

**7COM1079-0901-2025**  
Team Research and Development Project

# **Final Report**

**Is there a correlation between the reported happiness and economic performance, and does this correlation differ between more productive economies and less productive ones?**

**Group ID: Group A 14**  
Dataset Number: DS282

**Prepared by:**

Name	Student ID	Email
Gideon Ganah	24088169	gg24abh@herts.ac.uk
Kevin Kogo	24061635	kk24acz@herts.ac.uk
Daniel Terna Tule	24076755	dt24aav@herts.ac.uk
Edosa Izevbizua	23097619	ei24aas@herts.ac.uk
Kelvin Amaadi Baani-Anoff	24083020	kb24aby@herts.ac.uk

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Problem statement and research motivation . . . . .	2
1.2	The data set . . . . .	2
1.3	Research question . . . . .	2
1.4	Null hypothesis and alternative hypothesis (H <sub>0</sub> /H <sub>1</sub> ) . . . . .	2
<b>2</b>	<b>Background Research</b>	<b>2</b>
2.1	Research papers relevant to the topic . . . . .	2
2.2	Why the research question is of interest . . . . .	3
<b>3</b>	<b>Visualization</b>	<b>3</b>
3.1	Primary plot for the Research Question (output of R script) . . . . .	3
3.2	Additional visualizations (optional) . . . . .	7
3.3	Useful information for data understanding . . . . .	7
<b>4</b>	<b>Analysis</b>	<b>7</b>
4.1	Statistical test used and output . . . . .	7
4.2	Decision on the null hypothesis . . . . .	7
<b>5</b>	<b>Evaluation – Group’s Experience at 7COM1079</b>	<b>7</b>
5.1	What went well . . . . .	7
5.2	Points for improvement . . . . .	8
5.3	Group’s time management . . . . .	8
5.4	Project’s overall judgement . . . . .	8
5.5	Comment on GitHub log output . . . . .	8
5.6	Commit Messages . . . . .	8
5.6.1	Commit Messages [Description for visualizations], 3.2 and 3.3, polished.] .	8
5.6.2	Commit Messages [I have added the statistical analyses to the word document. that is the chapter 4] . . . . .	8
5.6.3	Commit message [Finalization of the documentation] . . . . .	9
<b>6</b>	<b>Conclusions</b>	<b>9</b>
6.1	Results explained . . . . .	9
6.2	Interpretation of the results . . . . .	9
6.3	Reasons / implications for future work and limitations . . . . .	9
<b>7</b>	<b>Reference List</b>	<b>10</b>
<b>8</b>	<b>Appendices</b>	<b>11</b>
	<b>Appendix A: R Code</b>	<b>11</b>
	<b>Appendix B: GitHub Log Output</b>	<b>14</b>

# 1 Introduction

## 1.1 Problem statement and research motivation

Economic growth has been considered an indicator of societal progress for a very long time, our research question aims to question the assumption that higher economic performance automatically leads to an increased sense of well-being or happiness. There exists emerging literature on subjective well-being called the Easterlin Paradox, which suggests that overtime as a countries level of economic development increases, the overall happiness of its members does not necessarily rise (Stevenson and Wolfers, 2008). The motivation behind this research is to clearly understand if economic development is linked to happiness and if this link is universal across countries with highly productive economies and those with lower productive economies.

## 1.2 The data set

Our study utilizes the 2015 World Happiness Report, which is sourced from the Gallup World Poll, The dataset comprises of 12 columns which define global well-being. The primary dependent variable is the Happiness Score (ratio data). Potential independent variables include economy (GDP per Capita), Family, Health, Freedom, Trust, and Generosity (all ratio type data) . Country and region provide geographical context (nominal data), Dystopia residual (ratio) represents a hypothetical worst-case baseline used to compare countries fairly.

