The Factors that affect the Happiness of Population

Ersa Mëzuraj

Department of Business Administration, Epoka University, 1032 Tirana, Albania; emezuraj22@epoka.edu.al

Abstract: Happiness Determinants, Multiple Linear Regression, Socioeconomic Factors, Robust Standard Errors

1. Introduction

Happiness is a fundamental aspect of human well-being and societal progress. Research on the factors influencing population happiness is becoming more popular in the social sciences and economics. Key drivers of happiness, such as social support, economic prosperity, health, and governance, have been continuously highlighted in the World Happiness Report (Wang et al., 2024). Nonetheless, the degree to which these elements influence general pleasure differs throughout nations and eras.

This paper seeks to answer the research question: What factors have influenced the happiness of countries around the world? Our objective is to determine the most important factors influencing happiness between 2015 and 2020 by using multiple linear regression analysis. In order to evaluate their influence on happiness ratings globally, this study takes into account a number of factors, including GDP per capita, health, freedom, generosity, government trust, social support, and the corruption perception index (CPI).

Interpreting the link between economic, social, and political issues and population happiness is the main goal of this study. This research might offer insights that might assist policymakers in creating plans to enhance well-being by examining these issues using statistical models and graphical representations.

This study adds to the body of research that happiness is influenced by both concrete and intangible elements, such as social support and trust in the government (Diener & Seligman, 2004). This research attempts to further knowledge of how societal institutions affect human well-being by concentrating on data from 2015 to 2020.

2. Literature review

The causes of happiness have been the subject of several studies, which have shown that a complex interaction between political, social, and economic variables affects well-being. Studies have repeatedly demonstrated that a number of factors, including family, dystopia residual, government trust, social support, GDP per capita, health, freedom, generosity, and the corruption

perception index (CPI), significantly influence happiness levels in different nations. This section examines the proven link between these factors and happiness.

Higher well-being is frequently associated with economic wealth. Happiness and GDP per capita have been found to positively and statistically significantly correlate, suggesting that people in affluent countries are generally happy (A. Easterlin, 1973; Aknin et al., 2013). But according to the Easterlin Paradox, GDP growth enhances well-being, but this impact fades when fundamental requirements are satisfied (Clark et al., 2018). According to research published in the Journal of Economic Behavior & Organization, non-economic variables start to play a bigger role in predicting happiness at a particular income level (Stevenson et al., 2008).

Given that poor health may drastically lower life satisfaction, health is a key factor in determining well-being (Deaton, 2013). Happiness scores are greater in nations with superior healthcare systems and longer life expectancies, according to research from the World Happiness Report (Helliwell et al., 2020).

Happiness has a high correlation with both personal and political liberty. According to the Freedom House Report, life happiness is better in nations with more civil freedoms and political rights (Ronald et al., 2008). Additionally, another paper claims that democratic nations have happier citizens. (Democraty and Happiness, 2022).

Generosity and social capital have a favorable impact on wellbeing. Acts of generosity, including volunteering and charity giving, are linked to higher levels of happiness, according to research published in the Journal of Happiness Studies (Aknin et al., 2013). People who serve others and give to charity tend to experience better levels of life happiness, according to the World Happiness Experience (Helliwell et al., 2020).

A major factor in determining the well-being of society is trust in governmental institutions. Happiness is positively connected with government openness and low levels of corruption, according to research published in the European Journal of Political Economy (Helliwell & Huang, 2006). Higher levels of trust and general contentment are typically seen in nations with stable governments and effective public services (Rothstein & Stolle, 2008).

Life pleasure is greatly impacted by strong social ties. Research shows that nations with stronger support systems and social cohesiveness tend to have happier citizens (Putnam, 2000). According to the World Happiness Report, people who have a support system in place during difficult times routinely score higher on measures of wellbeing (Helliwell et al., 2020).

The degree of corruption in a nation affects public satisfaction and trust. Better CPI ratings, which indicate less corruption, are linked to better levels of happiness, according to research

published in the Journal of Public Economics (Treisman, 2000). According to one study, corruption lowers happiness, although this impact is more noticeable in high-income nations (Arvin & Lew, 2014).

