

# Project 1:

# How happy is the world?



Team Members:  
Julia Begley  
Hasan Gokceoglu  
Isil Bulut  
Violetta Antonova  
Sumeyra Bharuchi

# Executive summary

- ▷ The World Happiness Report measures happiness levels around the world.
- ▷ Provides a valuable resource for researchers, policymakers, and anyone interested in :
  - Understanding global happiness trends.
  - The factors contributing to well-being.

- Country name
- Year
- Life Ladder
- Log GDP per capita
- Social support
- Healthy life expectancy at birth
- Freedom to make life choices
- Generosity
- Perceptions of corruption
- Positive affect
- Negative affect
- Confidence in national government

# Data Collection

- In this study we analyzed the World Happiness Report data between 2005-2021 to get insights about the factors affecting the happiness of countries around the world.

## World Happiness Report 2005-2021

Exploring global happiness trends and factors influencing well-being



```
# Study data files
happiness_csv = "World _Happiness_Report_2005-2021.csv"

# Read the mouse data and the study results
happiness_metadata = pd.read_csv(happiness_csv)

happiness_metadata
```

# Data

- ▶ The Report includes happiness scores (represented by 'Life Ladder' in the data) for countries, as well as several other socio-economic factors that are believed to impact happiness.

	Country name	Year	Life Ladder	Log GDP per capita	Social support	Healthy life expectancy at birth	Freedom to make life choices	Generosity	Perceptions of corruption	Positive affect	Negative affect	Confidence in national government
0	Afghanistan	2008	3.723590	7.302574	0.450662	50.500000	0.718114	0.173169	0.881686	0.414297	0.258195	0.612072
1	Afghanistan	2009	4.401778	7.472446	0.552308	50.799999	0.678896	0.195469	0.850035	0.481421	0.237092	0.611545
2	Afghanistan	2010	4.758381	7.579183	0.539075	51.099998	0.600127	0.125859	0.706766	0.516907	0.275324	0.299357
3	Afghanistan	2011	3.831719	7.552006	0.521104	51.400002	0.495901	0.167723	0.731109	0.479835	0.267175	0.307386
4	Afghanistan	2012	3.782938	7.637953	0.520637	51.700001	0.530935	0.241247	0.775620	0.613513	0.267919	0.435440
5	Afghanistan	2013	3.572100	7.657503	0.483552	52.000000	0.577955	0.066344	0.823204	0.547417	0.273328	0.482847
6	Afghanistan	2014	3.130896	7.650827	0.525568	52.299999	0.508514	0.109215	0.871242	0.491641	0.374861	0.409048
7	Afghanistan	2015	3.982855	7.634466	0.528597	52.599998	0.388928	0.085082	0.880638	0.491410	0.339276	0.260557
8	Afghanistan	2016	4.220169	7.629037	0.559072	52.924999	0.522566	0.047488	0.793246	0.501409	0.348332	0.324990
9	Afghanistan	2017	2.661718	7.629684	0.490880	53.250	1	-0.116068	0.954393	0.435270	0.371326	0.261179

# Data Clean-up

- For analysing the data, we used Python and several other data analysis libraries such as Pandas, Scipy
- For Visualisation studies, we used Matplotlib, Plotly, Cufflinks, Seaborn and APIs such as Geoapify
- For cleaning the data, we identified the empty cells for each dataset and removed them from the main dataset.

	Country name	Year
0	Afghanistan	2008
1	Afghanistan	2009
2	Afghanistan	2010
3	Afghanistan	2011
4	Afghanistan	2012
...	...	...
2084	Zimbabwe	2017
2085	Zimbabwe	2018
2086	Zimbabwe	2019
2087	Zimbabwe	2020
2088	Zimbabwe	2021
2089 rows × 12 columns		



