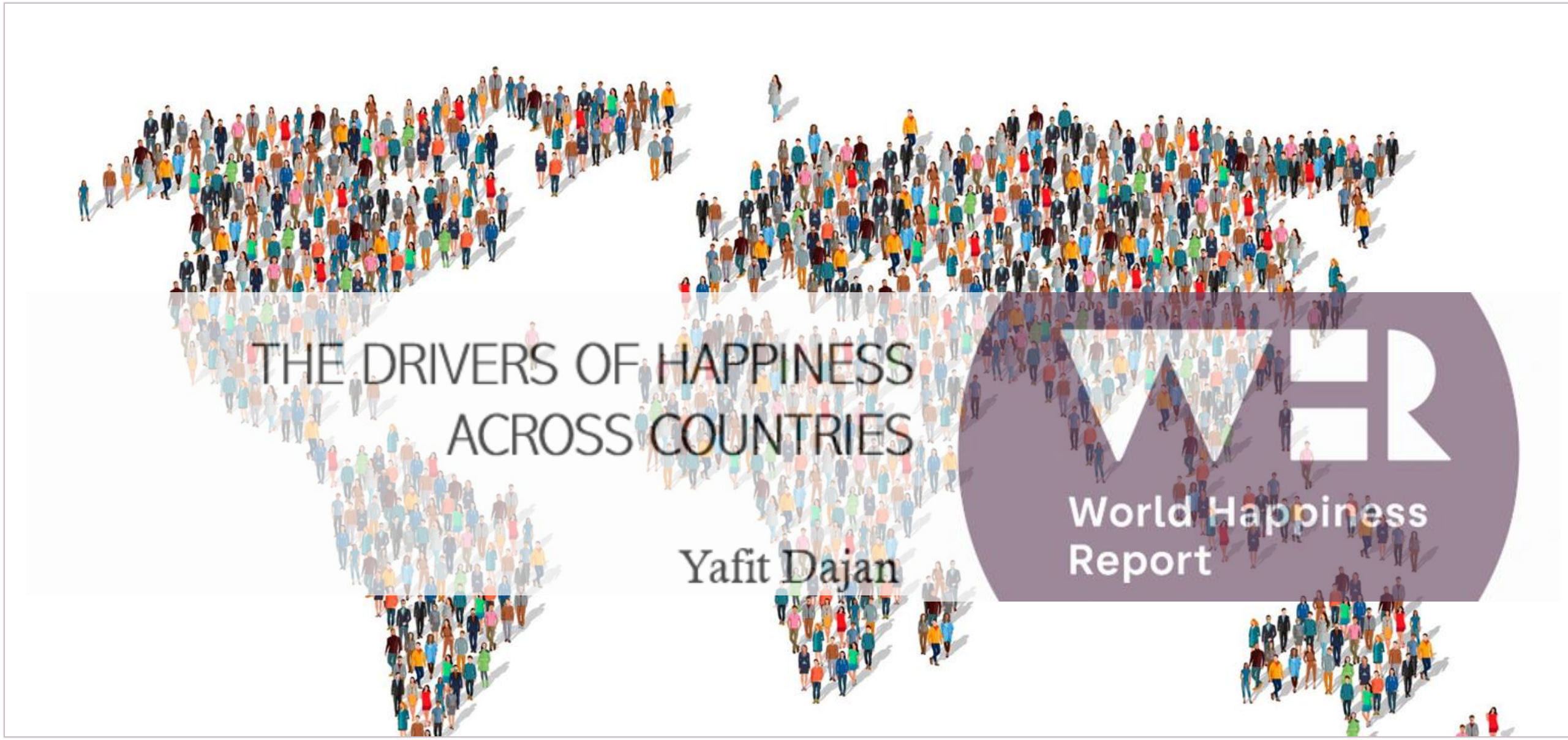

THE DRIVERS OF HAPPINESS
ACROSS COUNTRIES
Yafit Dajan
WR
World Happiness
Report

WHY DOES THE HAPPINESS REPORT MATTER?