Family structures are important for mental health. Strong familial ties have been found to have a major positive impact on happiness (Lucas & Dyrenforth, 2006). According to Demographic Studies, people who are married and have close family relationships are more satisfied with their lives than people who don't have these support networks (Waite & Gallagher, 2000).

Dystopia residual is the amount of variance in happiness ratings that cannot be explained after taking into consideration known influences. Research indicates that this residue encompasses cultural and psychological factors that impact well-being, such as historical context, optimism, and social resilience (Helliwell et al., 2020).

3. Data & Methodology

This study uses a dataset on worldwide happiness scores from 2015 to 2020 that was acquired via Kaggle ("Happiness and Corruption 2015-2020," 2022). Numerous socioeconomic and political variables that might affect happiness levels are included in the dataset. And they are: 'GDP per capita, health, freedom, generosity, government trust, social support, corruption perception index (CPI), family, and dystopia residual' ("Happiness and Corruption 2015-2020," 2022). Since the Country and Continent variables are categorical and do not directly contribute to numerical regression modeling, they were eliminated in order to guarantee a statistical analysis.

This study uses a Multiple Linear Regression (MLR) model to investigate the factors that influence happiness. This model enables the investigation of the combined effects of independent variables on the dependent variable, happiness score. The general form of the regression model is:

Happiness_Score = β0 + β1GDP_per_capita + β2Health + β3Freedom + β4Generosity + β5Government_Trust + β6Social_Support + β7CPI + β8Family + β9Dystopia_Residual + ε

β0 ⟶ the intercept

βi ⟶ the coefficients of the independent variables

Happiness_Score ⟶ The dependent variable. Average of answers to the main life assessment question.

GDP_per_capita → The independent variable. Economic Prosperity (USD per capita)

Health → The independent variable. Health Services (Index Score)

Freedom → The independent variable. Personal Autonomy (Index Score)

Generosity → The independent variable. Charitable Donations (Index Score)

Government_Trust → The independent variable. Political Confidence (Index Score)

Social_Support → The independent variable. Community Networks (Index Score)

CPI → The independent variable. Corruption Perception Index (Index Score)

Family → The independent variable. Household Stability (Index Score)

Dystopia_Residual → The independent variable. Unexplained Happiness (Score)

Several diagnostic tests were carried out in accordance with accepted econometric guidelines in order to confirm the reliability of the regression results.

High levels of correlation between independent variables can lead to **multicollinearity**, which can skew regression estimates. To evaluate this problem, the **Variance Inflation Factor (VIF)** is employed. A VIF number above 10 denotes severe multicollinearity, but values below this cutoff point imply that multicollinearity is not a major problem (O'Brien, 2007). All of the VIF values in this investigation were less than 10, indicating that multicollinearity had no effect on the model.

Regression residuals must have a **normal distribution** in order to be used for legitimate hypothesis testing. Applying the **Kolmogorov-Smirnov** test as advised by Razali & Yap (2011) produced a p-value larger than 0.05, which denotes normalcy. The Jarque-Bera test, which evaluates whether the skewness and kurtosis of residuals resemble those of a normal distribution, was also carried out (Jarque & Bera, 1980). However, the findings indicated departures from normalcy.

A fundamental tenet of linear regression is broken when the variance of the residuals is not constant, a phenomenon known as **heteroscedasticity**. Strong evidence of heteroscedasticity was found using the **Breusch-Pagan test**, which was suggested by Breusch & Pagan (1979) (p-value < 0.05). In order to rectify this, the White heteroscedasticity-consistent covariance estimator (White, 1980) was used to apply robust standard errors.

One of the most important assumptions in regression models is **linearity**. To determine if the model is accurately stated, the **Ramsey RESET test**—which was first presented by Ramsey (1969)—was used. The regression model is appropriately stated and exhibits a linear connection, according to the results (p-value > 0.05).

A **Weighted Least Squares (WLS)** regression model was generated in order to further address heteroscedasticity. By giving observations weights based on residual variance, WLS can assist in correcting heteroscedasticity (Greene, 2003). Nevertheless, the WLS model's Breusch-Pagan test continued to show heteroscedasticity, which prompted the adoption of robust standard errors.

Boxplots, histograms, and quantile-quantile (Q-Q) plots were examined in order to better investigate residual behavior. Gelman and Hill (2006) assert that these representations are useful instruments for determining normalcy. The results supported the use of robust standard errors by confirming that residuals were roughly normal with a few slight deviations.

