PSTAT 100 Final Project - World Happiness Report Analysis

Meghana Dhruv (mdhruv), Kira Jackson (kirajackson)
Hsinpin Wen (hcwen), Candis Wu (cwu383)

2024-06-04

Table of contents

Abstract

When writing an academic-style paper, it is important to include an abstract. The abstract is a short, few-line summary of the paper, and should include major findings and conclusions.

Dataset Introduction

The data set our group chose to work with is the World Happiness Report 2023. This report provides insights into the happiness and well-being of people in different countries from around 2008 to 2023 by utilizing quantitative methods to understand how different life experiences influence people's happiness and quality of life. The happiness and well-being

factors that were in this data set include each country's life ladder, GDP, social support, healthy life expectancy, freedom of choice, generosity, perceptions of corruption, and positive and negative affects. Further information about these factors as well as the year and country name of the dataset are as follows:

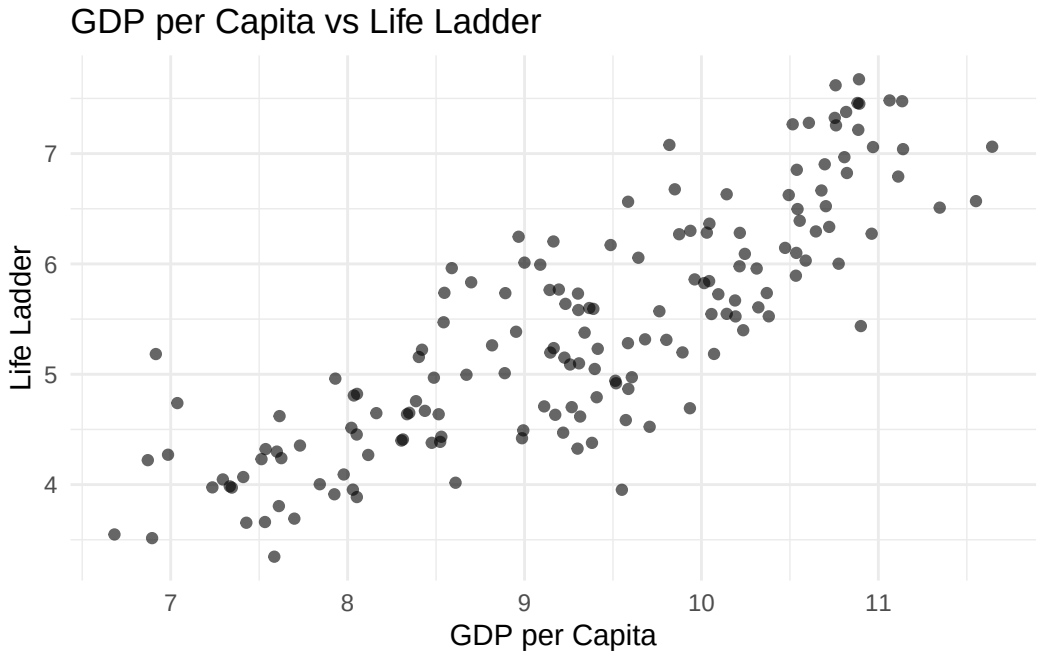
Variable Descriptions

`country_name`: Name of the country the data was collected in `year`: The year the data was collected in ranging from 2008-2023 `life_ladder`: Aggregate measure of how people within a country perceive and evaluate their own well-being and quality of life from 0 to 10 `gdp_per_capita`: A measure of economic prosperity and reflects the average income level of individuals in a country adjusted for population size `social_support`: The extent to which individuals feel they have someone to rely on in times of need, whether it be family, friends, or other forms of social support networks ranging from 0 to 1 `healthy_life_expectancy`: The average number of years an individual can expect to live in good health at birth which indicates the overall health and well-being of the country's population `freedom_of_choice`: A measure of individual autonomy and personal freedoms in making life choices `genorosity`: The portion of the national average charitable donations that cannot be explained by GDP per capita `perceptions_of_corruption`: The extent of how much individuals feel their government and businesses are corrupted ranging from 0 to 1 `positive_effect`: The average level of positive emotions in a country, based on responses to three questions in the Gallup World Poll: whether respondents smiled or laughed a lot yesterday, experienced enjoyment for much of the day yesterday, and learned or did something interesting yesterday. `negative_effect`: The negative affect variable represents the average level of negative emotions in a country, based on responses to three questions in the Gallup World Poll: whether respondents experienced worry, sadness, and anger for much of the day yesterday.