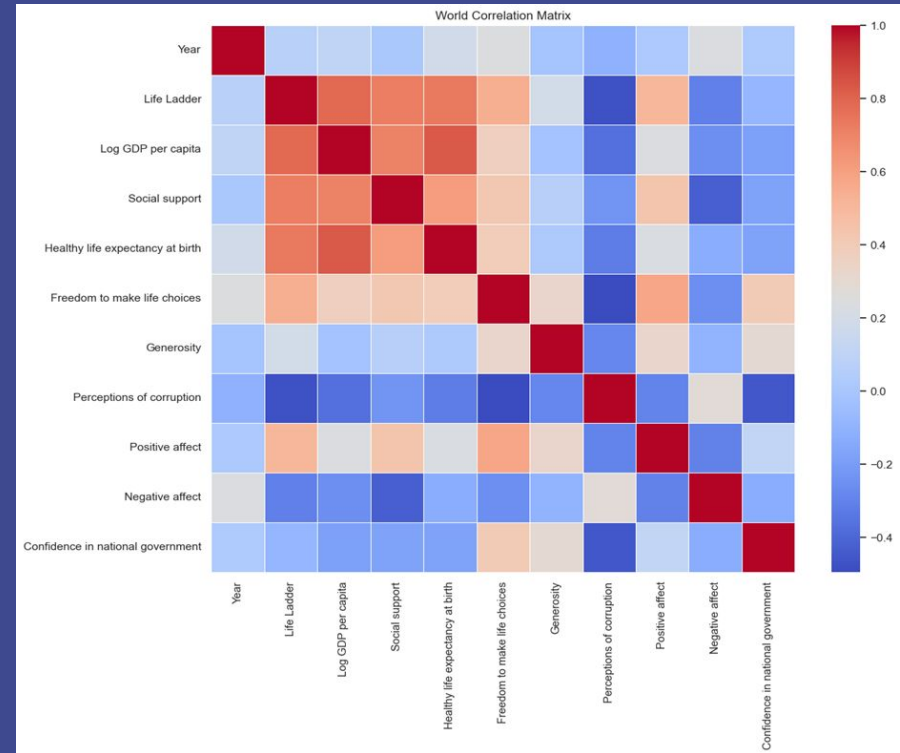
	Country name	Year
0	Afghanistan	2008
1	Afghanistan	2009
2	Afghanistan	2010
3	Afghanistan	2011
4	Afghanistan	2012
...	...	...
2084	Zimbabwe	2017
2085	Zimbabwe	2018
2086	Zimbabwe	2019
2087	Zimbabwe	2020
2088	Zimbabwe	2021
2031 rows × 12 columns		

# Our approach

- ▷ We used World Happiness Report data to perform correlation and trend analysis and try to gain insights into the factors affecting happiness across different countries.
- ▷ We wanted to look into:
  - Is there a positive relationship between social support and generosity?
  - Is there a positive relationship between happiness scores and their country's healthy life expectancy?
  - Is there a clear relationship between confidence in the government, freedom to make life choices, and perceived corruption?
  - How do these metrics differ globally or regionally?

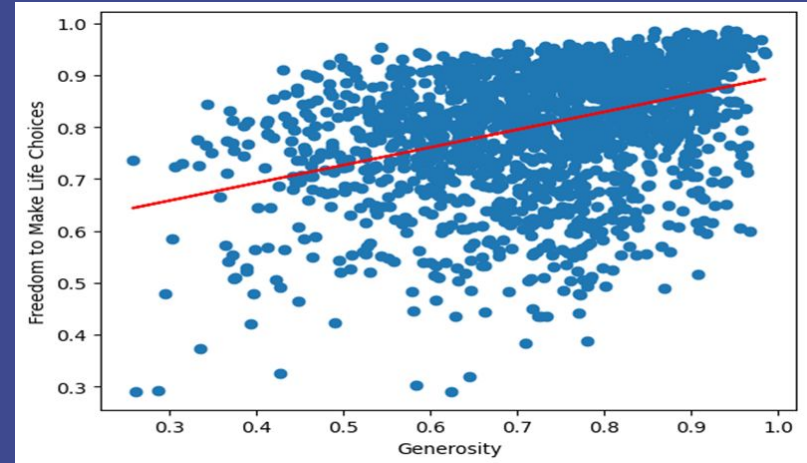
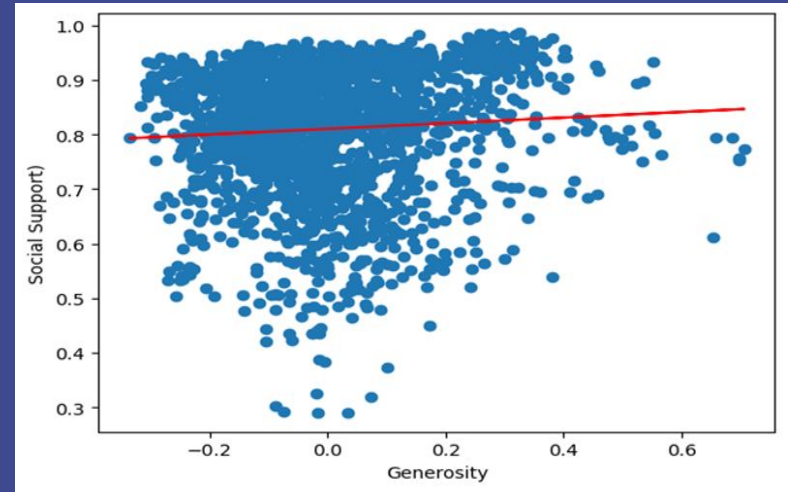
# World Happiness Correlation Matrix