## 1.3 Research question

Is there a correlation between the reported happiness and economic performance, and is this correlation stronger in more productive economies. The aim is to determine if stronger economic performance is linked to higher happiness scores and whether this pattern holds true in both top-performing and low-performing economies

## 1.4 Null hypothesis and alternative hypothesis (H<sub>0</sub>/H<sub>1</sub>)

The null hypothesis is that there is no correlation between the happiness score and the economic performance of a country, suggesting that an increase in a country income would not correlate with an increase in happiness. The alternative hypothesis is that there is a correlation between the happiness score and the economic performance of a country, and this correlation is stronger in high-performing economies, suggesting that an increase in income leads to an increase in happiness.

# 2 Background Research

## 2.1 Research papers relevant to the topic

This study relies on data available in the World Happiness Report 2015. As identified and highlighted in a previous study by Helliwell et al. in 2015, GDP per capita remains one of the six most influential factors used to determine happiness score's globally. Their analysis established that these factors collectively explain three-quarters of the variation in national scores, providing a strong baseline for the link between happiness score and economic output.

However, this association is controversial. Clark, Frijters, and Shields (2008) examine the Easterlin paradox. They argue that richer individuals are happier, but over time the aggregate national happiness score does not necessarily increase with economic growth. They also discuss diminishing marginal utility. This indicates that income is most important for meeting basic needs in poorer countries.

Conversley, Stevenson & Wolfers (2008) challenge this paradox using data from the Gallup World Poll. They observed that the effect of economic productivity does not stall and there is a strong linear log relationship between higher economic output and a greater sense of well-being. This aligns with our goal of examining whether a similar linear association also exists between countries classified as the richest and poorest countries in our dataset of World Happiness Report from 2015.

## 2.2 Why the research question is of interest

The standard economic theory (Clark et al . 2008) suggests that there is a steeper correlation in developing nations due to basic needs. Our research question investigates the possibility that there is a relationship between happiness score and economic productivity and this relationship is universal, the research question is intriguing because several researchers have come up with widely different conclusions as seen in section 2.1 above, Happiness in general is an important factor for everyday human life therefore we find the conflict with the current research to be an important quandary that we aim investigate .

## 3 Visualization

### 3.1 Primary plot for the Research Question (output of R script)

This section presents the figures generated from the R analysis. All figures are referenced automatically using their labels. A scatterplot is used as the main visualization because the research question concerns the correlation between two continuous variables: happiness score and economic productivity. A histogram is also included to show how happiness scores are distributed within each subgroup, helping assess the shape, skewness, and suitability for correlation analysis.