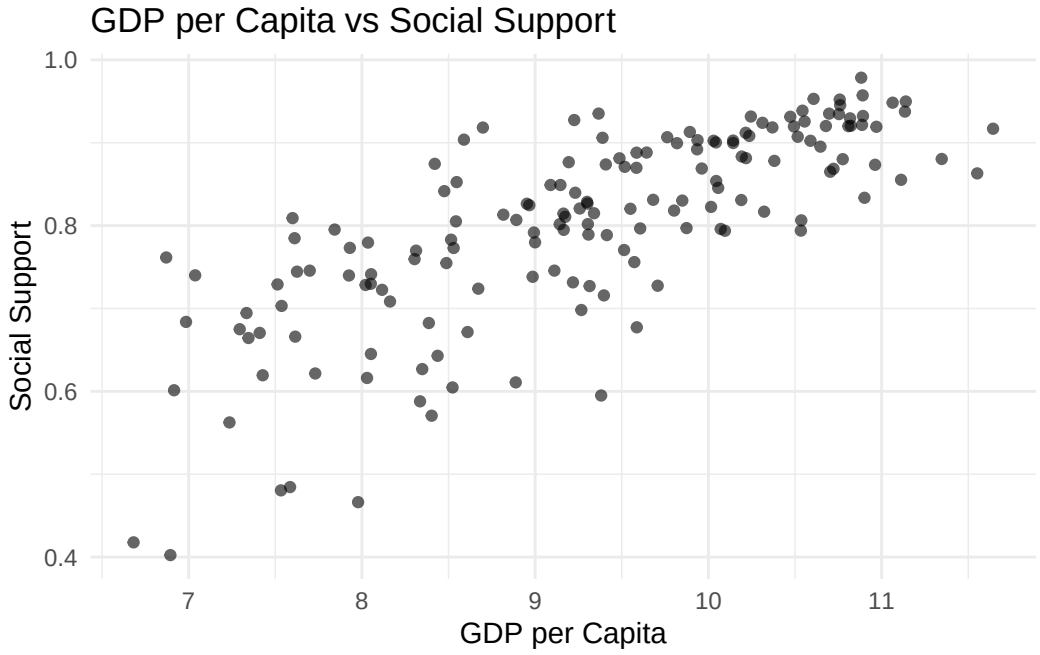
In this report, we will seek to answer the question: How does GDP per capita relate to the happiness and wellbeing indices reported in the World Happiness Report across different countries over time? **REPHRASE THIS QUESTION TO MATCH THE REPORT**

Country GDP vs Happiness Factors

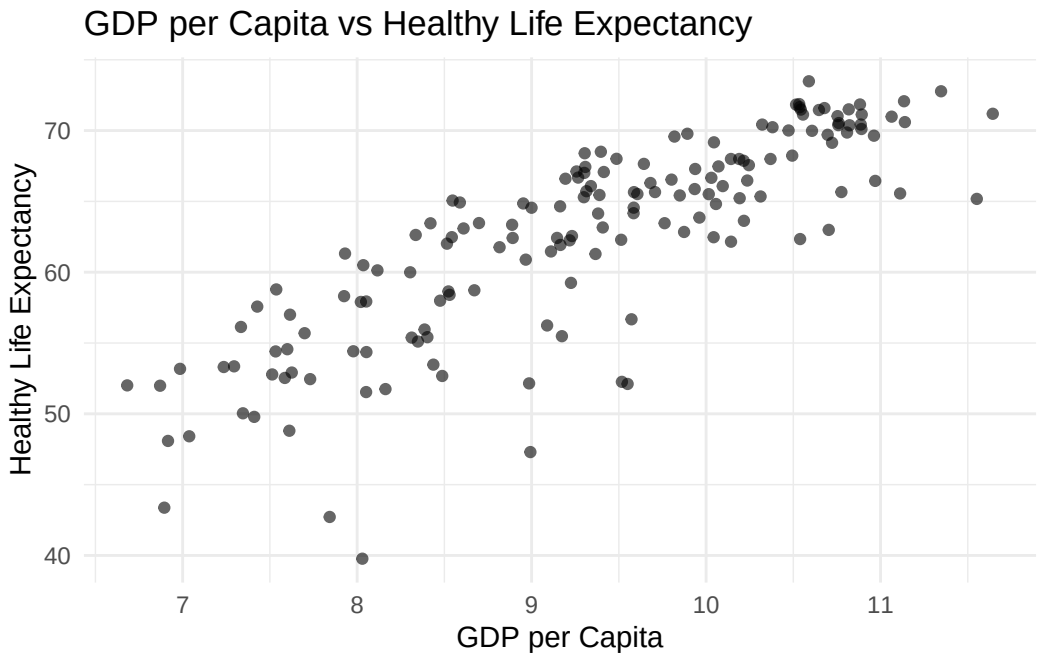
A country's Gross Domestic Product, or GDP, gives information about the size of its economy and how an economy is performing. It translates to a measure of a country's overall wealth, where a higher number indicates more money, and vice versa. Naturally, the economy of a country plays a large role in many aspects of how the country is doing. We will begin by exploring visualizations to examine the relationship between a country's GDP and a few of the factors contributing to the happiness and well-being of its people presented in the data set, namely **WHICH GRAPHS TO DO**. The relationship between GDP and all factors can be further explored in the Shiny app.