- This heatmap shows the correlations between all variables.
- The darkest reds show the strongest positive correlations and the darkest blues show the strongest negative correlations.
- The diagonal line of matrix shows the "1" correlation where each variable is compared to itself.



# What is the relationship between Generosity and Social Support?

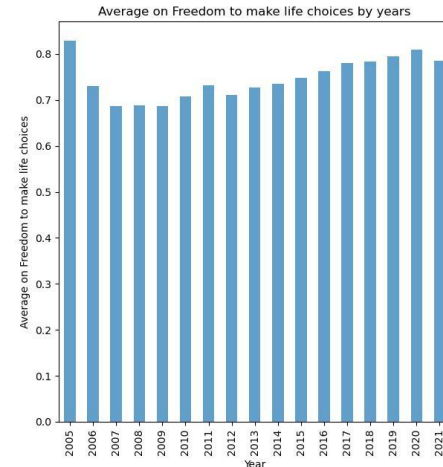
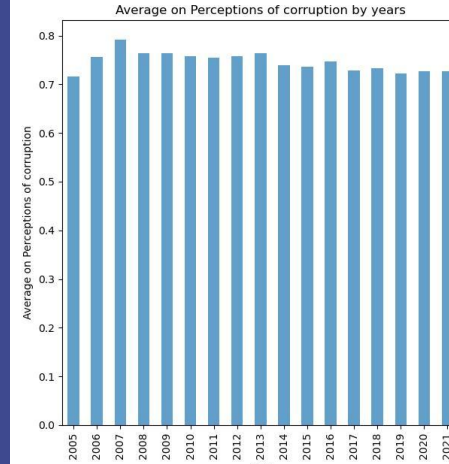
- ▷ The positive relationship between generosity and social support is none or very weak (0,07).
- ▷ The relationship between generosity and freedom to make life choices is weak but much higher than the one with social support (0,40).





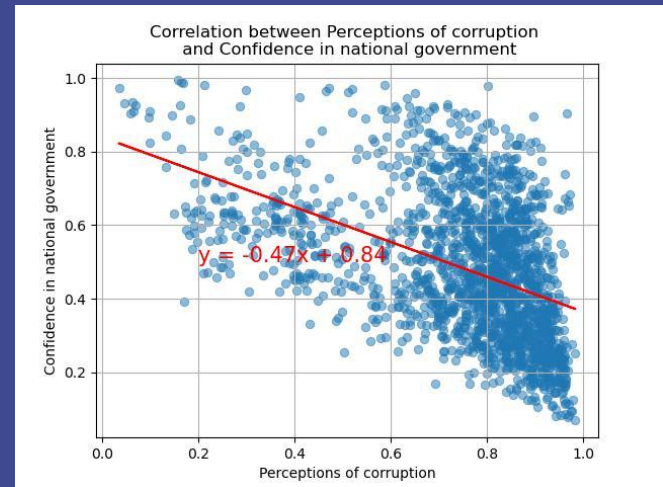
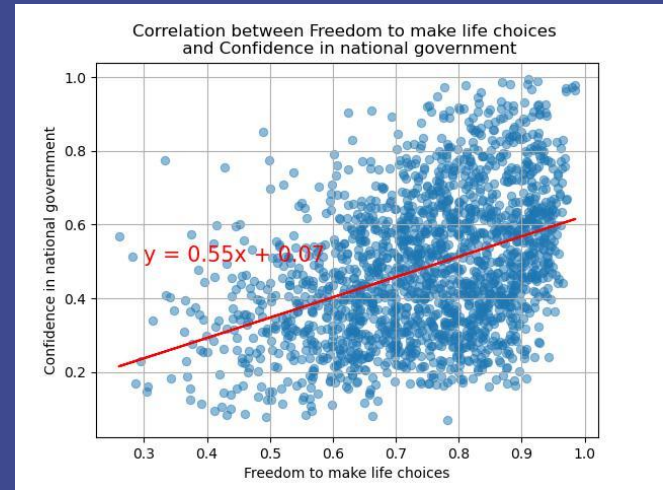
# What are the relationships between Freedom to make life choices and Perceptions of corruption, with Confidence in national government?