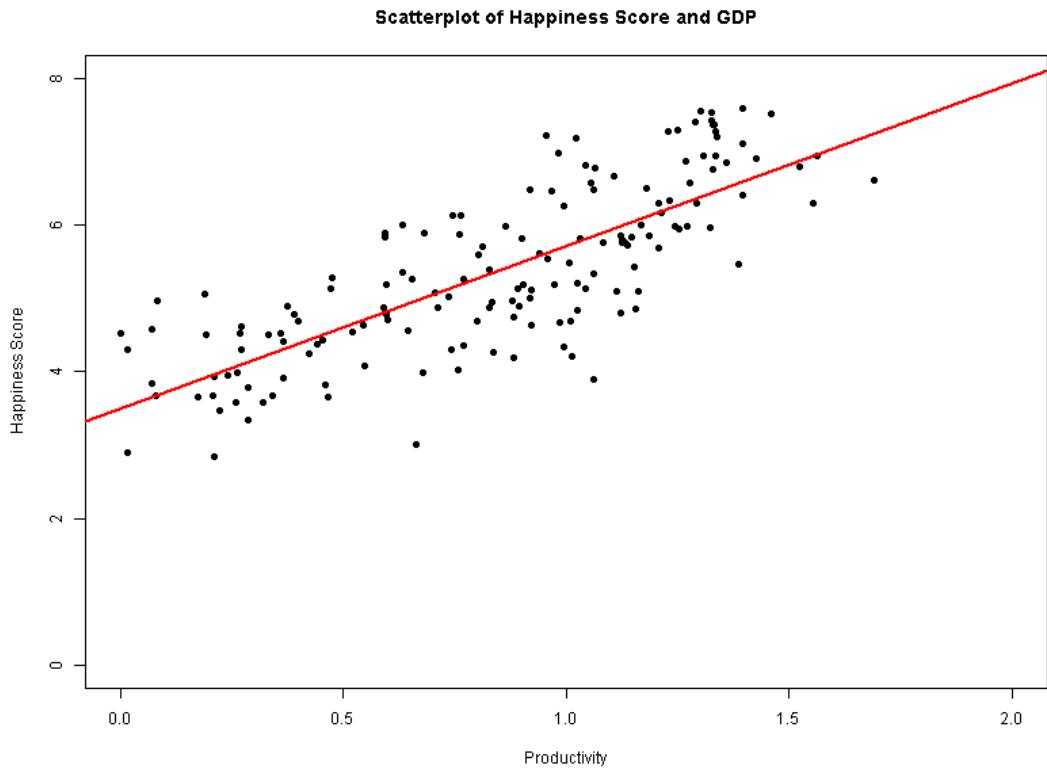


Figure 1: Scatterplot showing the correlation between productivity and happiness.

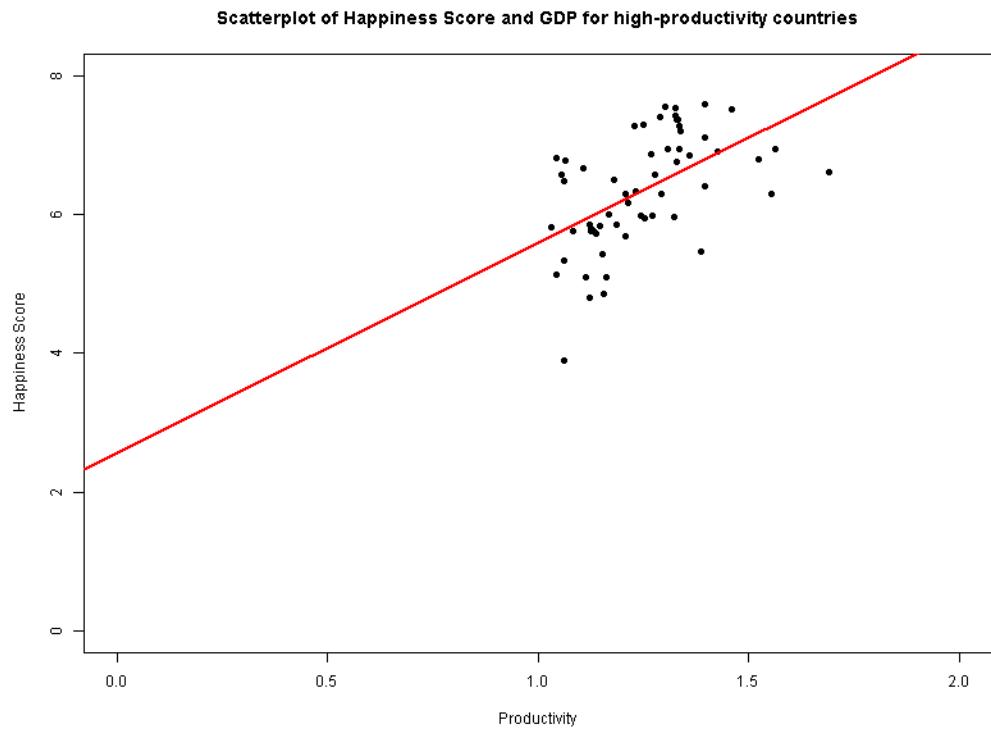


Figure 2: Scatterplot showing the correlation between productivity and happiness score among high-productivity nations.