- Shifting\Complement Economic Metrics
- Understand Human Flourishing
- Improve Public Policy

DATA

Rank	Country	Score	GDP	Social support	Healthy life expectanc	Freedom to mal	Generosity	Perceptic	of corruptio
1	Finland	7.632	1.305	1.592	0.874	0.681	0.202	0.393	
2	Norway	7.594	1.456	1.582	0.861	0.686	0.286	0.34	
3	Denmark	7.555	1.351	1.59	0.868	0.683	0.284	0.408	
4	Iceland	7.495	1.343	1.644	0.914	0.677	0.353	0.138	
5	Switzerlan	7.487	1.42	1.549	0.927	0.66	0.256	0.357	
6	Netherlanc	7.441	1.361	1.488	0.878	0.638	0.333	0.295	
7	Canada	7.328	1.33	1.532	0.896	0.653	0.321	0.291	
8	New Zealar	7.324	1.268	1.601	0.876	0.669	0.365	0.389	

https://www.kaggle.com/datasets/unsdsn/world-happiness

The dataset is derived from the World Happiness Report,
which collects and aggregates data from surveys conducted across 156 countries.

FEATURES

- Ladder score
- GDP
- Social support
- Healthy life expectancy
- Freedom to make life choices
- Generosity
- Perception of corruption

The ladder score is rated by people on a scale from 0 to 10, while the other variables are based on survey responses or national statistics and are scaled between 0 and 1

RESEARCH QUESTION

What Are the *Key Leading Factors* Behind Happiness Across Countries?

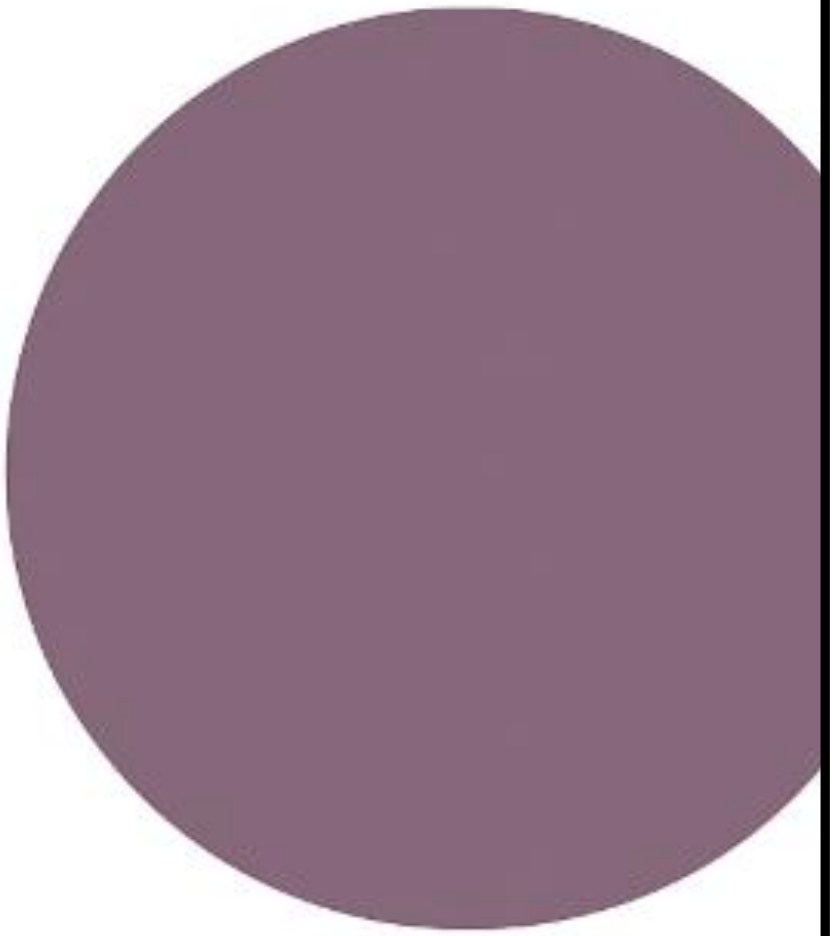
Research plan:

Initial Work:

- Visual data exploration
- Checked for outliers
- Normality testing (Shapiro Wilk test, Histogram and Q-Q plots inspection)
- **Spearman correlation matrix**
- Performed **Mann-Whitney U test**

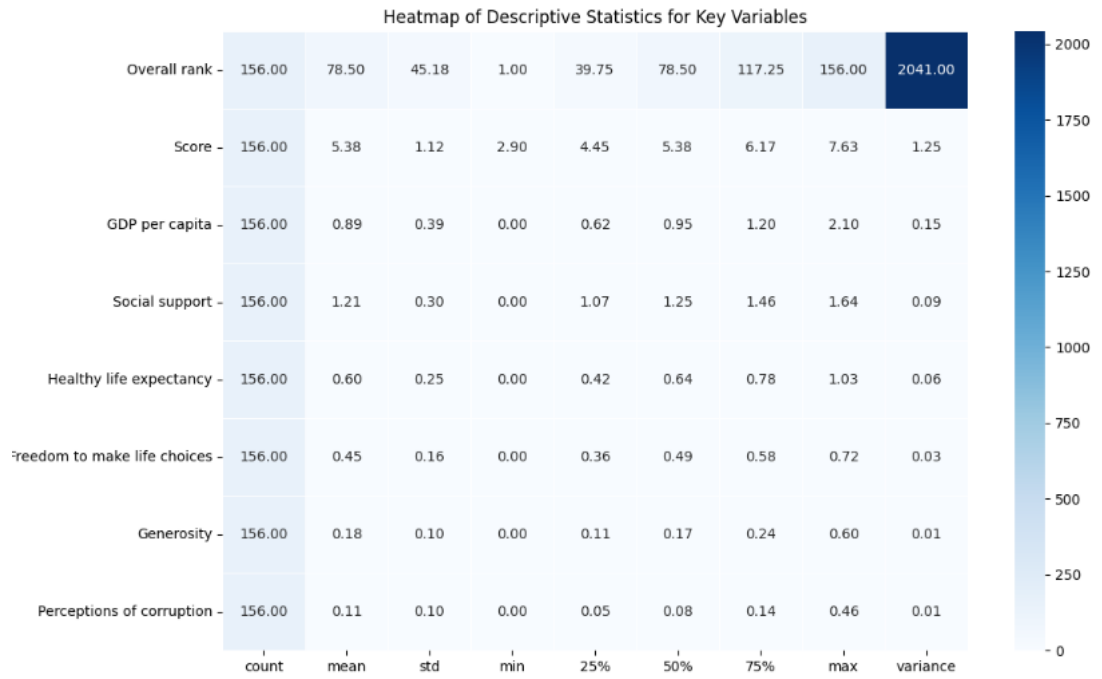
Next steps:

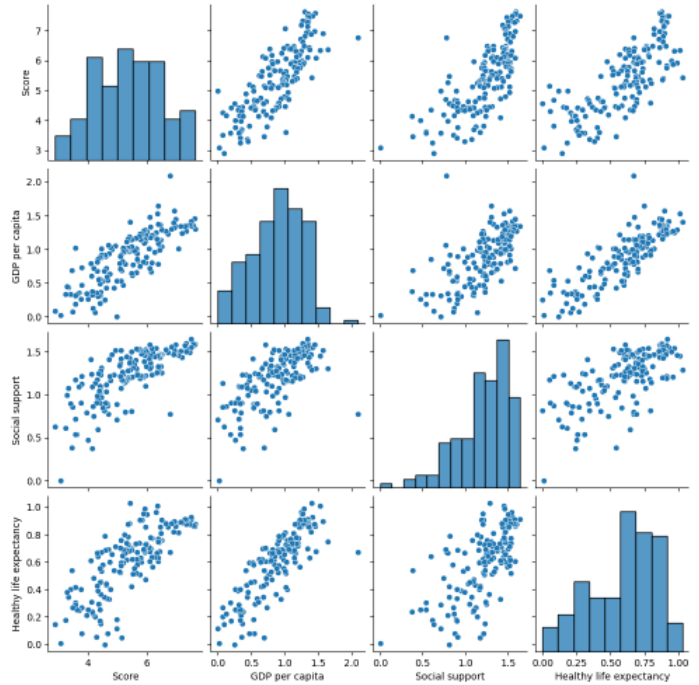
- Fit a **multiple linear regression model** to predict ladder score
- **Factor analysis** and **PCA** and for deeper insights.
- Open to suggestions..