4. Results & Discussion

The **descriptive statistics** reveal interesting insights about global happiness determinants. The happiness score ranges from 2.57 to 7.81, with an average of 5.47, indicating that most countries fall in the mid-range of the happiness scale. GDP per capita and social support show significant variation, suggesting economic prosperity and community support are crucial but unequally distributed factors. Notably, government trust has a low median value (0.089), reflecting widespread skepticism toward institutions, which may negatively impact overall well-being.

Here we have a collection of six histograms showing the distribution of happiness scores across different continents. The distributions vary across continents, revealing differences in reported happiness levels. Interestingly, the lowest score bin (around 2-3) has a surprisingly low frequency in most continents, suggesting very few people report extremely low happiness levels.

Distribution of Happiness Score by Continent
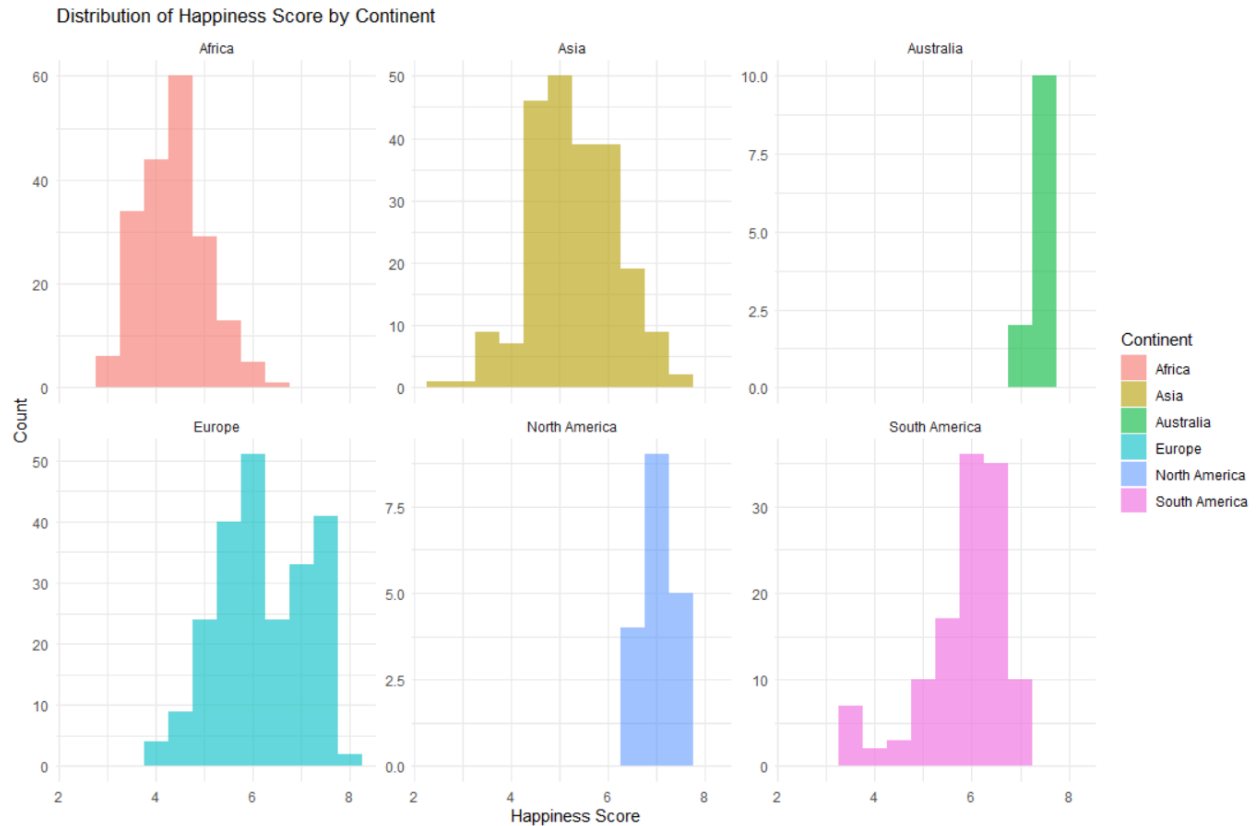
The initial regression model was estimated without adjustments for heteroscedasticity. The table below presents the coefficients, standard errors, t-values, and p-values for each independent variable:

Variable	Estimate	Std. error	T-value	Pr(>|t|)
Intercept	1.4827	0.0793	18.697	< 2e-16
GDP per capita	1.0479	|0.0857	12.222	< 2e-16
Health	0.9074	0.1291	7.031	4.46e-12
Freedom	0.9235	0.1477	6.254	6.58e-10
Generosity	0.8258	0.1543	5.353	1.14e-07
Government Trust	0.7203	0.2079	3.465	.000559
Social Support	1.1454	0.0792	14.461	< 2e-16

CPI Score	0.0038	0.0015	2.516	.516
Dystopia Residual	0.3445	0.0210	16.373	< 2e-16
Family	0.8475	0.0837	10.124	< 2e-16

R-squared = 0.831
Adjusted R-squared = 0.8291
F-statistic = 427.2 (p-value < 2.2e-16)

Variable	VIF
GDP per capita	4.0022
Health	3.2861
Freedom	1.7469
Generosity	1.2886
Governmnet Trust	1.8791
Social Support	9.4222
CPI Score	3.1171
Family	7.8314
Dystopia Residual	1.8720

Since Social Support (VIF = 9.42) and Family (VIF = 7.83) have relatively high VIF values, potential multicollinearity should be considered, but it does not exceed the critical threshold of 10.

This image is a **correlation matrix** visualizing the relationships between different factors potentially influencing happiness scores. It suggests several key relationships between factors and happiness. GDP per capita, health, freedom, generosity, government trust, social support, and family all appear to have positive correlations with the happiness score, indicated by the red hues. This implies that as these factors increase, happiness scores tend to increase as well. The strongest positive relationships (darker red, larger circles) seem to be between the happiness score and factors like GDP per capita, health, and social support.
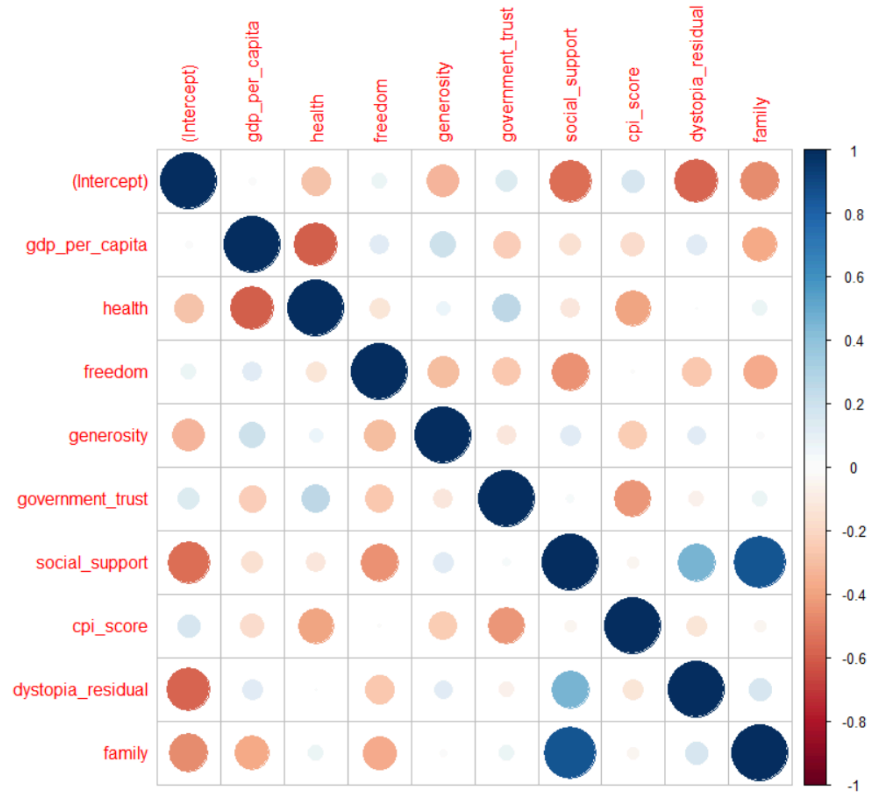
Breusch-Pagan Test: The p-value (0.00017) is less than 0.05, indicating strong evidence of heteroscedasticity in the model.

