

# PSTAT100 Final Project Report

George Castillo (g\_castillo), Brandon Yang (brandonyang), Anush Harish(anushharish)

July 30th 2025

## ***Abstract***

This report, conducted by three college-leveled data scientists, investigates the impact of amassed political tension within a selected number of countries that ultimately resulted in civil war. Collecting data from 2005 to 2023, our report draws upon the World Happiness Report where we explore trends on whether countries undergoing civil war experienced a decline in happiness and what factors contribute to this decline. By combining linear regression modeling, data clustering, principal component analysis, and other data science techniques, we are able to highlight common and altering patterns in how war affects life satisfaction. Ultimately, this report will serve to offer insight on which nations fare better in times of unrest and why.

## ***Introduction***

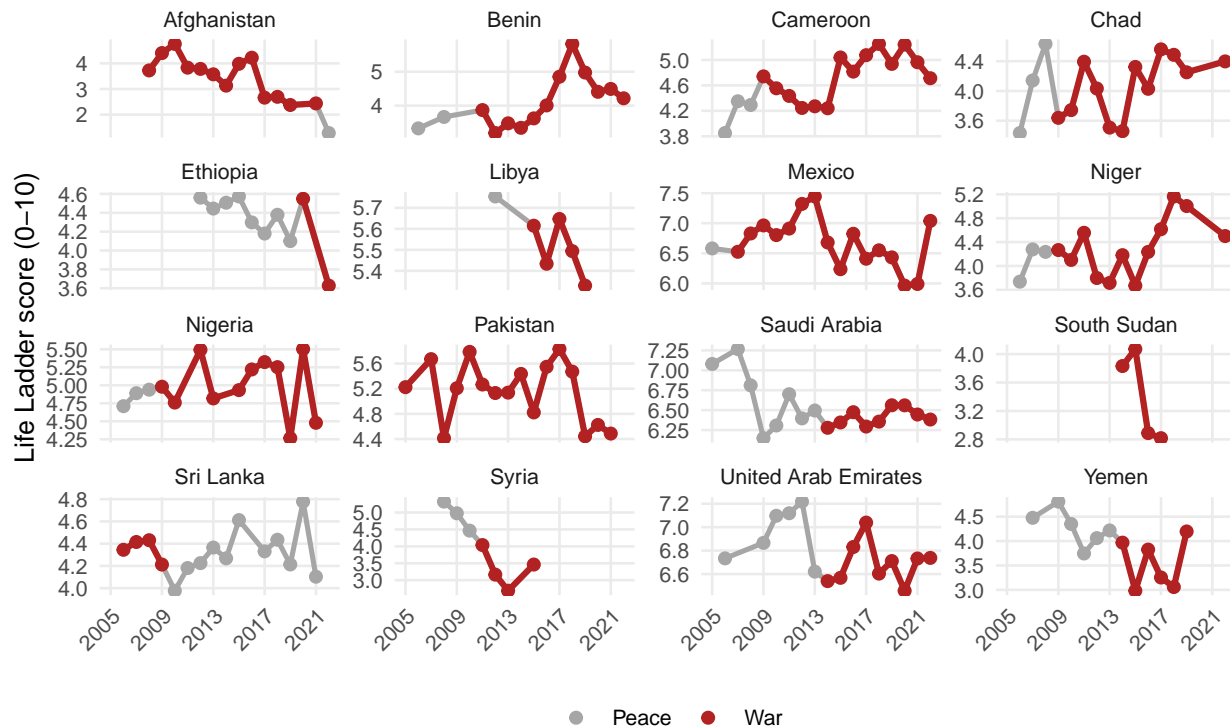
Happiness is often considered an indicator of well-being and can be measured in a variety of different ways, with quality of life satisfaction being one of them. The World Happiness Report, conducted by the Wellbeing Research Centre, and the UN Sustainable Development Solutions Network, collects yearly survey data from over 100 countries, capturing how citizens of these countries rate their overall life satisfaction on a scale from 0 to 10, along with other economic and social indicators.

In this report, we have collected data from the 2023 World Happiness Report on focused on countries that have experienced civil war between 2008 and 2023 to ask an relevant question: Do all countries suffer from civil war the same? What are the greatest factors in maintaining or reducing happiness in a country undergoing civil war?