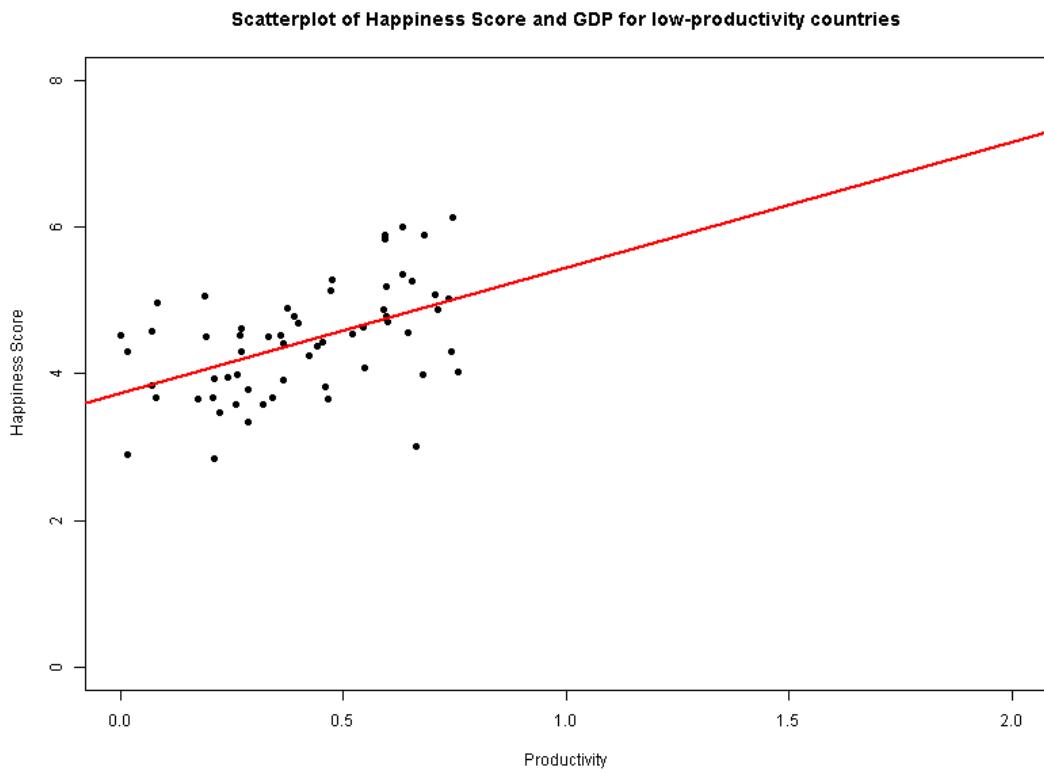


Figure 3: Scatterplot showing the correlation between productivity and happiness score among low-productivity nations.

