

Bi-module project: What makes people happy 😊?

Ruiz, Ivana

Data Analytics Bi-module project - allWomen

May 31st 2025

Table of content:

- 1. Some context
- 2. Data cleaning & EDA
- 3. Hypothesis testing
- 4. Data viz
- 5. Conclusions



Context

1. Identify the key contributors to happiness and offer insights for individuals and governments to foster well-being.

Analytical Approach

2. **Data Cleaning & EDA** – Address inconsistencies.
3. **Hypothesis testing.**
4. **Data analysis / Univariate & Bivariate Analysis** – Explore relationships between happiness and key factors.
5. **Conclusions**

The three major indicators to measure research impact are:

1. Does higher **GDP** increase happiness?
2. Does **Generosity** positively affect happiness?
3. Does **freedom** contribute to happiness?

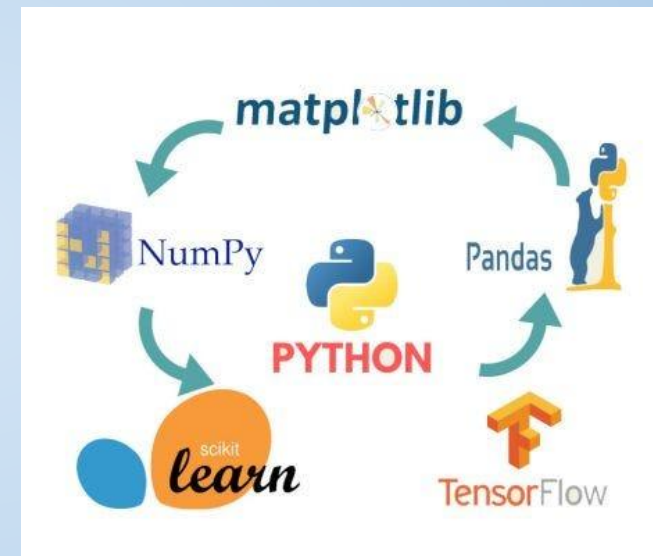
Dataset information

From Kaggle:

`Dataset-Happiness_Score.csv`



"If you don't reveal some insights soon, I'm going to be forced to slice, dice, and drill!"



Part 2: Data
Cleaning and EDA
(Exploratory Data
Analysis)

Part 2: Data cleaning and EDA

- column “Healthy life expectancy” we have more than 50% NaN (decided to dropped it)
- The data frame has 312 rows (2018 and 2019).
- 11 numerical variables and one categorical (“Country”)
- No duplicates.
- Rename columns.
- Import EDA function/ Replace NaN with the median in Corruption variable.

EDA

- Happiness Score: Range of Happiness suggests significant variation across different regions.
- GDP: Range 0.0 to 1.7 Indicates economic disparity among surveyed countries.
- Freedom: Mean ~0.4 (Scale 0-1)
- Corruption: Mean ~0.1.
- Social support: Scale 0-2 * Most countries seem to have moderate levels of social support.

	count	mean	std	min	25%	50%	75%	max
Unnamed: 0.1	312.000000	155.500000	90.210864	0.000000	77.750000	155.500000	233.250000	311.000000
Unnamed: 0	312.000000	77.500000	45.104737	0.000000	38.750000	77.500000	116.250000	155.000000
Overall rank	312.000000	78.500000	45.104737	1.000000	39.750000	78.500000	117.250000	156.000000
Score	312.000000	5.391506	1.114631	2.853000	4.514250	5.379500	6.176000	7.769000
GDP per capita	312.000000	0.898298	0.394592	0.000000	0.609500	0.960000	1.219500	2.096000
Social support	312.000000	1.211026	0.300310	0.000000	1.055750	1.265500	1.457500	1.644000
Healthy life expectancy	7.000000	0.038571	0.035213	0.000000	0.005000	0.048000	0.066000	0.080000
Freedom to make life choices	312.000000	0.423538	0.156024	0.000000	0.322500	0.449500	0.540250	0.724000
Generosity	312.000000	0.182926	0.096739	0.000000	0.108750	0.175500	0.245000	0.598000
Perceptions of corruption	311.000000	0.111299	0.095365	0.000000	0.050000	0.082000	0.140500	0.457000
year	312.000000	2018.500000	0.500803	2018.000000	2018.000000	2018.500000	2019.000000	2019.000000

Hypothesis testing:

* **Null Hypothesis (H_0):** There is no significant difference in the mean Happiness Score between countries with high and low Freedom to make life choices.