To answer this we will go through data that has been aligned with civil war periods (as determined by new reports and government officials) and pick at certain variables found within the 2023 World Happiness Report database. We will compare trends across nations and use statistical modeling to identity which variables are among the strongest when it comes to predicting happiness during wartime. By looking and analyzing these patterns throughout different conflict zones, we aim to better understand the human experience of war beyond the causalities and political ramifications, and towards the emotional cost for everyday citizens.

## Life-Ladder Trends with War Years Highlighted

Grey = peace-time · Red = conflict years (2005–2023)



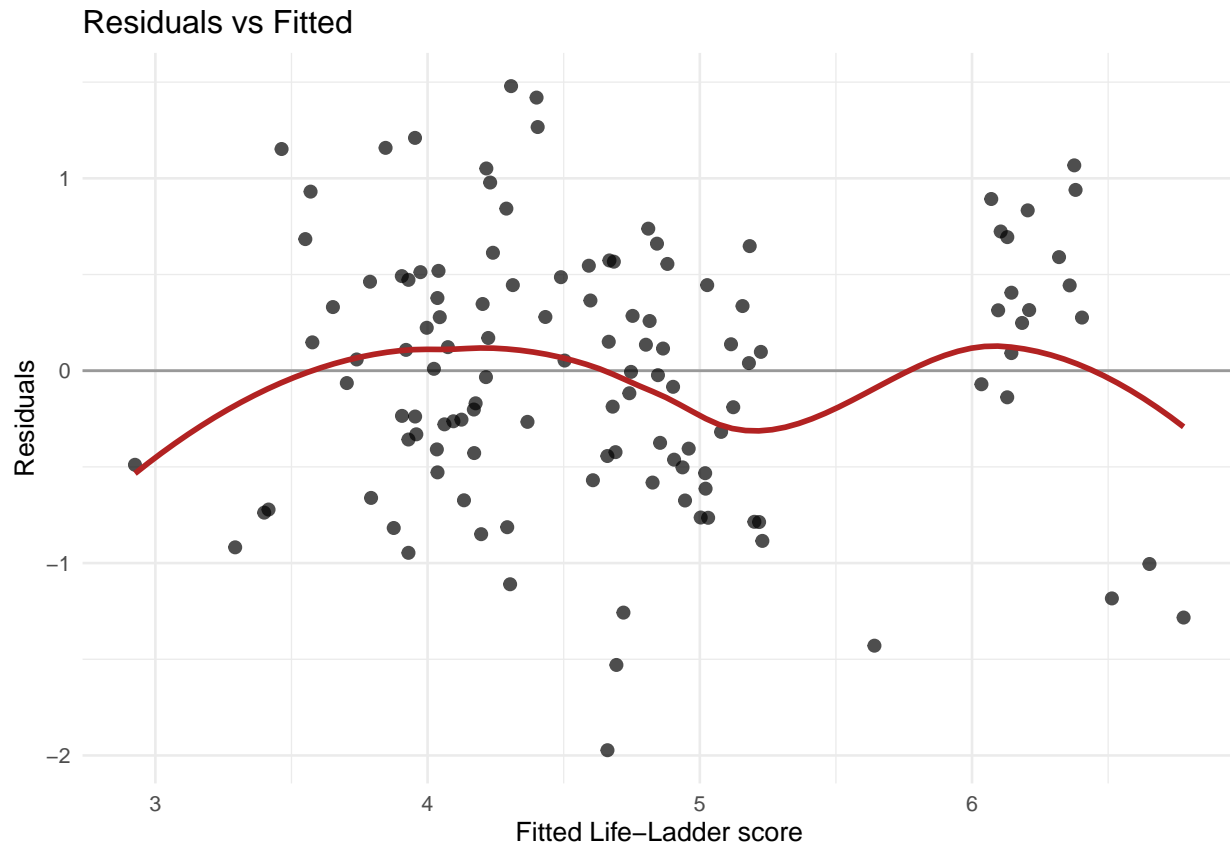
This first plot lays the general foundation of what exactly we are comparing. We are given 16 different countries that all run through the same time frame. Each country has a reported Life Ladder score which is the measure of self-reported life satisfaction scaled from 0 to 10. Each country contains two sets of lines and dots where the gray color represents peace-time trends and the red color represents civil conflict. A closer look reveals other important pieces of information. Countries such as Syria, Yemen, and South Sudan show sharp and prolonged declines in life satisfaction that align closely with the onset and continuation of armed conflict. In these cases, war appears to have a significant effect on subjective wellbeing. On the other hand, countries like Mexico and Pakistan show relatively stable life satisfaction scores despite enduring extended periods of internal violence or instability. This suggests that the impact of civil war on happiness is not universally felt to the same degree.