- Corruption levels from year 2007 to year 2021 went slightly down.
- Freedom to make life choices from year 2007 to year 2020 went up, but in year 2021 the value dropped down.



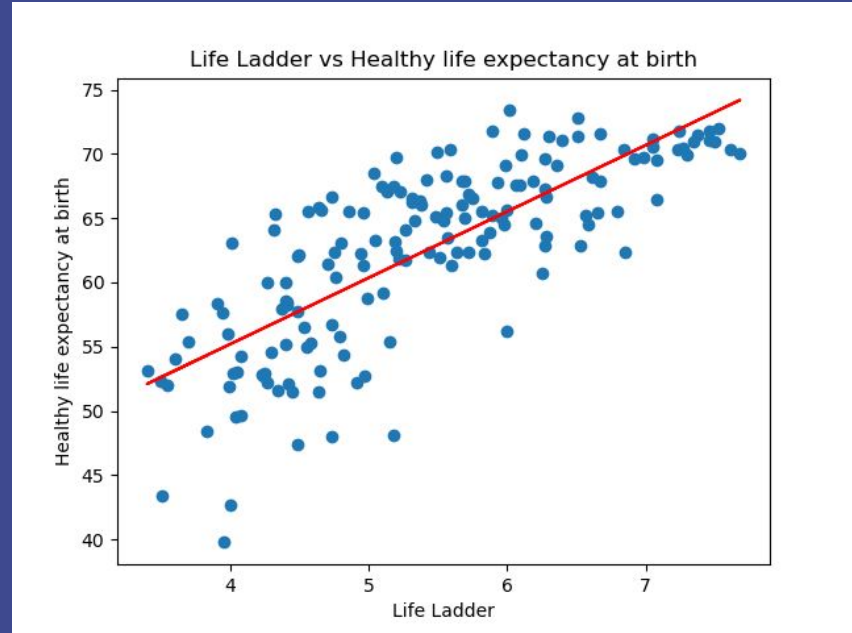
# What is the relationship between Freedom to make life choices and Perceptions of corruption, with Confidence in national government?