Linearity Test (Ramsey RESET Test): p-value = 0.2537 (> 0.05), indicating that the model is correctly specified and follows a linear relationship.

Kolmogorov-Smirnov Test: p-value = 0.5402 (> 0.05), indicating residuals are normally distributed.

Jarque-Bera Test: The test results indicate minor deviations from normality, suggesting that some residuals may be skewed.

Weighted Least Squares (WLS) Regression: Even after applying WLS, heteroscedasticity persisted, suggesting that robust standard errors should be used.

This is a **residual plot,** a diagnostic tool in regression analysis. Its purpose is to examine the quality of a linear regression model by plotting the residuals against the fitted values. Ideally, residuals should be randomly scattered around zero, indicating that the model's assumptions are met. Patterns in the residual plot suggest minor problems with the model, such as non-constant variance (heteroscedasticity) or outliers.

Residuals vs Fitted
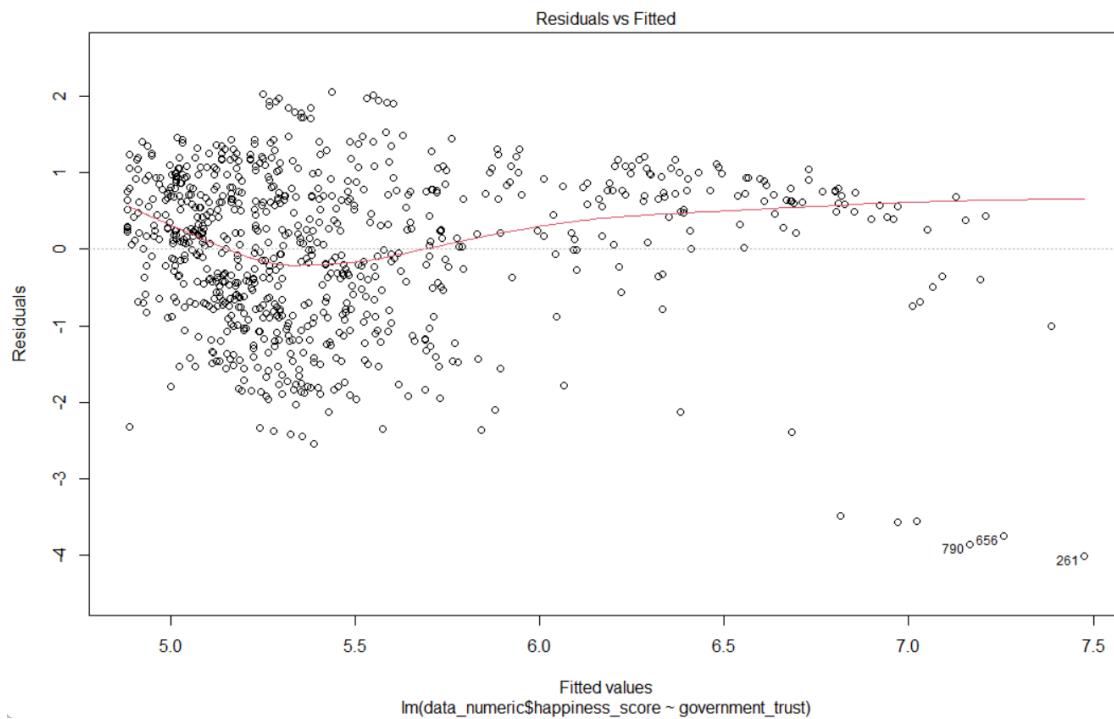
The plot below shows the quantiles of the standardized residuals against the quantiles of the theoretical normal distribution. If the residuals are normally distributed, the points on the **Q-Q plot** should fall approximately along the diagonal reference line. Deviations from this line indicate departures from normality.
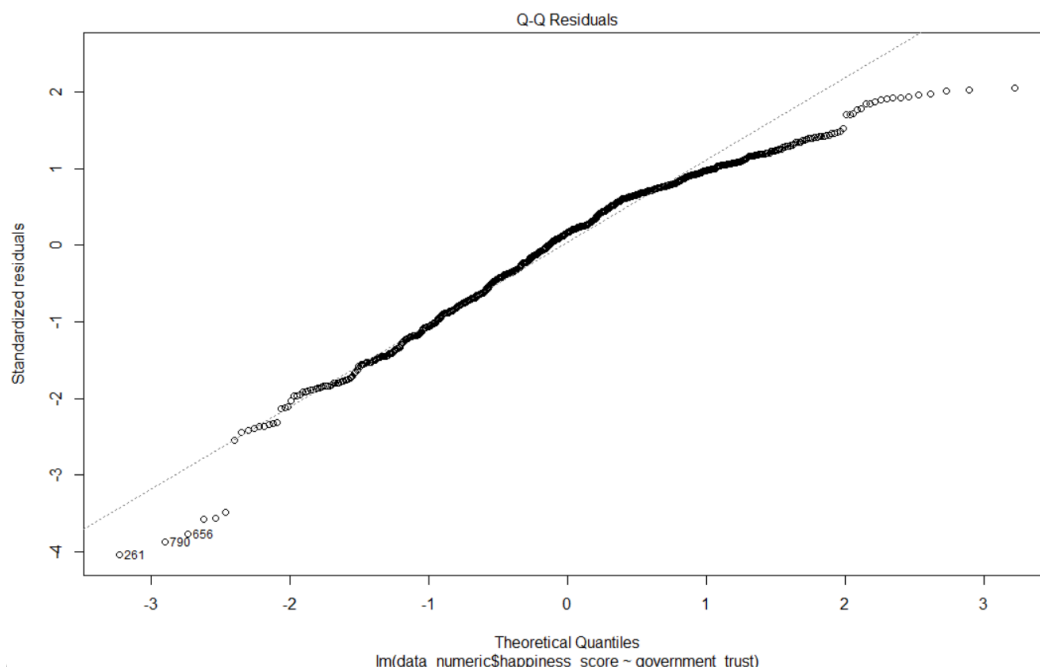
Q-Q Residuals