Clearly there is a trend in satisfaction levels when a country is in peacetime compared to when its at war. But while looking purely at the plots we can make certain conclusions about specific countries, there do seems to be internal vairables that are better explanations as to why certain trends differ greatly to other countries. For this reason, it is important to discuss these variables through other means.

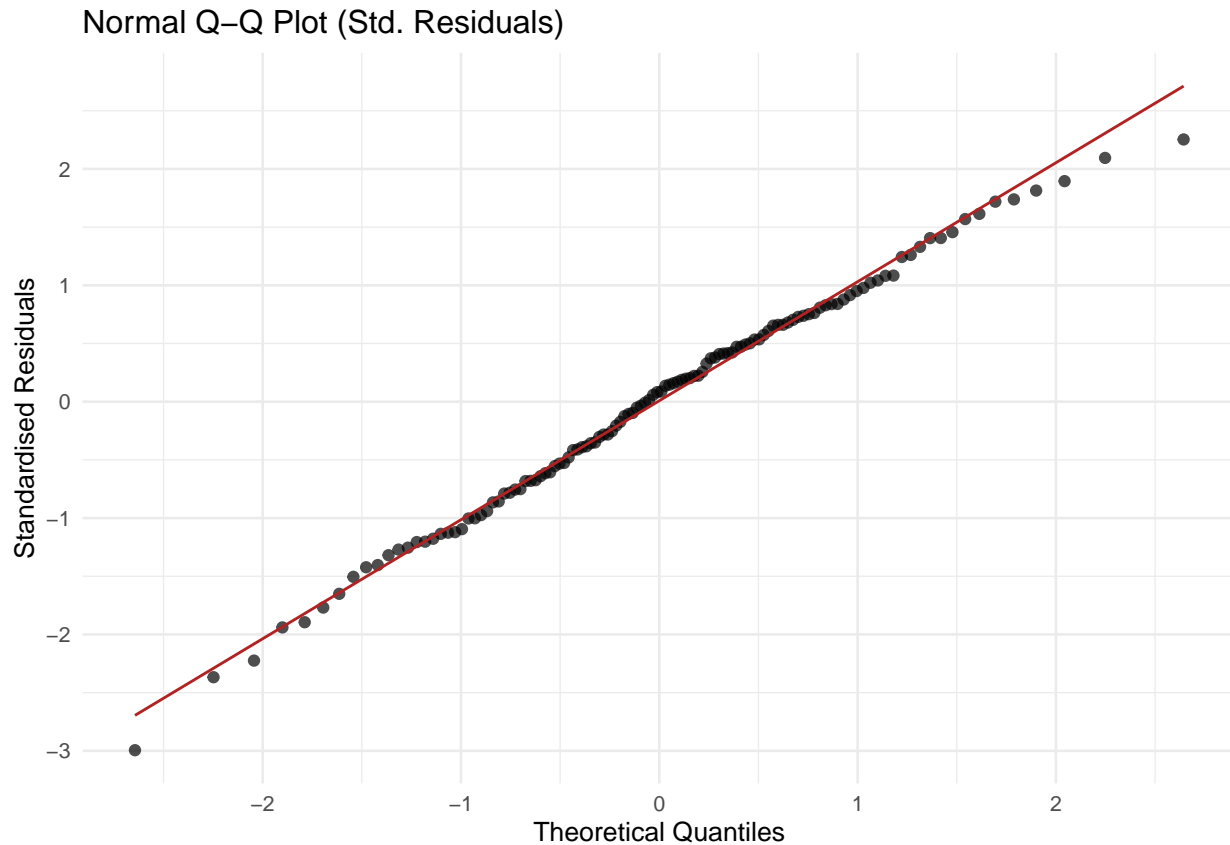
Table 1: OLS Coefficients (unstandardised).