FIRST INSIGHTS

Descriptive statistic



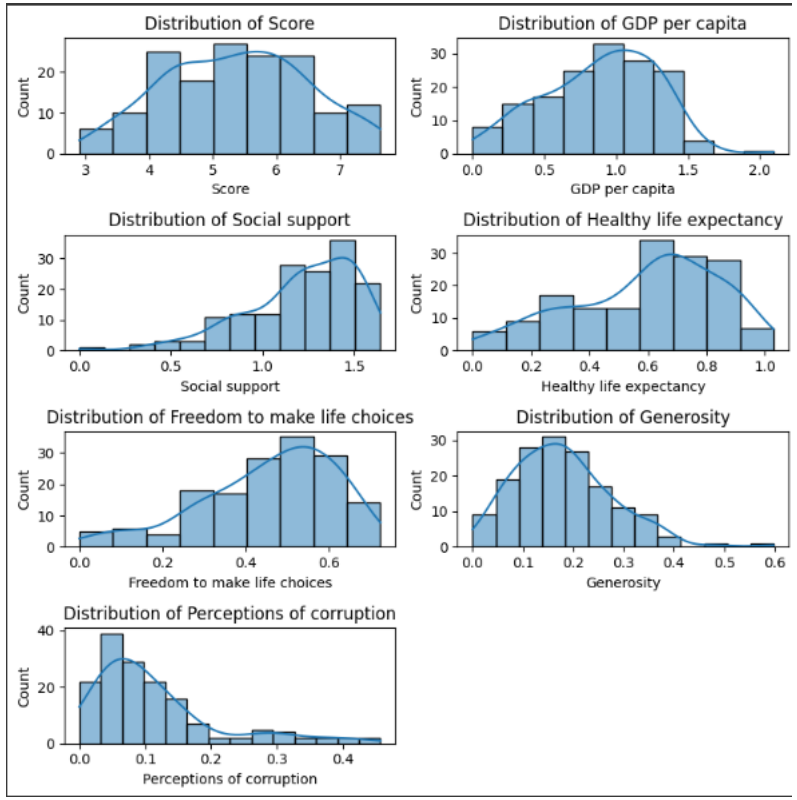


•**Low-to-moderate variability** (CV≤ 50%) – includes score, GDP, social support, health, and freedom.
•**High variability** (CV> 50%) – includes generosity and perceptions of corruption.

•**Positive linear relationships** are visible between **Happiness, GDP, Life expectancy** and **Social support.**