The next **histogram** displays the distribution of the residuals. Ideally, this histogram should be roughly bell-shaped (normal distribution) and centered around zero, indicating that the model's assumptions of normally distributed and unbiased errors are met. Still, the Kolmogorov-Smirnov Test shows that the residuals are normally distributed.

Residual Histogram
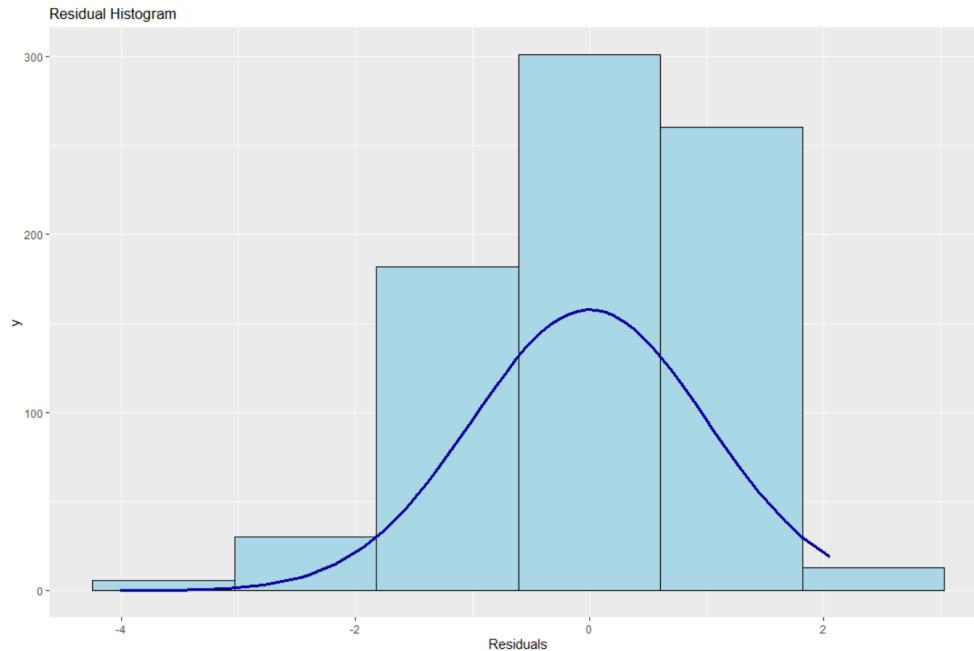
Since heteroscedasticity remained even after WLS, we used **robust standard errors** to improve the reliability of the coefficient estimates. The robust standard errors model is selected as the final regression model since it accounts for heteroscedasticity and provides more reliable statistical inference. All variables are statistically significant, as indicated by p-values below 0.05. Also, they are in positive correlation with the dependent variable. GDP per capita, social support, and dystopia residual have the strongest positive impact on happiness scores. Government trust and CPI score have weaker but still significant effects on happiness. The adjusted R-squared value of 0.8291 suggests that the model explains a high proportion of variance in happiness scores.