term	estimate	std.error	statistic	p.value	conf.low	conf.high
(Intercept)	0.32	1.20	0.27	0.79	-2.05	2.69
log_gdp	1.02	0.14	7.52	0.00	0.75	1.29
social	1.55	0.59	2.65	0.01	0.39	2.72
life_exp	-0.06	0.02	-2.53	0.01	-0.11	-0.01
freedom	0.71	0.55	1.28	0.20	-0.39	1.80
generosity	-0.07	0.61	-0.11	0.91	-1.29	1.15
corruption	-2.56	0.78	-3.28	0.00	-4.11	-1.02

Using this regression model, we are able to identify which factors influence happiness during war time. For this model we implemented only the years in which countries were actively involved in civil wars. The response variable in this model is the Life Ladder score with its predictors including six key components from the World Happiness Report: **log\_gdp** which tracks the GDP per capita reflecting economic performance, **social** which denotes the social support received by the person, **life\_exp** which shows life expectancy at birth, **freedom** which ties to the perceived freedom to make choices, **generosity** which tracks reports of prosocial behavior, and **corruption** which looks into the perceived corruption in a government. It is important to note that variables were renamed for clarity and rows with missing values were excluded in this model. The table presents estimated effect of each predictor on life satisfaction and we can see some interesting trends that arise. Social Support and freedom to make life choices prove to be the most robust positive association with life satisfaction, showing the extent in which internal aid and the right to freedoms have on a persons happiness. Log GDP per capita and health life expectancy also so this positive trend, further solidifying a relationship being economy and healthcare. On the other end, Generosity and perceived corruption show weaker effects on a persons happiness which may reflect inconsistencies on how these variables are experienced during wartime. Overall, the model emphasizes that subjective wellbeing in war is swayed by the strength of social support and individualism.