Normality Testing

- Kolmogorov-Smirnov
- *Shapiro-Wilk t*est
- Histograms and Q-Q Plot



4.1 Kolmogorov-Smirnov Test

(H_0): The sample distribution matches the reference normal distribution.
(H_1): The sample distribution does not match the reference normal distribution.

```
Kolmogorov-Smirnov Test for Score: Test Statistic=0.05605680817111691, p-value=0.6895308864938953
Fail to reject H₀: Score appears to be normally distributed (p > 0.05).

Kolmogorov-Smirnov Test for GDP per capita: Test Statistic=0.06842252836460072, p-value=0.4389334233429809
Fail to reject H₀: GDP per capita appears to be normally distributed (p > 0.05).

Kolmogorov-Smirnov Test for Social support: Test Statistic=0.1044977237433945, p-value=0.06159779950659561
Fail to reject H₀: Social support appears to be normally distributed (p > 0.05).

Kolmogorov-Smirnov Test for Healthy life expectancy: Test Statistic=0.10361769279265903, p-value=0.06527112886744046
Fail to reject H₀: Healthy life expectancy appears to be normally distributed (p > 0.05).

Kolmogorov-Smirnov Test for Freedom to make life choices: Test Statistic=0.09366928154810306, p-value=0.12141420175917983
Fail to reject H₀: Freedom to make life choices appears to be normally distributed (p > 0.05).

Kolmogorov-Smirnov Test for Generosity: Test Statistic=0.06943298865363323, p-value=0.42053826539284933
Fail to reject H₀: Generosity appears to be normally distributed (p > 0.05).

Kolmogorov-Smirnov Test for Perceptions of corruption: Test Statistic=0.1755482885735195, p-value=0.00011245982196460581
Reject H₀: Perceptions of corruption does not appear to be normally distributed (p <= 0.05).
```

*Shapiro-Wilk t*est for Generosity

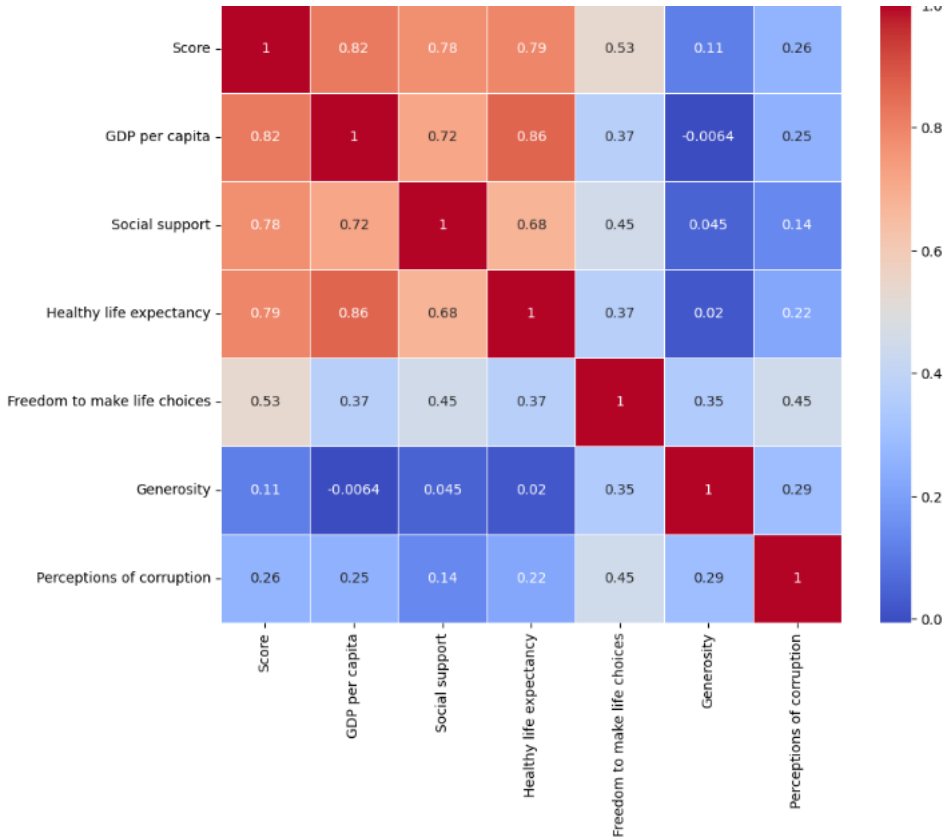
```
Shapiro-Wilk Test for Score: Test Statistic=0.9847009168993887, p-value=0.08279644536822618
Fail to reject H₀: Score appears to be normally distributed (p > 0.05).

Shapiro-Wilk Test for GDP per capita: Test Statistic=0.9783647566293384, p-value=0.014863139765569156
Reject H₀: GDP per capita does not appear to be normally distributed (p <= 0.05).

Shapiro-Wilk Test for Social support: Test Statistic=0.9157709523766362, p-value=7.14616542841997e-08
Reject H₀: Social support does not appear to be normally distributed (p <= 0.05).

Shapiro-Wilk Test for Healthy life expectancy: Test Statistic=0.9553927811486822, p-value=6.741699526378972e-05
Reject H₀: Healthy life expectancy does not appear to be normally distributed (p <= 0.05).

Shapiro-Wilk Test for Freedom to make life choices: Test Statistic=0.9454755060515656, p-value=9.620930219590556e-06
Reject H₀: Freedom to make life choices does not appear to be normally distributed (p <= 0.05).

Shapiro-Wilk Test for Generosity: Test Statistic=0.9580657359730878, p-value=0.00011795787790420376
Reject H₀: Generosity does not appear to be normally distributed (p <= 0.05).

Shapiro-Wilk Test for Perceptions of corruption: Test Statistic=0.8192201046017141, p-value=1.312741153798697e-12
Reject H₀: Perceptions of corruption does not appear to be normally distributed (p <= 0.05).
```

