortality | Numeric | Population dying and surviving the first 5 years of life | <https://ourworldindata.org/child-mortality> |
| 8 | Total Population | Numeric | Total Population of the Country | <https://ourworldindata.org/world-population-growth> |
| 9 | Cellular Subscriber | Numeric | Mobile cellular subscriptions (per 100 people), 2015  Mobile phone subscriptions, measured as the number per 100 people. | <https://ourworldindata.org/internet> |
| 10 | Internet Use | Numeric | Shows the share of the population that is accessing the internet for all countries of the world.  Share of individuals using the internet in 2015, measured as the percentage of the population.  Internet users are individuals whohave used the Internet (from any location) in the last 3 months.  The Internet can be used via a computer, mobile phone,personal digital assistant, games machine, digital TV etc. | <https://ourworldindata.org/internet> |
| 11 | Air Pollution | Numeric | PM2.5 air pollution, mean annual exposure (micrograms per cubic meter) | <https://ourworldindata.org/air-pollution> |
| 12 | Region | String | Region the country belongs to | <https://www.kaggle.com/unsdsn/world-happiness> |
| 13 | Happiness Rank | Numeric | Rank of the country based on the Happiness Score | <https://www.kaggle.com/unsdsn/world-happiness> |
| 14 | Happiness Score | Numeric | A metric measured in 2015 by asking the sampled people the question: "How would you rate your happiness on a scale of 0 to 10 where 10 is the happiest." | <https://www.kaggle.com/unsdsn/world-happiness> |
| 15 | Economy (GDP per Capita) | Numeric | Quality of life based on the amount people earn | <https://www.kaggle.com/unsdsn/world-happiness> |
| 16 | Family | Numeric | Quality of family life, nuclear and joint family | <https://www.kaggle.com/unsdsn/world-happiness> |
| 17 | Health (Life Expectancy) | Numeric | Ranking healthcare availability and average life expectancy in the country | <https://www.kaggle.com/unsdsn/world-happiness> |
| 18 | Freedom | Numeric | How much an individual is able to conduct themselves based on their free will | <https://www.kaggle.com/unsdsn/world-happiness> |
| 19 | Trust (Government Corruption) | Numeric | Trust in government to not be corrupt | <https://www.kaggle.com/unsdsn/world-happiness> |
| 20 | Generosity | Numeric | How much their country is involved in peacekeeping and global aid | <https://www.kaggle.com/unsdsn/world-happiness> |
| 21 | Forest Area Land Percentage | Numeric | Percentage of forest in a land |  |
| 22 | Electricity Access | Numeric |  |  |
| 23 | Infant Immunization Measles | Numeric | Percentage of children ages 12-23 months that have been vaccinated to this disease |  |
| 24 | Urban Population |  | Percentage of population that live in an urban or rural environment |  |
| 25 | Obesity |  | How many people suffer from obesity |  |
| 26 | Compulsory Education |  |  |  |
| 27 | Agriculture Land |  |  |  |
| 28 | Automotive Mortality |  |  |  |
| 29 | Population Growth |  |  |  |
| 30 | Entrepreneurship Cost |  | How much it is to create a business |  |
| 31 | HIV Disease |  |  |  |