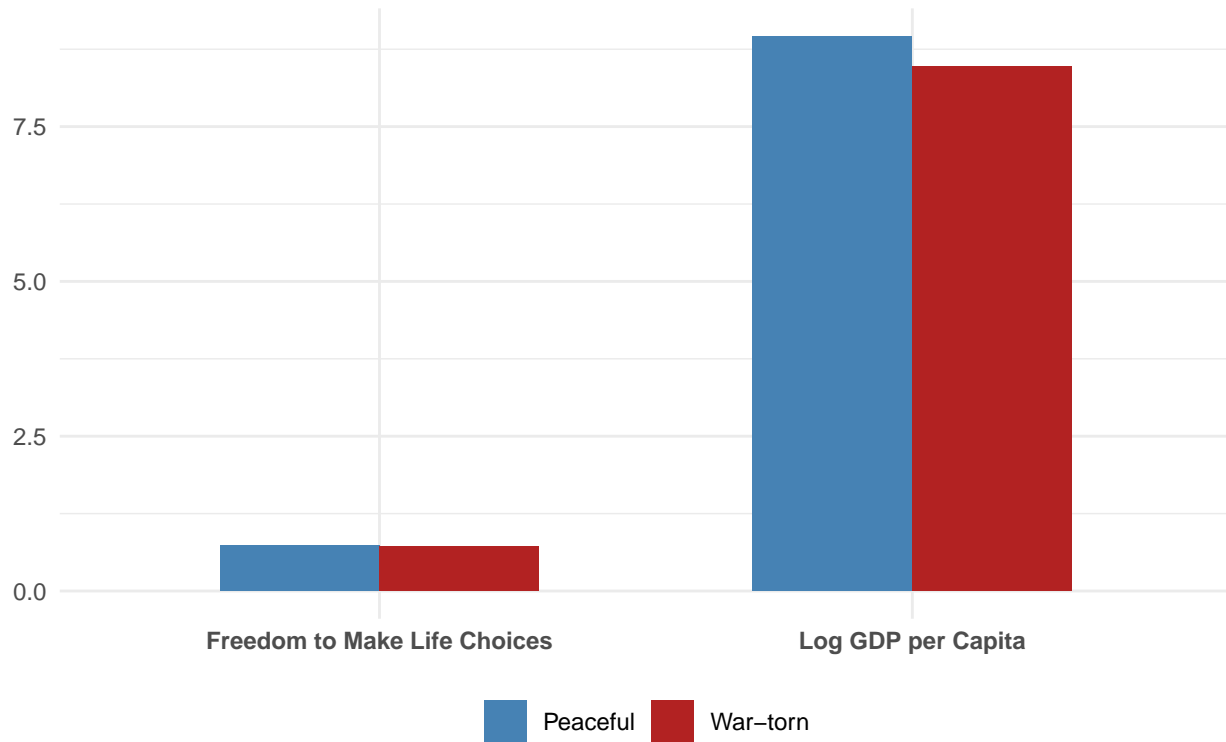


To better understand if our model is doing a good job at predicting happiness scores during wartime, its important to plot our residuals against the fitted values. Each point on this graph shows how far off the model's prediction was for a country in a given year. Here we placed the horizontal axis to show the predicted happiness scores that we got, and the vertical axis to show the residual, or the difference between what the model predicted and the actual score. A good model shows variability and no notable trend which is reflected well in the graph. The red curve also adds to help show if there's a pattern in the errors. Since our line is staying close to zero, we can assume our model is predicting consistently.



Now assume we are interested in the errors of our model. Specifically let's look at the differences between the predicted and actual happiness scores. To do this, we plot our model through a Q-Q plot which follows a normal distribution, a basic assumption behind statistical modeling. Looking deeper into Q-Q plots, each point represents one of the model's errors. The red diagonal line shows the ideal case, if the points closely follow this line, it means that the model's errors are normal, which is what we are looking for. Given our specific dataset, we can see that our points tend to stick to the red line. This is what we want to see and because of this we can be sure that our model's assumptions are reasonable. In terms of what this actually means, we are confident that our claims about what affects happiness during wartime are based on a reliable model.

Average Log GDP and Freedom (latest year)



Red = 16 conflict countries – Blue = remaining peaceful world

To complement the long-term trends and regression model, we could also look at a more recent case. Here, we grouped countries into two categories: those affected by civil conflict (“War-torn”) and those who are not (“Peaceful”). Keep in mind that since this data was gathered in 2023, these trends may not be aligned with current day situations. We compare the average values of two key variables: GDP per capita (log-transformed) and freedom to make life choices, both shown to be highly influential in our earlier model. The results show that on average, peaceful countries report higher GDP per capita and greater freedom, which aligns with expectations. as economic security and personal care are harder to maintain during conflict. It is important to note however that some countries like that of Mexico and UAE report relatively high values even during these difficult times. This further supports the argument that not all civil wars affect countries the same. There are internal conditions at play here that greatly buffer life satisfaction.