* **Alternative Hypothesis (H_1):** There is a significant difference in the mean Happiness Score between countries with high and low Freedom to make life choices.

Hypothesis testing:

Test if the mean Happiness Score is significantly different between countries with high and low Freedom to make life choices

1. Created two groups: high_freedom and low_freedom
2. alpha (α): 0.5
3. One-sided t-test to compare
4. Define the threshold using the Median
5. Z-score: 1.00, P-value: 0.1587
6. The mean Happiness Score is significantly different between countries with high and low Freedom to make life choices. So, we reject the null hypothesis



Univariate and Bivariate analysis

IQR Method to visualize the outliers

- * Box plots for visualize extreme values
- * The interquartile Range (IQR) method identifies outliers as values below $Q1 - 1.5 \times IQR$ or above $Q3 + 1.5 \times IQR$

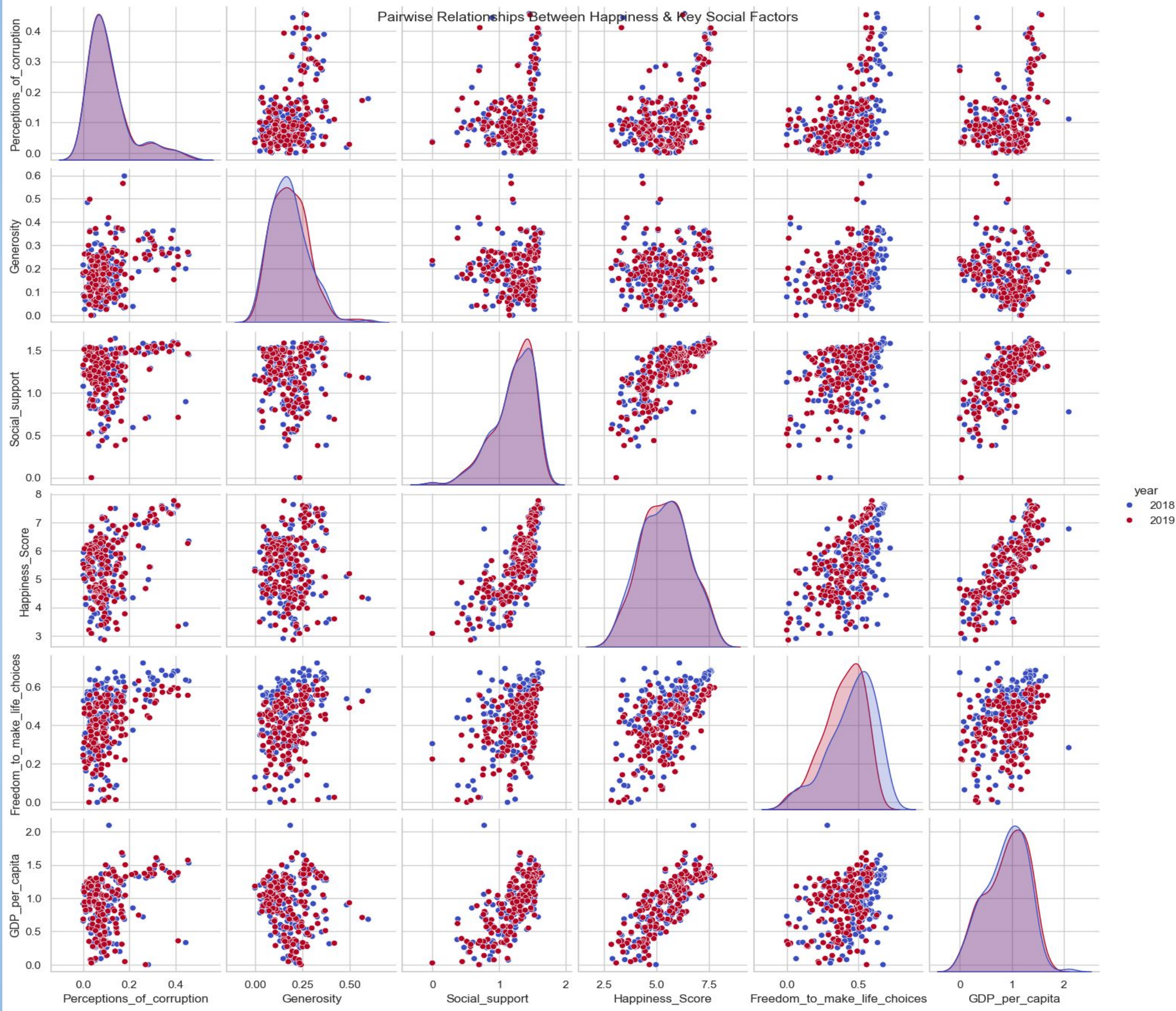
Outliers Across Happiness Dataset Variables



Univariate and Bivariate analysis

- Top outliers:
- Perceptions of corruption (Dropped)
 - Freedom to make life choices and,
 - Generosity.

