

## Dataset

Download the dataset (happiness\_development\_2017.csv) from the course website: [http://vda.univie.ac.at/Teaching/Vis/19w/data/happiness\\_development.csv](http://vda.univie.ac.at/Teaching/Vis/19w/data/happiness_development.csv)

It is composed from multiple sources:

(Happy) Happiness dataset: <https://www.kaggle.com/unsdsn/world-happiness/version/2#2017.csv>

(CIA) CIA Factbook: <https://www.cia.gov/library/publications/the-world-factbook/index.html>

(INA) Inequality-adjusted Human Development Index: <http://hdr.undp.org/en/content/inequality-adjusted-human-development-index-ihdi>

(HDI) Human Development Index: <http://hdr.undp.org/en/content/human-development-index-hdi>

Fields:

**Country (Key):** Name of the country.

**Happiness\_rank (Happy):** Rank of the country based on the score.

**Happiness\_score (Happy):** A metric measured in 2017 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."

**Economy (Happy):** The extent to which GDP contributes to the calculation of the Happiness Score.

**Family (Happy):** The extent to which Family contributes to the calculation of the Happiness Score.

**Health (Happy):** The extent to which Life expectancy contributed to the calculation of the Happiness Score.

**Freedom (Happy):** The extent to which Freedom contributed to the calculation of the Happiness Score.

**Generosity (Happy):** The extent to which Generosity contributed to the calculation of the Happiness Score.

**Corruption (Happy):** The extent to which Perception of Corruption contributes to Happiness Score.

**Dystopia\_residual (Happy):** The extent to which Dystopia Residual contributed to the calculation of the Happiness Score.

**Internet\_access\_population[%] (CIA):** Percentage of the population with internet access.

**Cellular\_subscriptions (CIA):** Number of cellular subscriptions per 100 inhabitants.

**Surplus\_deficit\_GDP[%] (CIA):** This entry records the difference between national government revenues and expenditures, expressed as a percent of GDP

**Family\_income\_gini\_coeff (CIA):** This index measures the degree of inequality in the distribution of family income in a country. [0, 100] where 0 means perfect equality and 100 means perfect inequality.

**GDP\_per\_capita[\$] (CIA):** This entry shows GDP on a purchasing power parity basis divided by population as of 1 July for the same year.

**Inflation\_rate[%] (CIA):** This entry furnishes the annual percent change in consumer prices compared with the previous year's consumer prices.

**Military\_expenditures[%] (CIA):** The spending's for defense programs in % of the GDP.

**Map\_reference (CIA):** The mapping region within the factbook. References the continent.

**Biggest\_official\_language (CIA):** Biggest officially accepted language. Multiple official languages are omitted for the sake of simplicity.

**Population (CIA):** The number of inhabitants.

**Human Development Index (HDI):** A composite index measuring average achievement in three basic dimensions of human development—a long and healthy life, knowledge and a decent standard of living. See *Technical note 1* at [http://hdr.undp.org/sites/default/files/hdr2018\\_technical\\_notes.pdf](http://hdr.undp.org/sites/default/files/hdr2018_technical_notes.pdf) for details on how the HDI is calculated.

**Life expectancy at birth [Years] (HDI):** Number of years a newborn infant could expect to live if prevailing patterns of age-specific mortality rates at the time of birth stay the same throughout the infant's life.

**Expected years of schooling [Years] (HDI):** Number of years of schooling that a child of school entrance age can expect to receive if prevailing patterns of age-specific enrolment rates persist throughout the child's life.

**Mean years of schooling [Years] (HDI):** Average number of years of education received by people ages 25 and older, converted from education attainment levels using official durations of each level.

**Gross national income (GNI) per capita [\$] (HDI):** Aggregate income of an economy generated by its production and its ownership of factors of production, less the incomes paid for the use of factors of production owned by the rest of the world, converted to international dollars using PPP rates, divided by midyear population.

**HDI rank for 2016 (HDI):** Ranking by HDI value for 2016, which was calculated using the same most recently revised data available in 2018 that were used to calculate HDI values for 2017.

**Inequality-adjusted HDI (INA):** HDI value adjusted for inequalities in the three basic dimensions of human development. See *Technical note 2* at [http://hdr.undp.org/sites/default/files/hdr2018\\_technical\\_notes.pdf](http://hdr.undp.org/sites/default/files/hdr2018_technical_notes.pdf) for details on how the IHDI is calculated.

**Inequality in life expectancy [%], 2015–2020 (INA):** Inequality in distribution of expected length of life based on data from life tables estimated using the Atkinson inequality index.

**Inequality in education [%] (INA):** Inequality in distribution of years of schooling based on data from household surveys estimated using the Atkinson inequality index.

**Inequality in income [%] (INA):** Inequality in income distribution based on data from household surveys estimated using the Atkinson inequality index.

**Quintile ratio, 2010-2017 (INA):** Ratio of the average income of the richest 20 percent of the population to the average income of the poorest 20 percent of the population.

**Palma ratio, 2010-2017 (INA):** Ratio of the richest 10 percent of the population's share of gross national income (GNI) divided by the poorest 40 percent's share. It is based on the work of Palma (2011).

**Gini coefficient, 2010-2017 (INA):** Measure of the deviation of the distribution of income among individuals or households within a country from a perfectly equal distribution. A value of 0 represents absolute equality, a value of 100 absolute inequality.