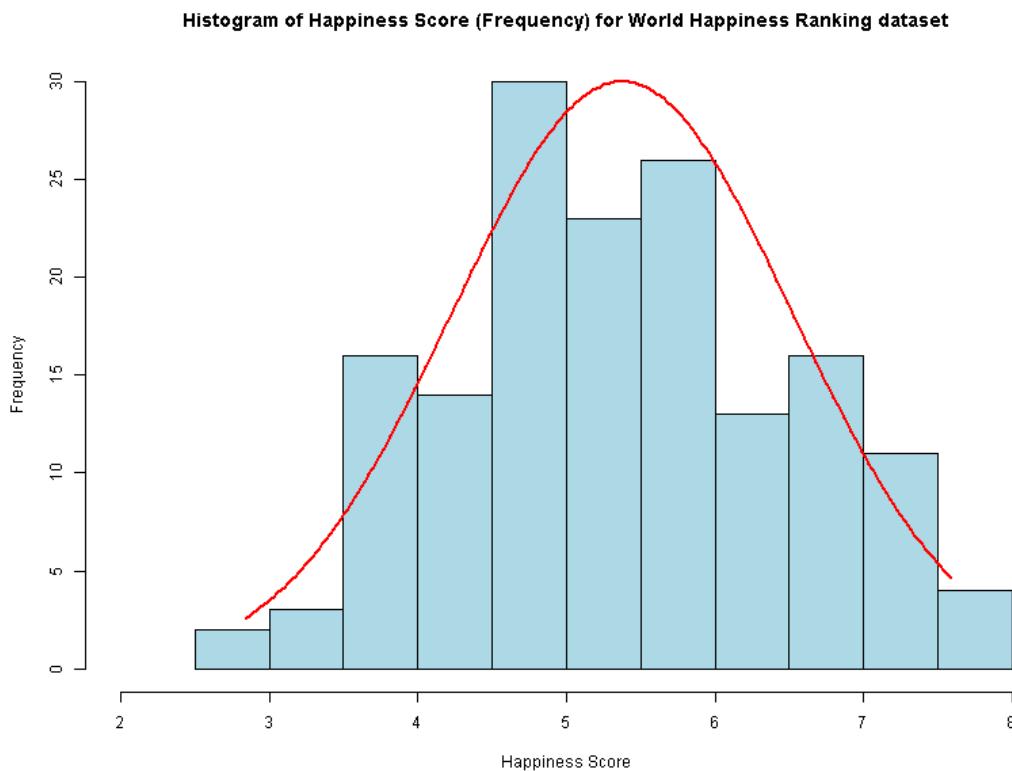


Figure 4: Histogram of happiness scores for the full dataset.

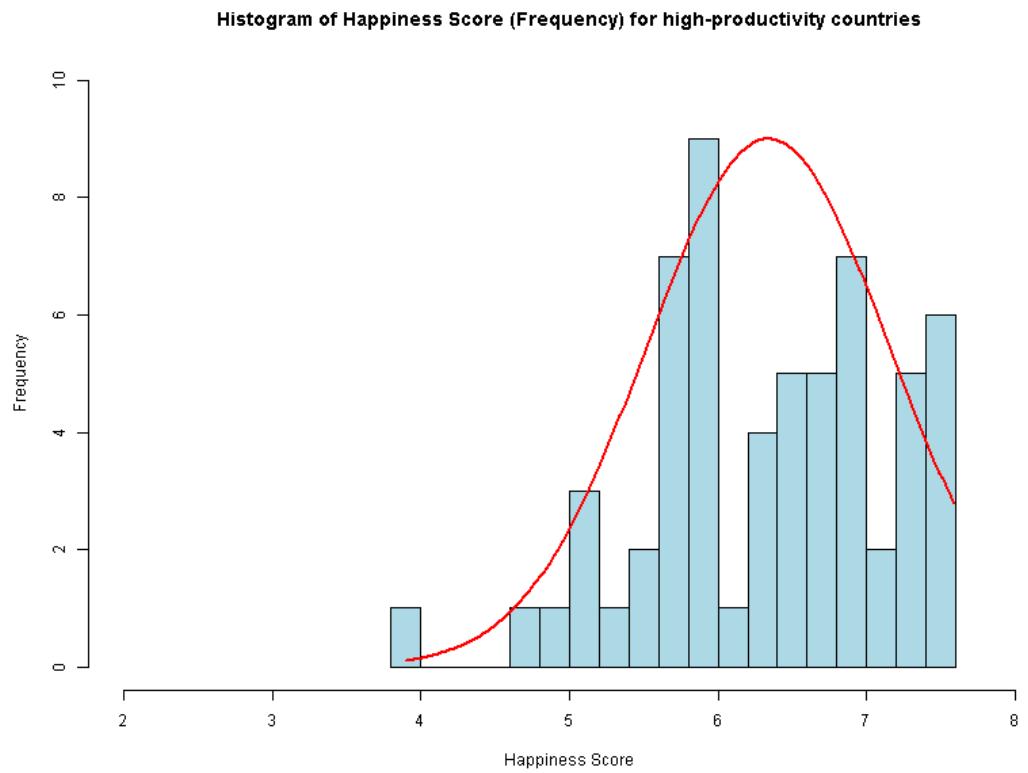


Figure 5: Histogram of happiness scores for high-productivity nations.