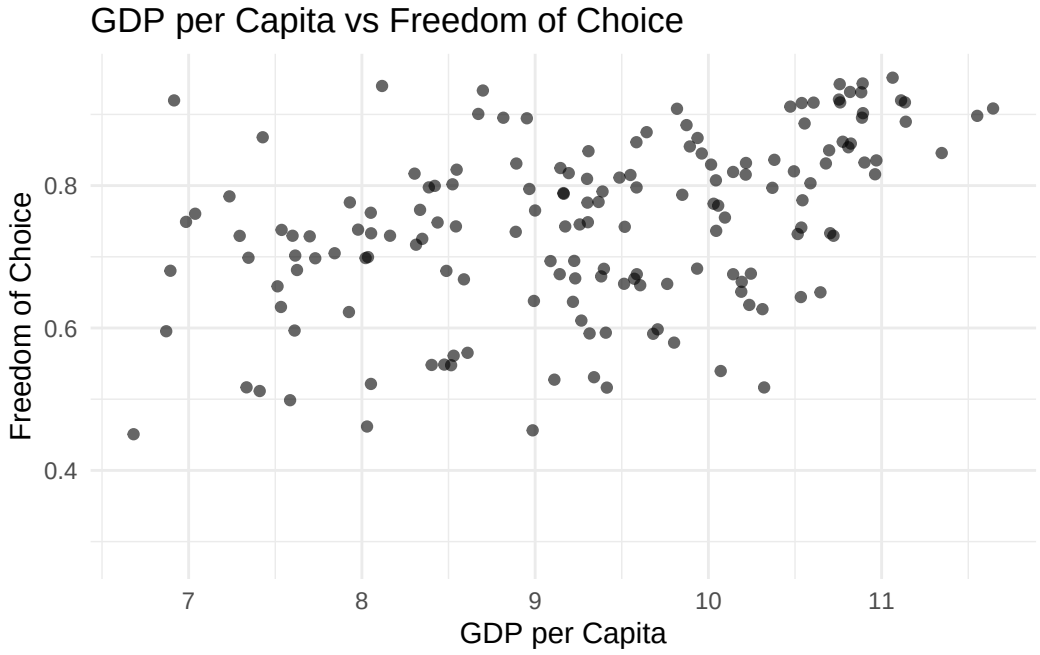
GRAPH ANALYSIS



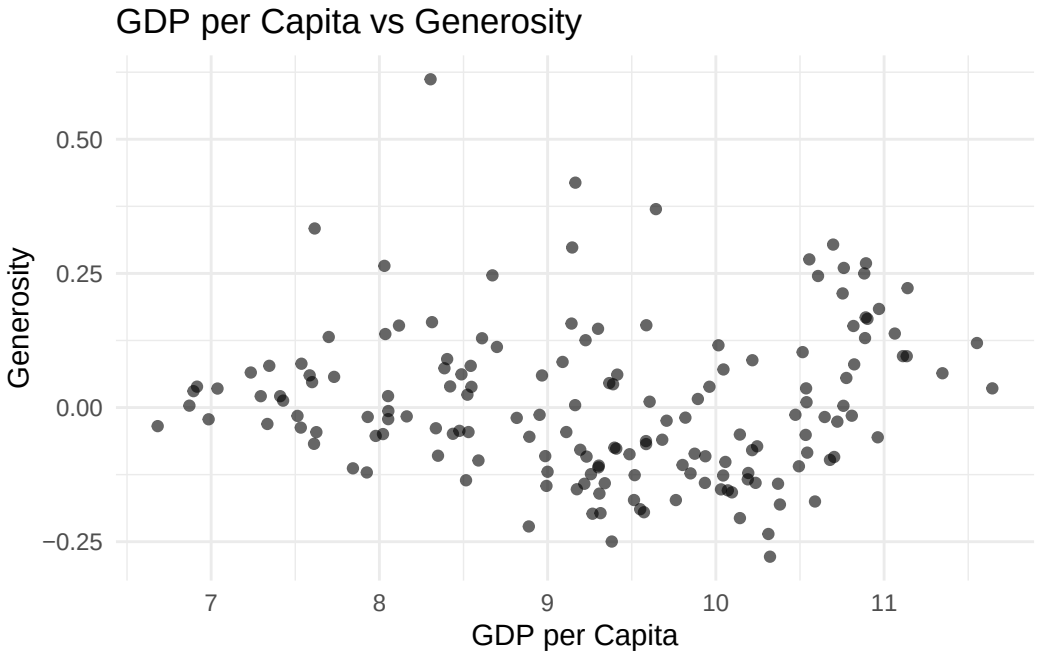
GRAPH ANALYSIS



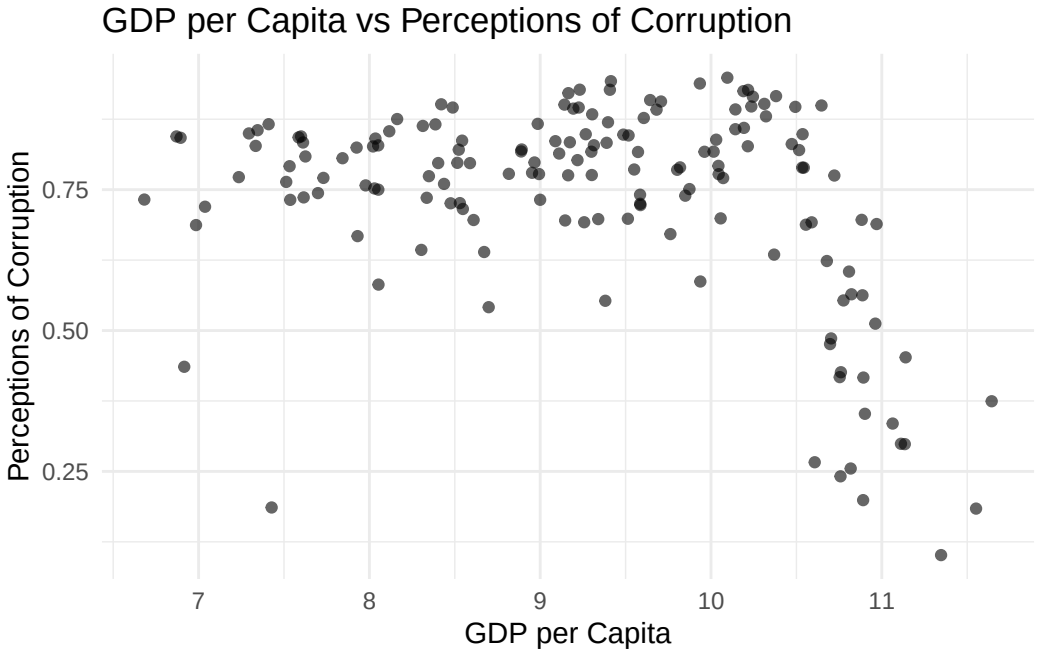
GRAPH ANALYSIS



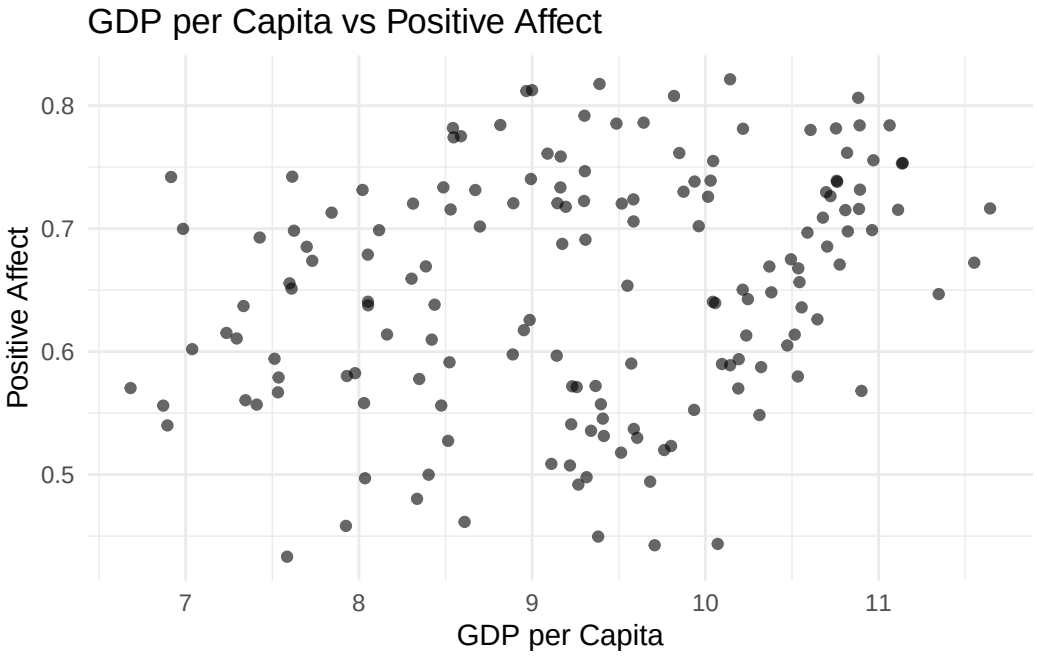
GRAPH ANALYSIS



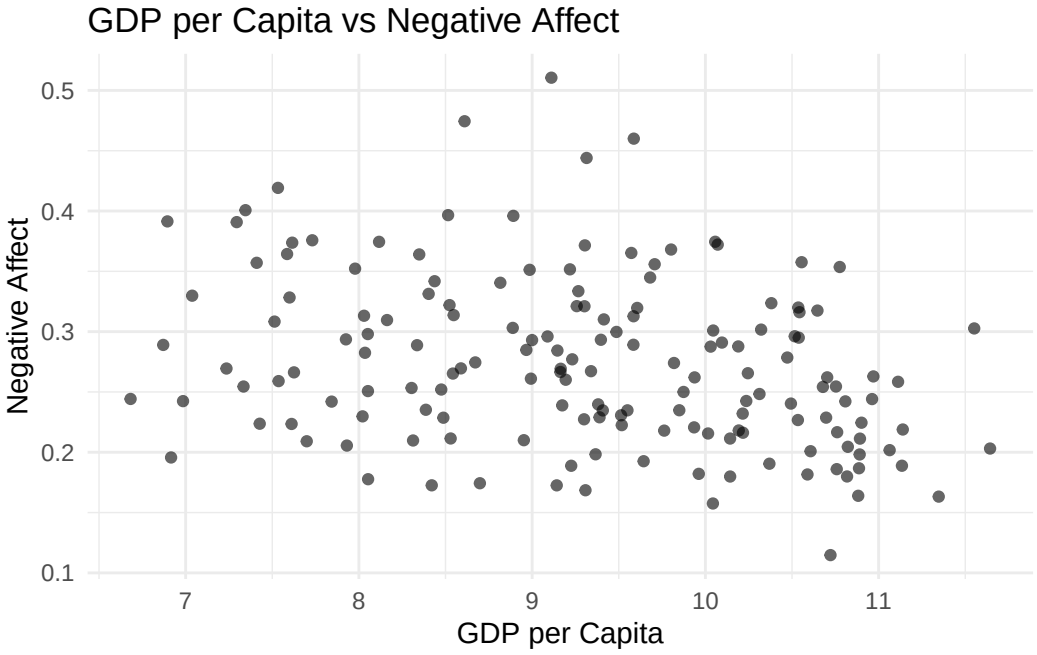
GRAPH ANALYSIS



GRAPH ANALYSIS



GRAPH ANALYSIS

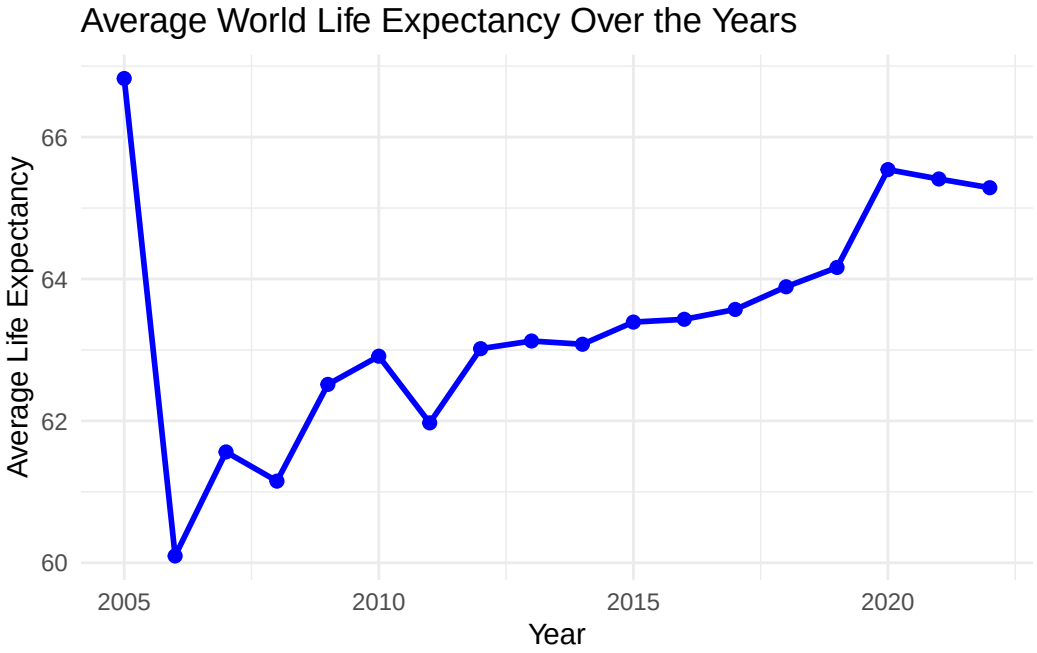


GRAPH ANALYSIS

Life Expectancy

ADD DISCRIPTION/INTRO: Intro to life expectancy, important because gauges health of overall country. Increase in life expectancy indicates better living conditions, improved lifestyle, better education, greater access to health

services. Countries with a higher GDP tend to have a longer life expectancy because blah blah blah. We will start by looking at the average life expectancy of the world over the years in the plot below.