Variable	Estimate	Std. error	T-value	Pr(>|t|)
Intercept	1.4827	0.0844	17.567	< 2.2e-16
GDP per capita	1.0479	|0.0826	12.687	< 2.2e-16
Health	0.9074	0.1473	6.161	1.16e-09
Freedom	0.9235	0.1437	6.254	2.29e-10

Generosity	0.8258	0.1577	5.237	2.10e-07
Government Trust	0.7203	0.2317	3.109	0.001945
Social Support	1.1454	0.0741	15.466	< 2.2e-16
CPI Score	0.0038	0.0015	2.462	0.0140
Dystopia Residual	0.3445	0.0231	14.910	< 2.2e-16
Family	0.8475	0.0756	11.216	< 2.2e-16

The final regression model is presented below:

Happiness_Score = 1.4827 + 1.0479GDP_per_capita + 0.1473Health + 0.9235Freedom + 0.8258Generosity + 0.7203Government_Trust + 1.1454Social_Support + 0.0038CPI + 0.8475Family + 0.3445Dystopia_Residual

5. Conclusion

This research looked at the main political and economic determinants of population happiness throughout the world. Using multiple linear regression analysis, we found that family stability, freedom, health, social support, and GDP per capita all significantly influence happiness scores. The results show that while GDP per capita is significant, well-being is also influenced by generosity, strong social networks, and government confidence. Low institutional trust and the sense of corruption, however, continue to be problems, reflecting larger societal issues that affect happiness. The findings imply that social support and the trust of governance are just as important to well-being as material prosperity.

The study emphasizes the necessity of well-rounded policymaking that supports social trust, political freedom, affordable healthcare, and economic stability. Nations with high happiness scores typically place a larger priority on community support and just government, going beyond economic metrics to promote human well-being. Resolving corruption and institutional trust may greatly increase social satisfaction. Future studies might examine the psychological and cultural aspects of happiness, expanding on our knowledge of what actually improves people's quality of life throughout the world.

References

Wang, Shiplett, Huang, & Helliwell. (2024, August 15). *The World Happiness Report*. https://worldhappiness.report/

Diener, E., & Seligman, M. E. (2004). *Beyond money.* Psychological Science in the Public Interest, 5(1), 1–31. https://doi.org/10.1111/j.0963-7214.2004.00501001.x

Aknin, Barrington-Leigh, Dunn, Helliwell, Burns, Biswas-Diener, Kemeza, Nyende, Ashton-James, & Norton. (2013, January 18). *Prosocial spending and well-being: cross-cultural evidence for a psychological universal*. National Institutes of Health (NIH). https://pubmed.ncbi.nlm.nih.gov/23421360/

A. Easterlin. (1973). *Does Economic Growth Improve the Human Lot? Some Empirical Evidence*. ScienceDirect. https://www.sciencedirect.com/science/article/abs/pii/B9780122050503500087

Clark, Fleche, Layard, Powdthavee, & Ward. (2018, February 3). *The origins of happiness.* Princeton University Press. https://press.princeton.edu/books/hardcover/9780691177892/the-origins-of-happiness?srsltid=AfmBOorXDDaQqz64L5QrqXYacRKzsfIeJwTeHka-Xuw6_MkNUiFeJx-L

Stevenson, B., Becker, G. S., Wolfers, J., & Rayo, L. (2008, March 1). *Economic Growth and Subjective Well-Being: Reassessing the Easterlin Paradox*. Brookings. https://www.brookings.edu/articles/economic-growth-and-subjective-well-being-reassessing-the-easterlin-paradox/