** Social support has the highest correlation with the score apart from GDP.



Should I follow the **DATA**
or my instincts?

DATA



Final dataset variables:

- ☐ Overall_Rank
- ☐ Country
- ☐ Happiness_Score
- ☐ GDP_per_Capita
- ☐ Social_support
- ☐ Freedom_to_make_life_choices
- ☐ Generosity
- ☐ year

Happiness_Dataset

4. Data viz and storytelling

What makes people happy?

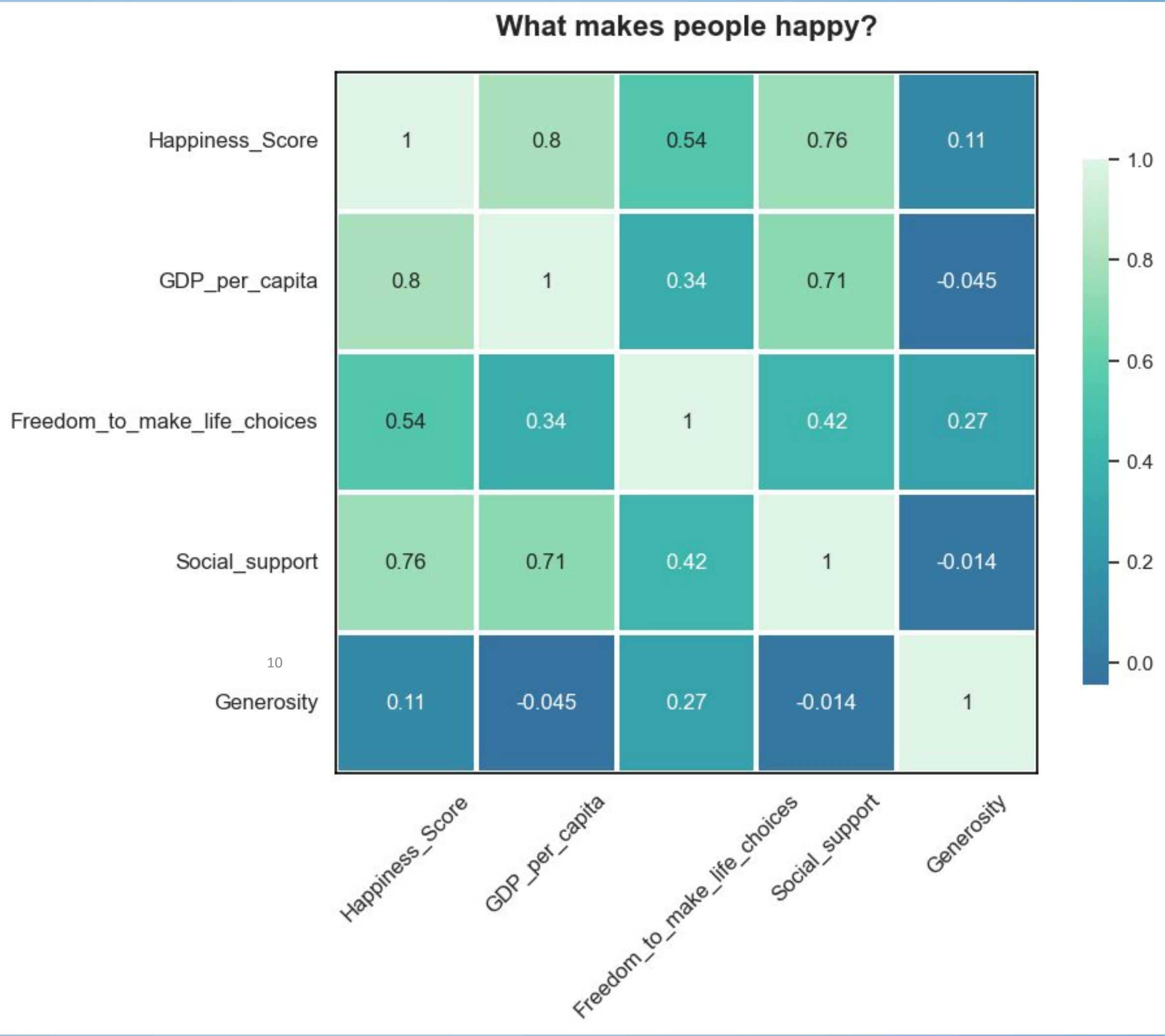
We are going to use Happiness_Score with other variables to measure this.