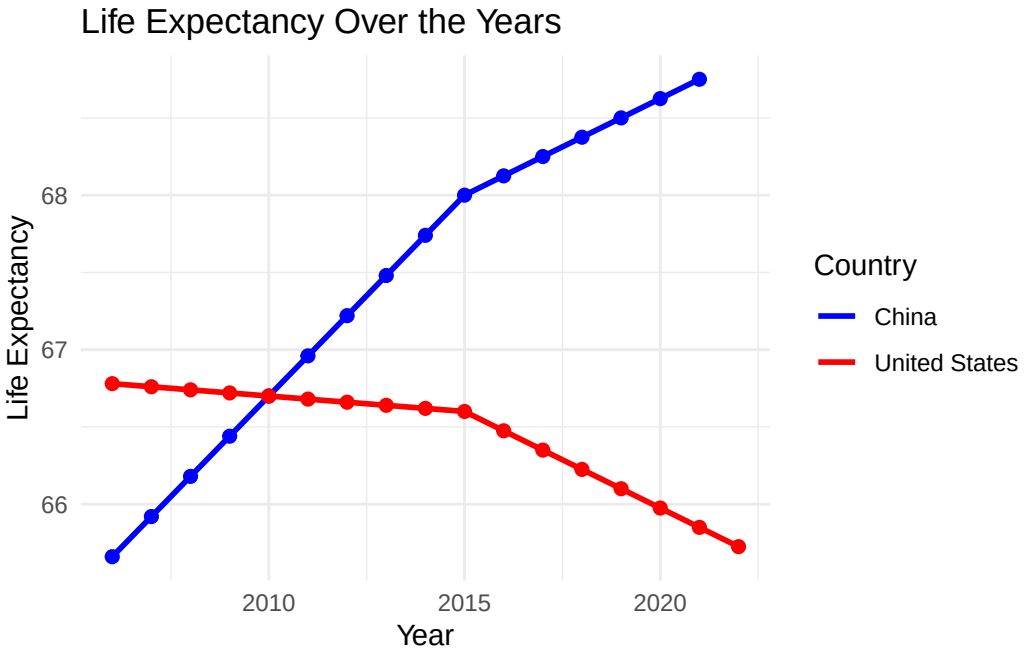


PLOT DESCRIPTION/ANALYSIS

China and United States Life Expectancy Comparison

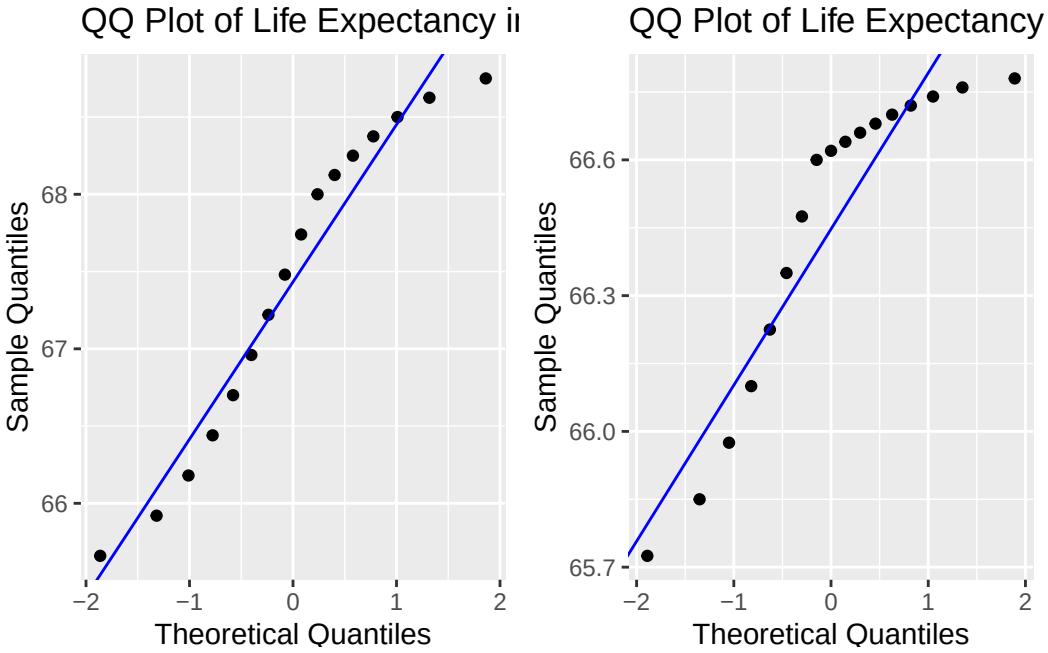
Next, we will narrow down our analysis to two economic global powerhouses: China and the United States. According to the US Department of State **ADD HYPERLINK https://www.state.gov/countries-areas/china/**, "China is currently the third-largest export market for U.S. goods (after Canada and Mexico), and the United States is China's largest export market." In examining global health metrics, comparing the life expectancy between the United States and China offers valuable insights into the health outcomes of two of the world's largest economies. Life expectancy, a key indicator of a population's overall health and well-being, reflects the average number of years a person can expect to live based on current mortality rates. By analyzing this metric, we can uncover the impact of various health behaviors, socio-economic conditions, and healthcare systems on the longevity of individuals in these countries. This comparison aims to highlight significant differences and underlying factors contributing to the health disparities between the United States and China.

```
Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
i Please use `linewidth` instead.
```



ADD MORE ANALYSIS/DESCRIPTION TO ABOVE PLOT Interestingly enough, we can see that while China's life expectancy has steadily increased over the years, the United States' life expectancy has been decreasing. Before we dive into this difference, we need to answer the following question: Is there a significant difference between the average life expectancy in the United States and China?