Deaton. (2013, September 24). *The Great Escape*. Princeton University Press. https://press.princeton.edu/books/hardcover/9780691153544/the-great-escape?srsltid=AfmBOop6PtBd-RlU2ObJe5Ple3LaJ6iEhioXVWptIh7yxXzceIh6E-Ow

Helliwell, J. F., Layard, R., Sachs, J. D., & De Neve, J. (2020, March 20). *World Happiness Report 2020*. The World Happiness Report. https://worldhappiness.report/ed/2020/

Ronald, Foa, Peterson, & Welzel. (2008, January). *Development, Freedom and Rising Happiness: A Global Perspective 1981-2006*. Research Gate. https://www.researchgate.net/publication/235643983_Inglehart_R_R_Foa_C_Peterson_C_Welzel_2008_Development_Freedom_and_Rising_Happiness_A_Global_Perspective_1981-2006_Perspectives_on_Psychological_Science_3_4264-85

Happiness and democracy. (2022, March 18). *Action for Happiness*. https://actionforhappiness.org/happiness-and-democracy

Helliwell, & Huang. (2006, February). *How's your government? International evidence linking good government and Well-Being*. Research Gate. https://www.researchgate.net/publication/5186775_How's_Your_Government_International_Evidence_Linking_Good_Government_and_Well-Being

Rothstein, & Stolle. (2008, July). *The State and Social Capital: An Institutional Theory of Generalized Trust*. Research Gate. https://www.researchgate.net/publication/272213133_The_State_and_Social_Capital_An_Institutional_Theory_of_Generalized_Trust

Putnam. (2000). *Bowling alone: The collapse and revival of American community*. APA PsycNet. https://psycnet.apa.org/record/2001-01469-000

Treisman. (2000, June). *The causes of corruption: a cross-national study*. ScienceDirect. https://www.sciencedirect.com/science/article/abs/pii/S0047272799000924

Arvin, & Lew. (2014, May). *Does income matter in the happiness-corruption relationship?* Research Gate. https://www.researchgate.net/publication/262576984_Does_income_matter_in_the_happiness-corruption_relationship

Lucas, & Dyrenforth. (2006). *Does the existence of social relationships matter for subjective Well-Being?* Semantic Scholar. https://www.semanticscholar.org/paper/Does-the-Existence-of-Social-Relationships-Matter-Lucas-Dyrenforth/3b97fbd3aa28061f7deda39f1508192a7ca199c0

Waite, & Gallagher. (2000). *The case for marriage: Why married people are happier, healthier, and better off financially*. PsychPage. http://www.psychpage.com/family/brwaitgalligher.html

Happiness and Corruption 2015-2020. (2022). [Dataset]. In *Kaggle*. Elias Turk. https://www.kaggle.com/datasets/eliasturk/world-happiness-based-on-cpi-20152020?select=WorldHappiness_Corruption_2015_2020.csv

O'brien, R. M. (2007). *A caution regarding rules of thumb for variance inflation factors*. Quality & Quantity, 41(5), 673–690. https://doi.org/10.1007/s11135-006-9018-6

Razali, & Yap. (2011, January). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. Research Gate. https://www.researchgate.net/publication/267205556_Power_Comparisons_of_Shapiro-Wilk_Kolmogorov-Smirnov_Lilliefors_and_Anderson-Darling_Tests

Jarque, & Bera. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. Elsevier. https://www.sciencedirect.com/science/article/abs/pii/0165176580900245

Breusch, & Pagan. (1979, September). A simple test for heteroscedasticity and random coefficient variation. JSTOR. https://www.jstor.org/stable/1911963

Ramsey. (1969). Tests for specification errors in classical linear Least-Squares regression analysis. JSTOR. https://www.jstor.org/stable/2984219

Greene, W.H. (2003) Econometric Analysis. 5th Edition, Prentice Hall, Upper Saddle River. - References - Scientific Research Publishing. (n.d.). https://www.scirp.org/reference/referencespapers?referenceid=1875244

Gelman, A., & Hill, J. (2006). Data analysis using regression and Multilevel/Hierarchical models. https://doi.org/10.1017/cbo9780511790942