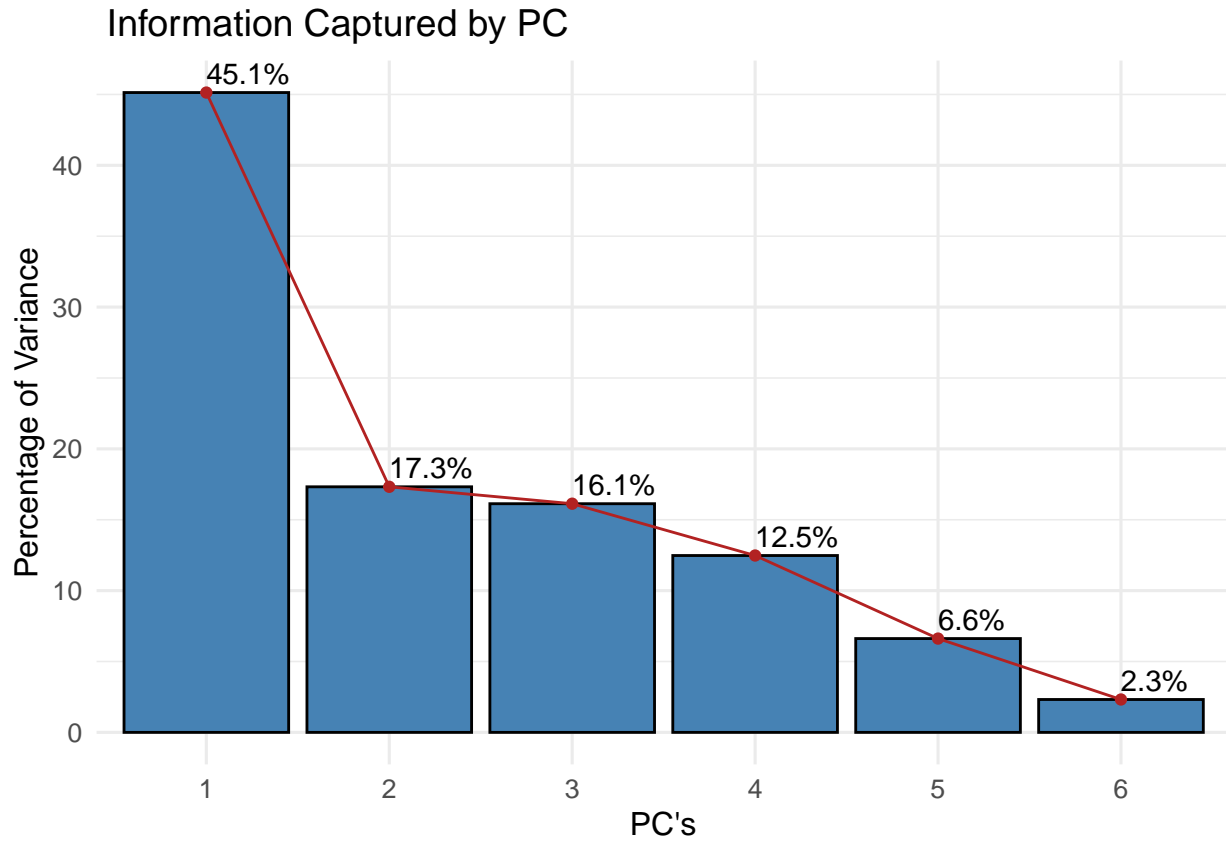


Table 2: Loadings on the first three Principal Components

	PC1	PC2	PC3
gdp	0.49	-0.16	-0.46
social_support	0.42	-0.36	0.03
life_expectancy	0.52	0.14	-0.39
freedom	0.41	0.09	0.52
generosity	-0.29	0.29	-0.59
corruption	-0.24	-0.86	-0.14

To further explore patterns in national wellbeing during wartime, we can also apply a Principle Component Analysis (PCA) using key predictors of happiness: GDP per capita, social support, life expectancy, freedom, generosity, and perceived corruption. The use of a PCA here helps us uncover underlying structures that may explain certain variation in happiness across countries. The plot displays the distribution of total variance by each principle competent (PC). The first PC's share a significant amount of variance suggesting that much of the original variables can be captured in just one dimension. The table that follows show each loading of each variable on the first three PC's. PC1 reflects a general wellbeing index with strong positive loadings on GDP, social support, life expectancy, and freedom. From this we can draw that countries scoring high on this axis tend to have better development and institutional strength even during wartime. PC2 compares generosity and freedom against GDP and life expectancy, potentially capturing cultural or societal resilience in lower-income contexts. PC3 is greatly influenced by corruption and generosity which reflects the variability in trust and society. These results allow us to compare countries and see what kinds of traits are most important under conflict.



