Correlation analysis on World Happiness

*General Topic:* Correlation Analysis on variables that influence a country’s happiness

*Research Question:* Investigate the relationship between variables like GDP per capita, life expectancy, etc on the outcome of United Nation’s World Happiness Report

*Variables:*   
Happiness Score: The Happiness Score of a nation is determined by the responses of survey participants reflecting on whether they perceive themselves as living their optimal lives, where 10 represents the highest level of satisfaction and 0 the lowest. These individual scores are then compiled from samples that are representative of the entire population.

GDP per capita: Logarithmized Gross Domestic Product (GDP) per capita, adjusted for purchasing power parity (PPP), indicates the overall economic output relative to the population and commodities across various countries. This metric, sourced from the World Development Indicators, is standardized to 2017 international dollar values, resulting in measurements expressed in thousands of international dollars.

Social Support: The variable "social support" measures the perception of having reliable individuals to turn to in times of need. The national average is calculated based on responses to the question "If you were in trouble, do you have relatives or friends you can count on to help you whenever you need them?" where answers are recorded as binary values (0 for no, 1 for yes), and then averaged.

Healthy Life Expectancy: The Healthy Life Expectancy at birth is obtained from the data repository of the World Health Organization's (WHO) Global Health Observatory.

Freedom to make life choices: In measuring the freedom to make life choices, respondents were asked with the question "Are you satisfied or dissatisfied with your freedom to choose what you do with your life?" The value assigned to each country represents the national average of responses to this question, with responses recorded in binary format (0 for dissatisfied, 1 for satisfied).

Generosity: A metric between 0 and 1, generosity was measured by asking respondents whether they had donated money to a charity in the past month. The World Happiness Report (WHR) team then conducted a regression analysis of the national average responses against GDP per capita. The residuals from this regression were aggregated to establish a score for each country. This approach takes into account that countries with higher GDPs, indicative of stronger economies, typically exhibit a greater proportion of individuals donating to charity due to their anticipated higher standard of living, while countries with lower GDPs may demonstrate the inverse relationship.

Perceptions of corruption: For the "Perceptions of corruption" variable, respondents were asked two distinct questions as part of the survey: "Is corruption widespread within the government or not?" and "Is corruption widespread within businesses or not?" The composite corruption score is subsequently computed by averaging the responses to these two questions, with values falling between 0 and 1, indicating the extent of perceived corruption.

Analysis:

The World Happiness Report is a report on a country’s happiness by the United Nations, based on data collected from Gallup surveys. In this analysis, we will try to establish which variables affect a country's scores and to what degree, and if variables affect other variables besides the happiness score.

I have used Python and libraries like pandas, numpy, matplotlib and seaborn to draw the analysis. You may find the code on my GitHub [here](https://github.com/dakshashra/happiness_eda).

A graph of blue dots

Description automatically generatedGDP per capita: GDP per capita shows a relatively strong correlation between itself and the happiness score. With a correlation coefficient of 0.79, we can say that it is a strong contributor to the happiness score associated with a country.