- * 1. Is the GDP per capita which makes you happy?
- * 2. Why the most generous countries are the ones with less happiness score?
- * 3. Is the Freedom of Life Choices which makes you happy?



Correlation matrix

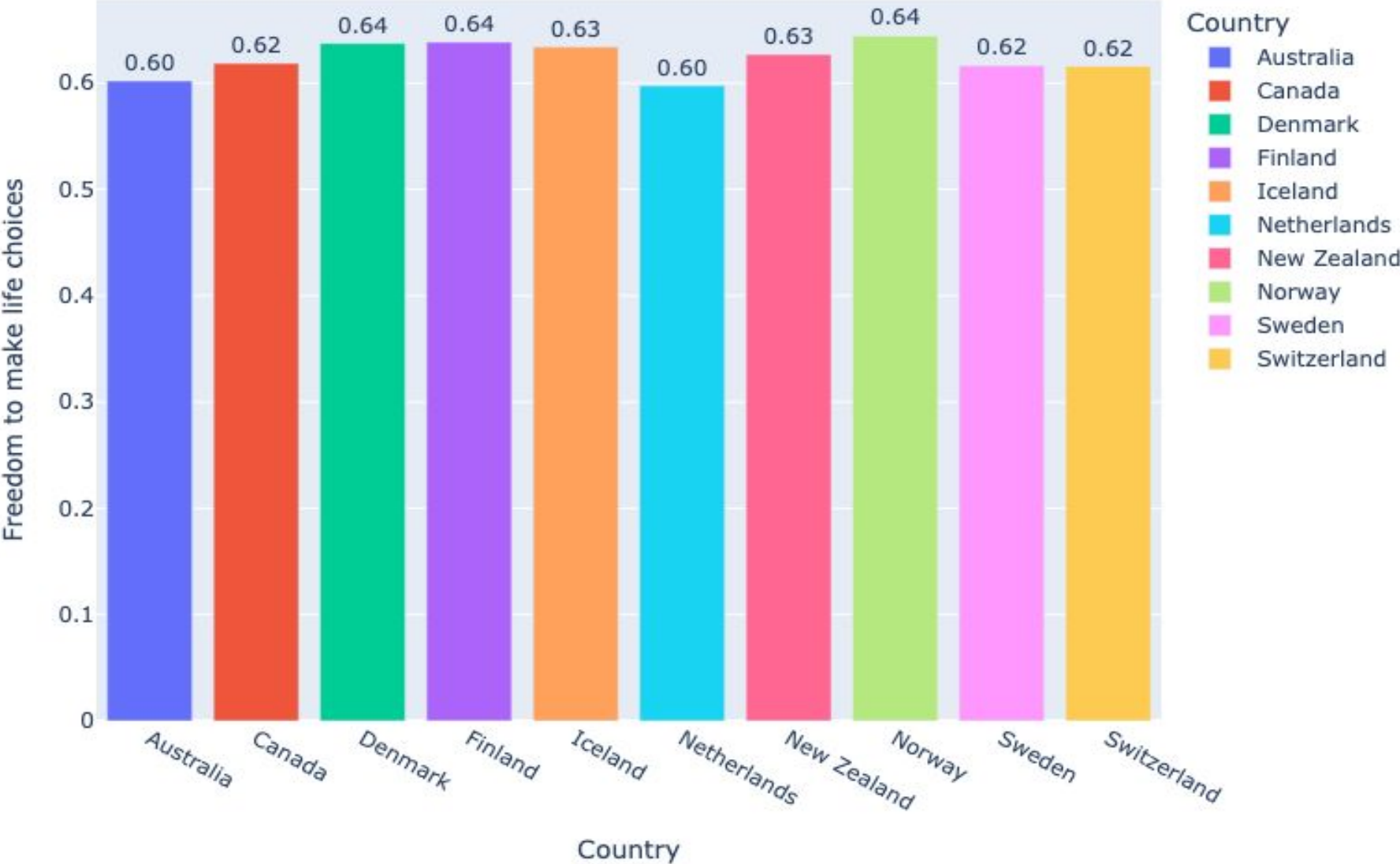
- * Money helps, but it's not everything—GDP has a strong influence, but social and psychological factors play key roles too.
- * Freedom contributes to happiness, but it's not always tied to GDP.
- * The strongest happiness predictor here is personal freedom, showing that feeling in control of one's life matters significantly.