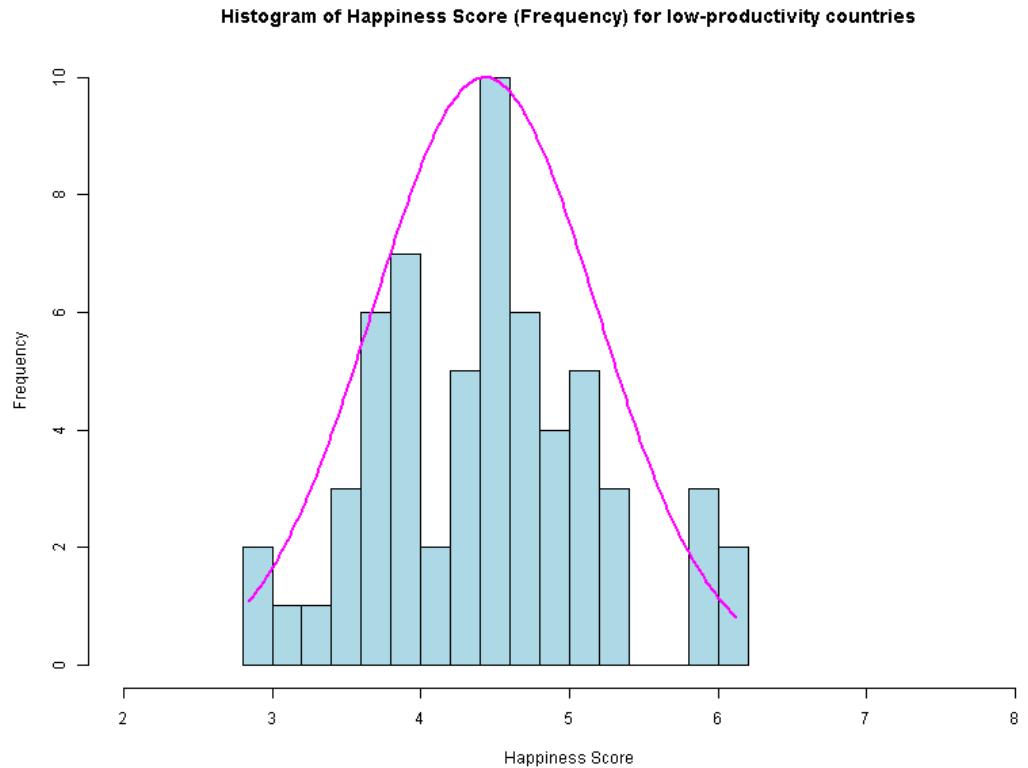


Figure 6: Histogram of happiness scores for low-productivity nations.

### **3.2 Additional visualizations (optional)**

- Figure 1 shows a positive Trend Line Shows positive correlation between the independent and dependent variable for all countries
- Figure 2 shows a very steep positive trend line and tighter clustering shows a stronger positive correlation in high productivity countries
- Figure 3 shows a less steep positive trendline with less clustering shows a weaker correlation in low productivity countries
- Figure 4 shows a symmetric and near normal histogram without major outliers for all countries

5 shows a left skewed and not normal distribution for high productivity countries

- Figure 6 shows a not normal distribution with some outliers for the low productivity countries

### **3.3 Useful information for data understanding**

Key observations include a positive linear trend in all scatterplots, with a noticeably steeper slope in high-productivity countries, suggesting a stronger correlation between productivity and happiness. The histograms show that high-productivity nations cluster toward higher happiness, while low-productivity countries exhibit more variability, indicating that additional factors may influence their happiness scores.

## **4 Analysis**

### **4.1 Statistical test used and output**

Our research sought to ascertain whether there exists a correlation between the reported happiness score and the economic performance of countries. We found that the entire happiness score has a nearly normal distribution but not a perfectly normal one, so we used spearman test for the general correlation between happiness score and economic productivity. The spearman test is a non parametric statistical test that measures the correlation between two variables.