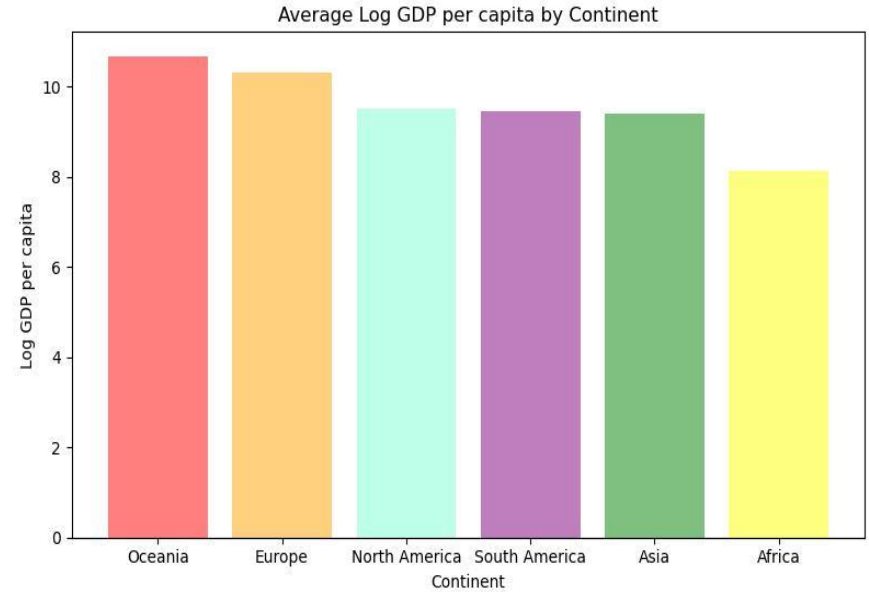
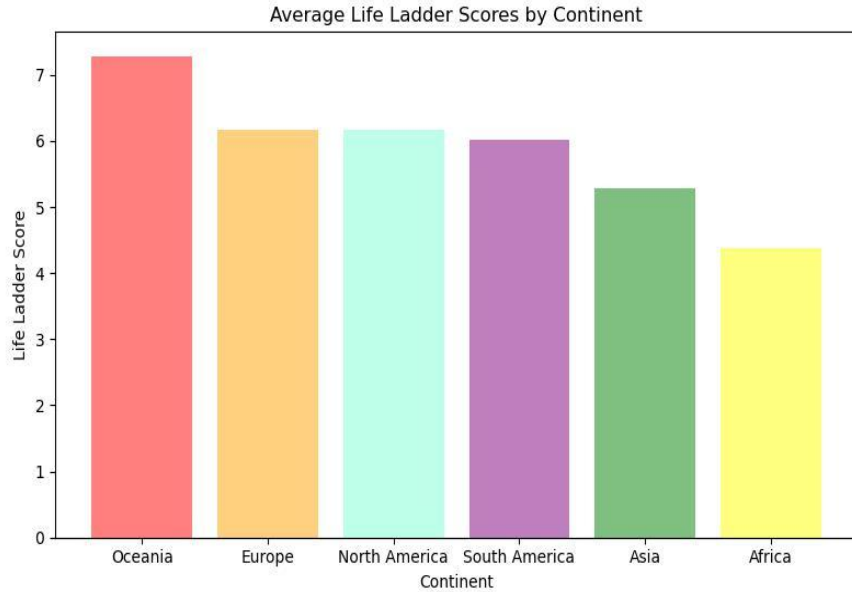
- The correlation between Freedom to make life choices and Confidence in national government is positive and the correlation coefficient is 0.4.
- The correlation between Perceptions of corruption and Confidence in national government is negative and the correlation coefficient is -0.46.



Is there a positive relationship between happiness scores and their country's healthy life expectancy?

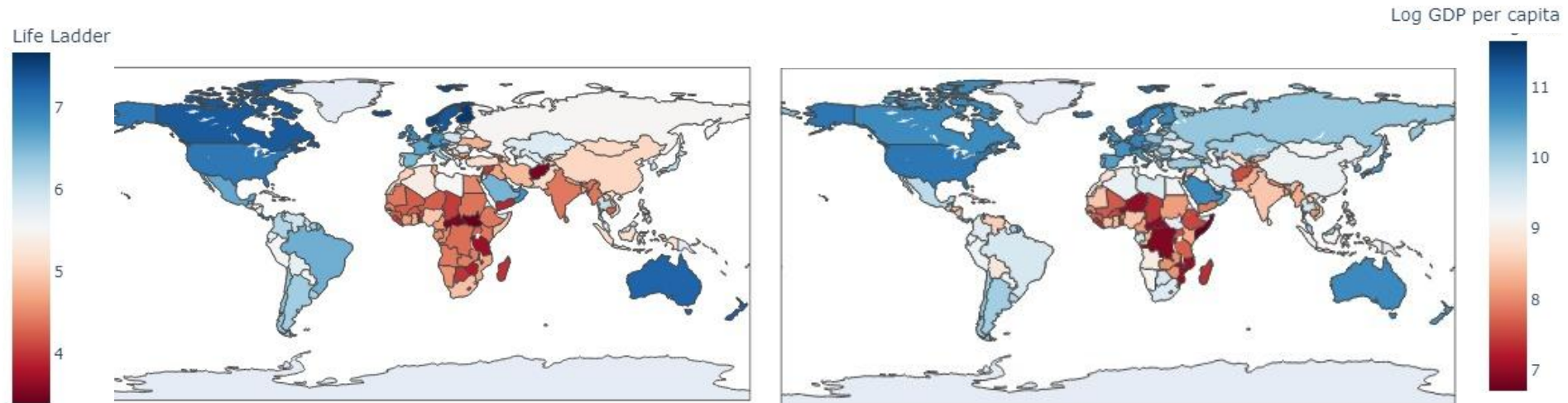
- ▷ The correlation coefficient value between Life Ladder and Healthy life expectancy at birth is 0.77.
- ▷ This means the association between Average Life Ladder and Healthy life expectancy at birth is very high.