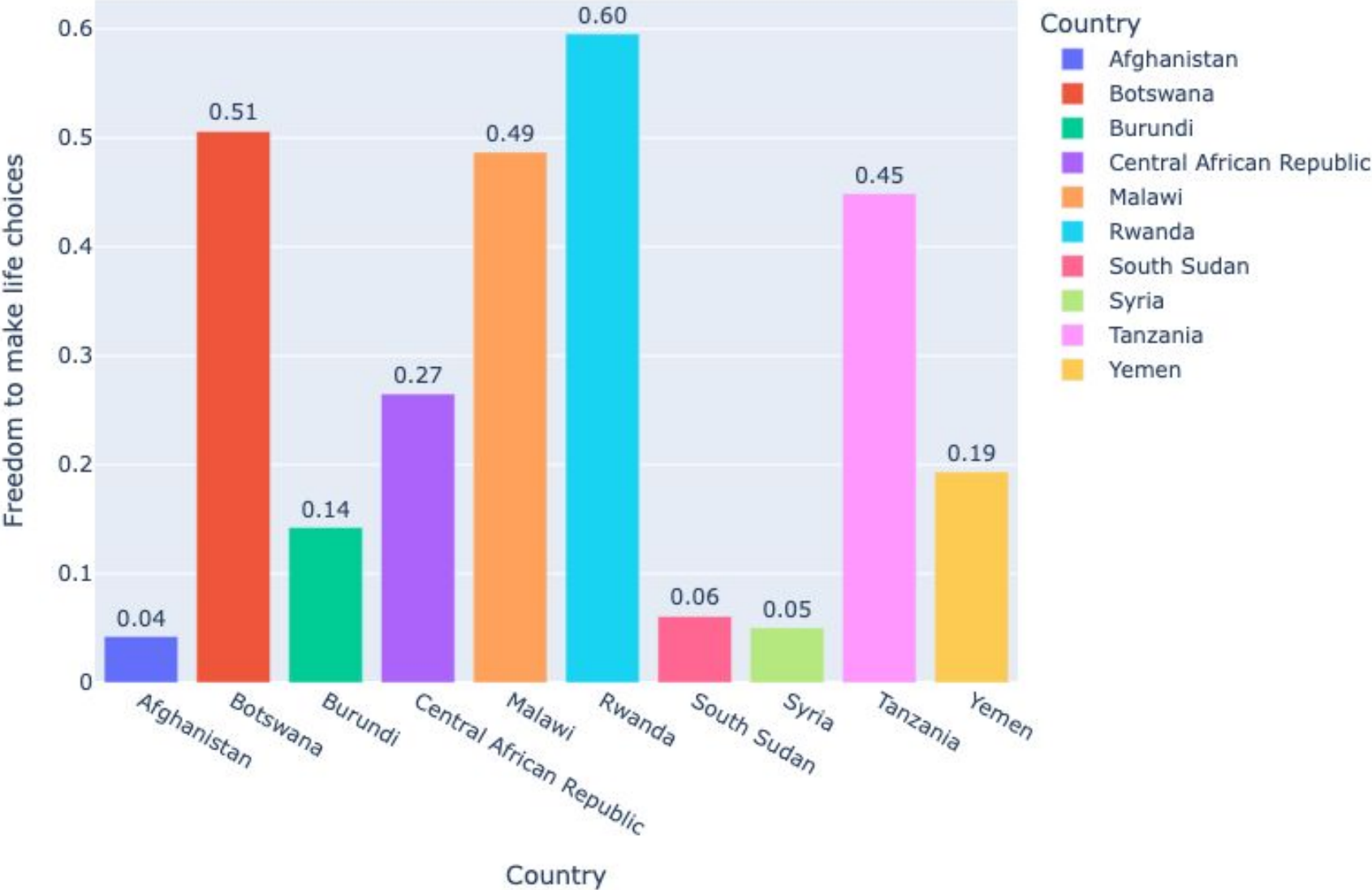
Top 10 Happiest Countries by Freedom and Happiness Score