To answer this question, we will perform a t-test. **INFO/ASSUMPTIONS ABOUT T-TESTS?** We will begin by plotting QQ Plots of the life expectancies of China and the US to check for normality.



explain the QQ plot We can see from the QQ plots above that ___ are not quite normally

distributed. However, … large sample indicates we can still perform a t-test despite the non-normal distribution. **NEED TO FIND REFERENCE THAT THIS IS A LARGE SAMPLE - look in https://worldhappiness.report/ed/2023/**

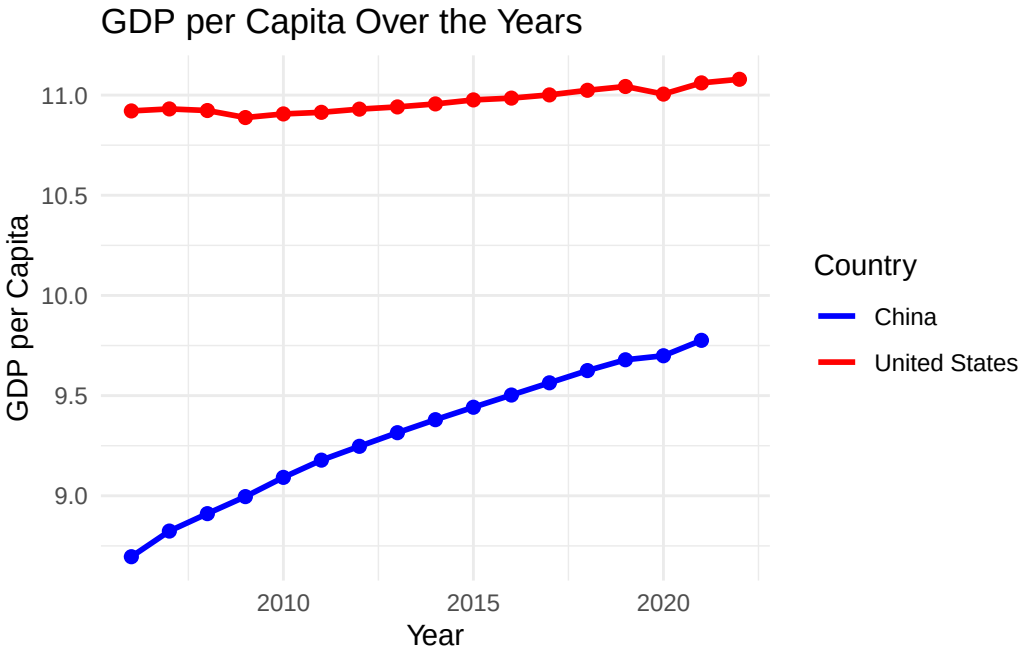
However, it is important to note that using t-tests and ANOVA tests with a large sample size can give meaningful results even if the data is not perfectly normal, but its not immune to errors so we will be interpreting the results with caution.

```
    Welch Two Sample t-test

data:  china_data$healthy_life_expectancy and us_data$healthy_life_expectancy
t = 3.68, df = 18.219, p-value = 0.001685
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.4234717 1.5480356
sample estimates:
mean of x mean of y
 67.43281  66.44706
```

INTERPRET THE T-TEST Upon conducting the above t-test, the resulting p-value was found to be 0.001685, indicating a statistically significant difference in healthy life expectancy between the two countries as it is less than the alpha value (?) 0.05. blah blah more analysis

From the earlier plot Section of GDP versus life expectancy, there is a positive linear trend between a country's GDP and healthy life expectancy. However, we are not seeing that in the case of the United States. Taking a look at the plot of China and the United States' GDP below, we can see that the US has had a consistently high GDP over the years. So, why has the United States' life expectancy been decreasing despite its high GDP?