## *Conclusion*

Initially, this report set out to explore a complex and relative question: Do all countries suffer from civil war the same way, and what are the greatest factors in maintaining or reducing happiness during such conflicts? Using data from the 2023 World Happiness Report, we were able to investigate how life satisfaction varies across countries experiencing internal conflict. Our visuals showed diverse patterns with some countries like Syria and Yemen displaying steep and prolonged declines in happiness while others like Mexico and Pakistan showed relatively stable life satisfaction despite ongoing conflicts. This comes to show that the impact of war on happiness is not relative to one another, each country processes their experiences differently. Modeling showed key predictors of life satisfaction during wartime. Variables such as economic strength (GDP per capita), social support, and freedom to make life choices are highlighted to be key drivers when it comes to determining happiness. The snapshot comparison between “war-torn” and peaceful countries in recent years reinforced these findings, showing that peaceful nations tend to have higher GDP and greater freedom, though exceptions reveal that the story is flawed. Principal Component Analysis further clarifies these factors into core underlying dimensions, emphasizing that happiness in conflict zones is shaped by a combination of interrelated social and economic influences. Overall, our results underscore that civil war affects populations differently depending on internal country dynamics. While war generally reduces happiness, countries that maintain stronger economic conditions and social freedoms appear better able to preserve well-being amidst conflict. By moving beyond death tolls and political narratives to focus on the lived human experience, this analysis contributes to a more compassionate understanding of war’s true cost—one that encompasses the vulnerabilities of everyday lives.

## Appendixes

(In order that they appeared)

```
happiness <- read.csv("WHR_2023/whr-2023.csv")
selected_countries <- c(
  "Syria", "Yemen", "Saudi Arabia", "United Arab Emirates",
  "Afghanistan", "Pakistan", "Mexico", "Ethiopia",
  "South Sudan", "Nigeria", "Niger", "Chad",
  "Cameroon", "Benin", "Libya", "Sri Lanka"
)

happiness_filtered <- happiness %>%
  filter(`Country.name` %in% selected_countries)

war_periods <- tribble(
  ~Country, ~Start, ~End,
  "Syria", 2011, 2023,
  "Yemen", 2014, 2023,
  "Saudi Arabia", 2014, 2023,
  "United Arab Emirates", 2014, 2023,
  "Iraq", 2005, 2023,
  "Afghanistan", 2005, 2021,
  "Pakistan", 2005, 2021,
  "Ethiopia", 2020, 2022,
  "South Sudan", 2013, 2018,
  "Nigeria", 2009, 2023,
  "Niger", 2009, 2023,
  "Chad", 2009, 2023,
  "Cameroon", 2009, 2023,
  "Benin", 2009, 2023,
  "Libya", 2011, 2011,
  "Libya", 2014, 2020,
  "Mexico", 2006, 2023,
  "Sri Lanka", 2006, 2009
)

happiness2 <- happiness_filtered %>%
  left_join(
    war_periods %>% select(Country, Start, End),
    by = c("Country.name" = "Country"),
    relationship = "many-to-many"
  ) %>%
  mutate(in_war = !is.na(Start) & year >= Start & year <= End)

plot_df <- happiness2 %>%
  group_by(`Country.name`, year) %>%
  summarise(`Life.Ladder` = first(`Life.Ladder`),
            in_war = any(in_war),
            .groups = "drop")

ggplot(plot_df, aes(year, `Life.Ladder`, group = `Country.name`)) +
  geom_line(colour = "grey65", linewidth = 0.9) + # baseline
  geom_line(data = filter(plot_df, in_war), # war overlay
            colour = "firebrick", linewidth = 1.1) +
```

```

geom_point(aes(colour = in_war), size = 1.8) +
scale_colour_manual(values = c(`FALSE` = "grey65",
                                `TRUE` = "firebrick"),
                    labels = c("Peace", "War"), name = NULL) +
facet_wrap(~ `Country.name`, ncol = 4, scales = "free_y") + # + free y-scales
scale_x_continuous(breaks = seq(2005, 2023, 4)) +
labs(title = "Life-Ladder Trends with War Years Highlighted",
      subtitle = "Grey = peace-time · Red = conflict years (2005-2023)",
      x = NULL, y = "Life Ladder score (0-10)") +
theme_minimal(base_size = 10) +
theme(legend.position = "bottom",
      panel.grid.minor = element_blank(),
      axis.text.x = element_text(angle = 45, hjust = 1))

```

```

model <- happiness2 %>%
  filter(in_war) %>%
  rename(
    life_ladder = Life.Ladder,
    log_gdp     = Log.GDP.per.capita,
    social      = Social.support,
    life_exp    = Healthy.life.expectancy.at.birth,
    freedom     = Freedom.to.make.life.choices,
    generosity  = Generosity,
    corruption  = Perceptions.of.corruption
  ) %>%
  drop_na(life_ladder, log_gdp, social, life_exp, freedom, generosity, corruption) %>%
  lm(
    life_ladder ~ log_gdp + social + life_exp + freedom + generosity + corruption,
    data = .
  )

tidy(model, conf.int=TRUE) |>
  mutate(across(where(is.numeric), \(x) round(x, 2))) |>
  kable(booktabs=TRUE, caption = "OLS Coefficients (unstandardised).")