Findings:

Most variables exhibit **non-normal distributions** in the **Shapiro–Wilk test** . Although the **Kolmogorov–Smirnov test** failed to reject the normality assumption for most. The **happiness score** appears to be the only variable that closely follows a normal distribution.

Non-parametric tests

Spearman Correlation Matrix



Mann-Whitney U test

(H_0): The distribution of happiness scores is the same for both groups (e.g., countries with high and low social support).
(H_1): The distribution of happiness scores is different between the two groups.

```
Mann-Whitney U Test for GDP per capita: U-statistic = 5509.0, p-value = 2.2831859588712224e-18
Reject H₀: There is a significant difference in happiness scores between high and low GDP per capita groups.

Mann-Whitney U Test for Social support: U-statistic = 5473.0, p-value = 7.00964721691908e-18
Reject H₀: There is a significant difference in happiness scores between high and low Social support groups.

Mann-Whitney U Test for Healthy life expectancy: U-statistic = 5369.0, p-value = 1.6361291019341514e-16
Reject H₀: There is a significant difference in happiness scores between high and low Healthy life expectancy groups.

Mann-Whitney U Test for Freedom to make life choices: U-statistic = 4555.5, p-value = 8.197418156509714e-08
Reject H₀: There is a significant difference in happiness scores between high and low Freedom to make life choices groups.

Mann-Whitney U Test for Perceptions of corruption: U-statistic = 3685.5, p-value = 0.02163472889022436
Reject H₀: There is a significant difference in happiness scores between high and low Perceptions of corruption groups.
```

```
Mann-Whitney U Test for Generosity: U-statistic = 3270.5, p-value = 0.41901521548437626
Fail to reject H₀: There is no significant difference in happiness scores between high and low Generosity groups.
```

Mann-Whitney U Test:
Most predictors are associated with **notable shifts in happiness scores** when countries are divided by the median level of each factor.
However, **generosity** does **not** show a difference in happiness between high and low generosity groups ($p = 0.42$).

- Strongest correlations with happiness are :
  GDP per capita ($\rho = 0.82$), **Healthy life expectancy ($\rho = 0.79$)**, and **Social support ($\rho = 0.78$)**.
- Weaker correlations:
  Freedom ($\rho = 0.53$), **Perceptions of corruption ($\rho = 0.26$)**, **Generosity ($\rho = 0.11$)**.

Conclusions & Future Work

- **GDP per capita, healthy life expectancy, and social support** appear to be the most influential factors associated with happiness.
- Apply **multiple linear regression** and **PCA** to identify key influencing factors
- Investigate **relationships between variables**, not only with happiness but among themselves
- For a more **objective comparison**, consider using the **Misery Index** (inflation + unemployment) for comparison.
- The survey has some limitation which we will mention later on.