Upon further exploration, we discovered that Americans have a lower life expectancy than other rich countries in the world. Below is a table showing the top 10 countries with the highest GDP's and their corresponding life expectancies:

Country Name	Average GDP per Capita	Average Healthy Life Expectancy
Luxembourg	11.64	71.19
Qatar	11.55	65.18
Singapore	11.35	72.77
Ireland	11.14	70.6
Switzerland	11.13	72.07
United Arab Emirates	11.11	65.56
Norway	11.06	70.98
United States	10.97	66.45
Kuwait	10.96	69.64
Netherlands	10.89	71.13

We can see from this table that apart from the Middle Eastern countries (Qatar, UAE, and Kuwait), the United States has the lowest average healthy life expectancy. We've excluded Middle Eastern countries from this observation because of the difference in sociopolitical, demographic, and economic environments compared to traditionally "Western" countries, since these countries' economies are driven by oil and other natural resources and an expansive migrant worker population. (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10378486/) **MAKE THIS A HYPERLINK**

Further research on the general health of Americans reveals that the United States has a higher rate of smoking, obesity, homicides, opioid overdoses, suicides, road accidents,

poverty, economic inequality, and lack of healthcare access compared to other rich countries. **ADD HYPERLINK FOR SOURCE: https://ourworldindata.org/us-life-expectancy-low** These factors contribute significantly to the nation's overall health outcomes and life expectancy. Importantly, many of these causes disproportionately affect young people which has a notable impact on the average life expectancy.

In conclusion, these findings reveal the relationship between health behaviors, societal challenges, and life expectancy in the United States. While the nation boasts considerable advancements in healthcare and medical technology, disparities in health outcomes persist due to many of the country's prevalent social, economic, and environmental issues. (add a more hopeful line?)