### **4.2 Decision on the null hypothesis**

We have rejected the null hypothesis which states that there is no correlation between the happiness score and the economic productivity of a country. Our p-value from the spearman test on the whole dataset was 2.2e-16 . This p-value is way below the threshold of 0.05 and hence from this we concluded that the positive correlation between happiness and economic productivity is statically and could not have occurred by chance

## **5 Evaluation – Group's Experience at 7COM1079**

### **5.1 What went well**

The group established a professional working relationship early, we aimed to avoid random dataset allocation. We learnt that historically led to lower marks. Online tools like (WhatsApp, Google Meet, Microsoft Teams) were adopted to coordinate tasks when in-person meetings were not feasible. Early role sharing and dataset selection aided workflow. Two tutorial

demo presentation sessions (RQ refinement and Testing methods) produced actionable feedback that improved our work. These practices align with good team formation and role allocation, supporting faster progress and clearer deliverables (Tuckman, 1965).

## 5.2 Points for improvement

Scheduling more face-to-face meetings to build stronger rapport, to better match tasks to individual strengths would have helped. Online meetings sustained progress, closer interpersonal familiarity would likely have allowed more flexible role swapping and peer coaching. Intra-group conflict and unclear role allocation can reduce performance (Jehn, 1995). Role sharing was choice based, with more rapport it could have been to individuals' strengths to maximize output or challenged individuals with parts they were uncomfortable with so they can improve.

## 5.3 Group's time management

Time constraints and competing coursework limited the frequency of meetings, so we relied on short online sessions and asynchronous messages on WhatsApp. For future projects, a shared calendar with fixed weekly check-ins and short agendas could help mitigate this and reduce waiting time and accelerate decision points (Salas et al., 2015).

## 5.4 Project's overall judgement

Overall, the project was successful. Dr Imran Khan (Tutor) recounted working with friends professionally and its adverse effects. Hence, we not being prior friends but meeting and working professionally helped. The group selected a dataset through considered discussion, iteratively refined our research question using tutor feedback, and produced reproducible analysis workflows.

## 5.5 Comment on GitHub log output

Project development and version control history, as seen in the Git log, highlights several key collaborative milestones. The foundational project template and structure were established in the initialization of the repository file commit (cb70625). Commits were done with almost no conflicts. (Full Git log in Appendix B).

## 5.6 Commit Messages

### 5.6.1 Commit Messages [Description for visualizations], 3.2 and 3.3, polished.]

Daniel Generated visualization from R code and images , this allowed us to get a better sense of the general direction of our analysis , helping us see clearly the spread of the data especially when we applied stratification

### 5.6.2 Commit Messages [I have added the statistical analyses to the word document. that is the chapter 4]

- Gideon completed the statistical analysis , we together as a group came to the conclusion that we will use the spearman test over the pearson because our dataset is not normal , it does appear normal but some bins are not proportionate hence our decision to use spearman , the results we got led us to the conclusion that the correlation with happiness score and economic productivity is stronger in higher productive economies

### **5.6.3 Commit message [Finalization of the documentation]**

Gideon , Daniel and Kevin came together to proof-read the work and correct any errors , we looked at every step from the beginning to the end and recalculated the spearman test just to be sure , we then pushed the final version before submission of the assignment . This served as the final version of our work which we are satisfied with.

## **6 Conclusions**

### **6.1 Results explained**

We initially considered the whole dataset, we found a correlation coefficient of 0.7957028 and a p-value of 2.2e-16 using the Spearman test. We then stratified the dataset into two groups , the first group contained the top 60 countries ranked by economic performance and the other group contained the bottom 60. We found a positive correlation coefficient of 0.6090581 and a p-value of 4.629e-07 for the top 60 countries and for the bottom 60 we found a positive correlation of 0.49197 and a p-value of 8.107e-05.