- The continents show in the same order for both measures with Oceania as highest for both.
- North America, South America and Asia have similar Log GDP per capita but Asia has a disproportionately low average Life Ladder score.

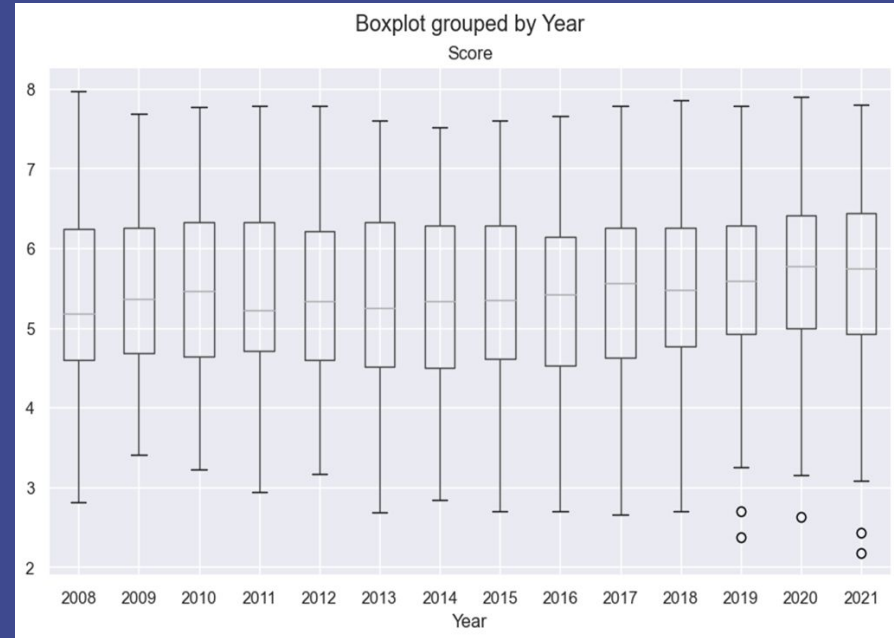
## Average Life Ladder Score by Country



- We expected Log GDP per capita to show a clearer relationship with average Life Ladder scores.
- There are a lot more differences between the 2 maps than we expected.

# Box plots Grouped by Year

- ▶ 50% of Happiness score changed between 4 - 6 out of 8 over the years.
- ▶ There were some negative outliers during Covid years (2019-2021).
- ▶ There were lower happiness levels in 2008 and 2011.
- ▶ We ran the Anova test and our  $H_0$  hypothesis was rejected, so the happiness score change between 2008-2021 is significant.



# Results & conclusions

- ▷ Happiness score changed significantly over time, it is interesting that the world was happier during Covid years compared to years 2008-2011.
- ▷ The factors we would expect to increase happiness (like a higher healthy life expectancy at birth, social support or freedom to make life choices), did really increase happiness.
- ▷ The happiest countries - Norway, Denmark and Finland - are also among the top 10 with the least reported corruption.
- ▷ It was surprising that the influence of Log GDP per capita on happiness was more complex than expected.

## Next steps:

1. Try to predict what will happen in the next years, based on collected data.
2. Decide what is the best way to clean the data, to have minimal impact from losing data.
3. Analyse the data in new groups and categories, such as East Europe, Central Asia, etc.
4. Explore any changes within particular countries year-by-year.



# March 20th is the International Day of Happiness