Tables

```
# Aggregate the data to get average GDP per capita per country
avg_gdp <- happiness_clean %>%
  group_by(country_name) %>%
  summarize(avg_gdp_per_capita = mean(gdp_per_capita, na.rm = TRUE)) %>%
  filter(!is.na(avg_gdp_per_capita)) %>%
  arrange(desc(avg_gdp_per_capita))

avg_gdp
```

```
# A tibble: 162 x 2
   country_name              avg_gdp_per_capita
   <chr>                                  <dbl>
 1 Luxembourg                              11.6
 2 Qatar                                   11.6
 3 Singapore                               11.3
 4 Ireland                                 11.1
 5 Switzerland                             11.1
 6 United Arab Emirates                    11.1
 7 Norway                                  11.1
 8 United States                           11.0
 9 Kuwait                                  11.0
10 Hong Kong S.A.R. of China               10.9
# i 152 more rows
```

```
# Compute average GDP per capita and average life ladder score for each country
# LIFE LADDER
avg_gdp_avg_life_ladder <- happiness_clean %>%
  group_by(country_name) %>%
```

```
    summarize(avg_gdp_per_capita = round(mean(gdp_per_capita, na.rm = TRUE), 4),
              avg_life_ladder = round(mean(life_ladder, na.rm = TRUE), 4)) %>%
    filter(!is.na(avg_gdp_per_capita), !is.na(avg_life_ladder)) %>%
    arrange(desc(avg_gdp_per_capita))

  # Print the updated table
  print(avg_gdp_avg_life_ladder)
```

```
# A tibble: 162 x 3
   country_name              avg_gdp_per_capita avg_life_ladder
   <chr>                                  <dbl>           <dbl>
 1 Luxembourg                              11.6            7.06
 2 Qatar                                   11.6            6.57
 3 Singapore                               11.3            6.51
 4 Ireland                                 11.1            7.04
 5 Switzerland                             11.1            7.47
 6 United Arab Emirates                    11.1            6.79
 7 Norway                                  11.1            7.48
 8 United States                           11.0            7.06
 9 Kuwait                                  11.0            6.27
10 Hong Kong S.A.R. of China               10.9            5.44
# i 152 more rows
```

```
  # SOCIAL SUPPORT
  avg_gdp_avg_social_support <- happiness_clean %>%
    group_by(country_name) %>%
    summarize(avg_gdp_per_capita = round(mean(gdp_per_capita, na.rm = TRUE), 4),
              avg_social_support = round(mean(social_support, na.rm = TRUE), 4)) %>%
    filter(!is.na(avg_gdp_per_capita), !is.na(avg_social_support)) %>%
    arrange(desc(avg_gdp_per_capita))

  print(avg_gdp_avg_social_support)
```

```
# A tibble: 161 x 3
   country_name              avg_gdp_per_capita avg_social_support
   <chr>                                  <dbl>              <dbl>
 1 Luxembourg                              11.6              0.917
 2 Qatar                                   11.6              0.863
 3 Singapore                               11.3              0.880
 4 Ireland                                 11.1              0.950
 5 Switzerland                             11.1              0.938
 6 United Arab Emirates                    11.1              0.855
```

```
 7 Norway                                  11.1              0.948
 8 United States                           11.0              0.919
 9 Kuwait                                  11.0              0.873
10 Hong Kong S.A.R. of China               10.9              0.834
# i 151 more rows
```

```r
# HEALTHY_LIFE_EXPECTANCY
avg_gdp_avg_healthy_life_expectancy <- happiness_clean %>%
  group_by(country_name) %>%
  summarize(avg_gdp_per_capita = round(mean(gdp_per_capita, na.rm = TRUE), 4),
            avg_healthy_life_expectancy = round(mean(healthy_life_expectancy, na.rm =
  filter(!is.na(avg_gdp_per_capita), !is.na(avg_healthy_life_expectancy)) %>%
  arrange(desc(avg_gdp_per_capita))

# Print the updated table
print(avg_gdp_avg_healthy_life_expectancy)
```

```
# A tibble: 160 x 3
   country_name         avg_gdp_per_capita avg_healthy_life_expectancy
   <chr>                             <dbl>                       <dbl>
 1 Luxembourg                         11.6                        71.2
 2 Qatar                              11.6                        65.2
 3 Singapore                          11.3                        72.8
 4 Ireland                            11.1                        70.6
 5 Switzerland                        11.1                        72.1
 6 United Arab Emirates               11.1                        65.6
 7 Norway                             11.1                        71.0
 8 United States                      11.0                        66.4
 9 Kuwait                             11.0                        69.6
10 Netherlands                        10.9                        71.1
# i 150 more rows
```

```r
# FREEDOM_OF_CHOICE
avg_gdp_avg_freedom_of_choice <- happiness_clean %>%
  group_by(country_name) %>%
  summarize(avg_gdp_per_capita = round(mean(gdp_per_capita, na.rm = TRUE), 4),
            avg_freedom_of_choice = round(mean(freedom_of_choice, na.rm = TRUE), 4)) %
  filter(!is.na(avg_gdp_per_capita), !is.na(avg_freedom_of_choice)) %>%
  arrange(desc(avg_gdp_per_capita))

# Print the updated table
print(avg_gdp_avg_freedom_of_choice)
```

```
# A tibble: 162 x 3
   country_name              avg_gdp_per_capita avg_freedom_of_choice
   <chr>                                  <dbl>                 <dbl>
 1 Luxembourg                              11.6                 0.908
 2 Qatar                                   11.6                 0.898
 3 Singapore                               11.3                 0.846
 4 Ireland                                 11.1                 0.890
 5 Switzerland                             11.1                 0.917
 6 United Arab Emirates                    11.1                 0.920
 7 Norway                                  11.1                 0.952
 8 United States                           11.0                 0.835
 9 Kuwait                                  11.0                 0.816
10 Hong Kong S.A.R. of China               10.9                 0.832
# i 152 more rows
```

```r
# GENEROSITY
avg_gdp_avg_generosity <- happiness_clean %>%
  group_by(country_name) %>%
  summarize(avg_gdp_per_capita = round(mean(gdp_per_capita, na.rm = TRUE), 4),
            avg_generosity = round(mean(generosity, na.rm = TRUE), 4)) %>%
  filter(!is.na(avg_gdp_per_capita), !is.na(avg_generosity)) %>%
  arrange(desc(avg_gdp_per_capita))

# Print the updated table
print(avg_gdp_avg_generosity)
```

```
# A tibble: 162 x 3
   country_name              avg_gdp_per_capita avg_generosity
   <chr>                                  <dbl>          <dbl>
 1 Luxembourg                              11.6         0.0358
 2 Qatar                                   11.6         0.120
 3 Singapore                               11.3         0.0639
 4 Ireland                                 11.1         0.223
 5 Switzerland                             11.1         0.0958
 6 United Arab Emirates                    11.1         0.0958
 7 Norway                                  11.1         0.138
 8 United States                           11.0         0.184
 9 Kuwait                                  11.0        -0.0554
10 Hong Kong S.A.R. of China               10.9         0.165
# i 152 more rows
```

```r
# perceptions_of_corruption
avg_gdp_avg_perceptions_of_corruption <- happiness_clean %>%
  group_by(country_name) %>%
  summarize(avg_gdp_per_capita = round(mean(gdp_per_capita, na.rm = TRUE), 4),
            avg_perceptions_of_corruption = round(mean(perceptions_of_corruption, na.r
  filter(!is.na(avg_gdp_per_capita), !is.na(avg_perceptions_of_corruption)) %>%
  arrange(desc(avg_gdp_per_capita))

# Print the updated table
print(avg_gdp_avg_perceptions_of_corruption)
```

```
# A tibble: 158 x 3
   country_name              avg_gdp_per_capita avg_perceptions_of_corruption
   <chr>                                  <dbl>                         <dbl>
 1 Luxembourg                              11.6                         0.374
 2 Qatar                                   11.6                         0.184
 3 Singapore                               11.3                         0.102
 4 Ireland                                 11.1                         0.452
 5 Switzerland                             11.1                         0.298
 6 United Arab Emirates                    11.1                         0.299
 7 Norway                                  11.1                         0.335
 8 United States                           11.0                         0.689
 9 Kuwait                                  11.0                         0.512
10 Hong Kong S.A.R. of China               10.9                         0.352
# i 148 more rows
```