```

```

aug <- augment(model)
p1 <- ggplot(aug, aes(.fitted, .resid)) +
  geom_hline(yintercept = 0, colour = "grey60") +
  geom_point(size = 1.8, alpha = 0.7) +
  geom_smooth(method = "loess", formula = y ~ x, colour = "firebrick", se = FALSE) +
  labs(title = "Residuals vs Fitted",
       x = "Fitted Life-Ladder score",
       y = "Residuals") +
  theme_minimal(base_size = 10)

```

p1

```

p2 <- ggplot(aug, aes(sample = .std.resid)) +
  stat_qq(size = 1.5, alpha = 0.7) +
  stat_qq_line(colour = "firebrick") +
  labs(title = "Normal Q-Q Plot (Std. Residuals)",
       x = "Theoretical Quantiles",
       y = "Standardised Residuals") +
  theme_minimal(base_size = 10)

```

p2

```
latest <- happiness_filtered %>%
  group_by(`Country.name`) %>%
  slice_max(order_by = year, n = 1, with_ties = FALSE) %>%
  ungroup()

latest <- latest %>%
  mutate(group = if_else(`Country.name` %in% c("Syria", "Yemen", "Saudi Arabia", "United Arab Emirates", "
    "War-torn", "Peaceful"))

avg_df <- latest %>%
  drop_na(`Log.GDP.per.capita`, `Freedom.to.make.life.choices`) %>%
  group_by(group) %>%
  summarise(avg_gdp = mean(`Log.GDP.per.capita`, na.rm = TRUE),
            avg_freedom = mean(`Freedom.to.make.life.choices`, na.rm = TRUE),
            .groups = "drop") %>%
pivot_longer(cols = c(avg_gdp, avg_freedom),
             names_to = "metric", values_to = "value") %>%
  mutate(metric = recode(metric,
                        avg_gdp = "Log GDP per Capita",
                        avg_freedom = "Freedom to Make Life Choices"))

ggplot(avg_df, aes(metric, value, fill = group)) +
  geom_col(position = "dodge", width = .6) +
  scale_fill_manual(values = c("War-torn" = "firebrick",
                                "Peaceful" = "steelblue"),
                    name = NULL) +
  labs(title = "Average Log GDP and Freedom (latest year)",
       x = NULL, y = NULL,
       caption = "Red = 16 conflict countries - Blue = remaining peaceful world") +
  theme_minimal(base_size = 11) +
  theme(legend.position = "bottom",
        axis.text.x = element_text(face = "bold"))

pca_input <- happiness2 %>%
  filter(in_war) %>%
  select(
    life_satisfaction = Life.Ladder,
    gdp = Log.GDP.per.capita,
    social_support = Social.support,
    life_expectancy = Healthy.life.expectancy.at.birth,
    freedom = Freedom.to.make.life.choices,
    generosity = Generosity,
    corruption = Perceptions.of.corruption
  ) %>%
  drop_na()

pca_mod <- prcomp(pca_input[-1], scale. = TRUE)

fviz_eig(
  pca_mod,
  addlabels = TRUE,
```

```

barfill  = "steelblue",
barcolor = "black",
linecolor = "firebrick"
) +
  labs(
    title = " Information Captured by PC",
    x = "PC's",
    y = "Percentage of Variance"
  ) +
  theme_minimal(base_size = 12)

kable(
  round(pca_mod$rotation[, 1:3], 2),
  booktabs = TRUE,
  caption = "Loadings on the first three Principal Components"
)

```

## *Sources*

Stack Overflow: Error code corrections

GeeksforGeeks: Understanding and utilizing different packages

ChatGPT: Formatting and error code corrections