Top 10 unhappiest Countries by Freedom and Happiness Score

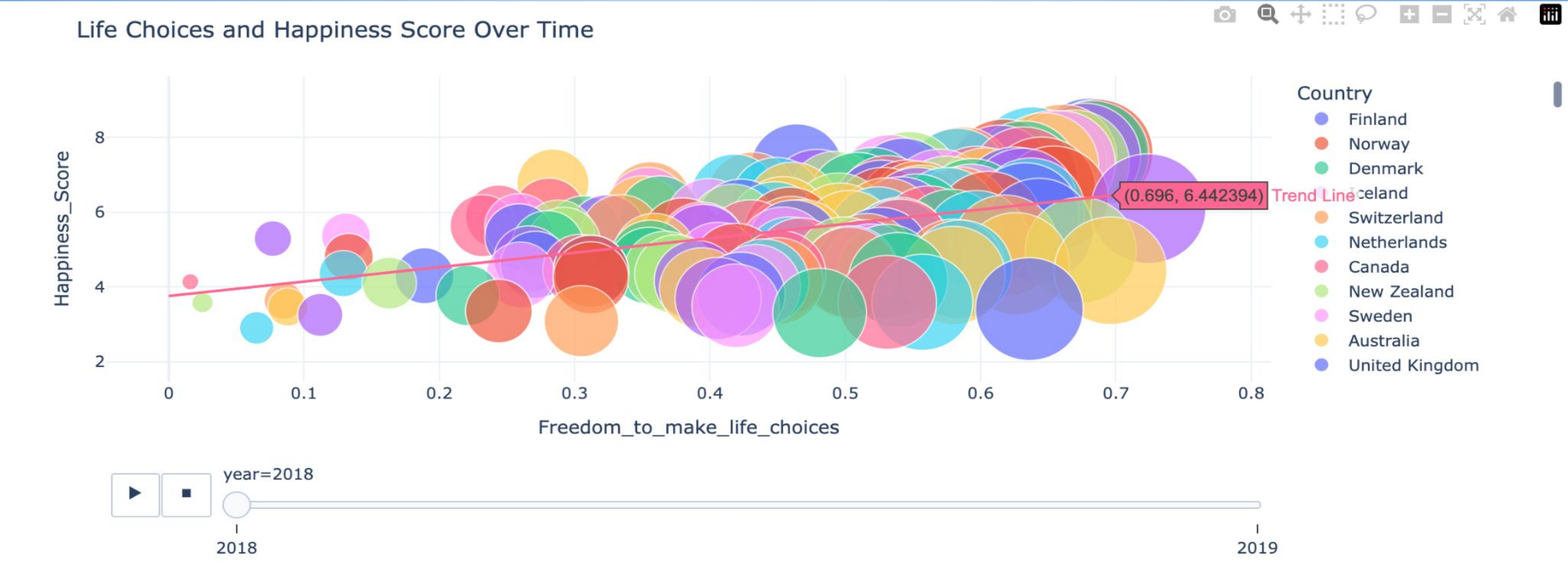
Top 10 Happiest Countries by Freedom to make life choices by Happiness Score



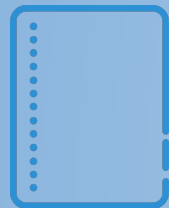
Top 10 Unhappiest Countries by Freedom to make life choices by Happiness Score



Freedom to make life choices and Happiness Score over time

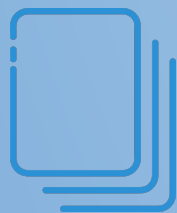


Conclusions



GDP boosts happiness

GDP boosts happiness (0.80 correlation), but financial prosperity alone isn't enough—social and emotional aspects matter too.



Generosity and social support

Generosity and social support—some of the poorest nations report high generosity, suggesting cultural influences on happiness. Places like Myanmar, Haiti, and Indonesia rank high in generosity, but their happiness scores remain low—suggesting that people give despite hardships rather than because they're thriving.



Freedom strongly impacts happiness

Freedom is the strongest predictor of happiness—it's about autonomy, not just wealth. But we need another factors just as Social Support to thrive and be happy.



Happiest vs. unhappiest

Nordic countries exhibit high happiness and freedom, supported by strong welfare systems. Least happy nations struggle with political instability, economic hardship, and low freedom.

**THANK
YOU**

@achitheraptor

Do you have any questions?