### **6.2 Interpretation of the results**

Our results show a strong correlation between happiness score and economic performance, the p-values show us that the result is statistically significant, this tells us that economic performance of a country contributes significantly to the populations overall sense of happiness. After stratifying the data into two groups,the top 60 countries based on GDP have a slightly higher correlation than the bottom 60, which informs us that with increased economic productivity the correlation with happiness is stronger.

### **6.3 Reasons / implications for future work and limitations**

Although the findings provide meaningful insights into the relationship between happiness score and productivity across countries, the limitation is that the data may not capture the complexity of these constructs fully because happiness can also be influenced by other factors like culture. Future work could prob into specific components of happiness score like, health, freedom of choice and social support.

## 7 Reference List

1. Clark, A. E., Frijters, P., & Shields, M. A. (2008). Relative income, happiness, and utility: An explanation for the Easterlin paradox and other puzzles. *Journal of Economic Literature*, 46(1), 95-144.
2. Helliwell, J. F., Huang, H.,& Wang, S. (2015). The Geography of World Happiness. In J. F. Helliwell, R. Layard, & J. Sachs (Eds.), *World Happiness Report 2015* (pp. 12-41). New York: Sustainable Development Solutions Network.
3. Jehn, K.A., 1995. A multimethod examination of the benefits and detriments of intragroup conflict. *Administrative Science Quarterly*, 40(2), pp.256–282.
4. Salas, E., Shuffler, M.L., Thayer, A.L., Bedwell, W.L. and Lazzara, E.H., 2015.Understanding and improving teamwork in organizations: a scientifically based practical guide. *Human Resource Management*, 54(4), pp.599–622. Available at: <https://doi.org/10.1002/hrm.21628> [Accessed 07/12/2025].
5. Tuckman, B. (1965). Developmental sequence in small groups. *Psychological Bulletin*, 63(6), pp.384–399. Available at: <https://doi.org/10.1037/h0022100> [Accessed 07/12/2025].
6. Stevenson, B., & Wolfers, J. (2008). Economic growth and subjective well-being: Re-assessing the Easterlin paradox. *Brookings Papers on Economic Activity*, 2008(1), 1-87.

## 8 Appendices

### Appendix A: R code used for analysis and visualisation

```
1 # Load Dataset
2 df <- read.csv( 2015_original_dataset.csv )
3
4 # # Ensure output directories exist
5 dir.create( assets/HistogramPlots , recursive = TRUE, showWarnings = FALSE)
6 dir.create( assets/ScatterPlots , recursive = TRUE, showWarnings = FALSE)
7
8 # Data Wrangling - Renaming Columns for clarity
9 colnames(df)[3] <- HappRank      # Happiness Rank
10 colnames(df)[4] <- HappScore     # Happiness Score
11 colnames(df)[5] <- StdError      # Standard Error
12 colnames(df)[6] <- GDPperCapita  # Economy (GDP Per Capita)
13 colnames(df)[8] <- LifeExpectancy # Health (Life Expectancy)
14 colnames(df)[10] <- GovtTrust    # Trust (Government Corruption)
15 colnames(df)[12] <- Dystopia     # Dystopia Residual
16
17
18 top60 <- head(df[order(df$GDPperCapita, decreasing = TRUE), ], 60)
19 bottom60 <- tail(df[order(df$GDPperCapita, decreasing = TRUE), ], 60)
20
21 h <- df$`HappScore`
22 h1 <- top60$`HappScore`
23 h2 <- bottom60$`HappScore`
24
25 png( assets/HistogramPlots/totalHistogram.png , width = 800, height = 600)
26
27 # Histogram using frequency (default)
28 h_hist <- hist(h,
29   breaks = 15,
30   col = lightblue ,
31   main = "Histogram of Happiness Score (Frequency) for World Happiness
32   Ranking dataset",
32   xlab = "Happiness Score",
33   ylab = "Frequency",
34   ylim = c(0, 30.0),
35   xlim = c(2.0, 8.0)
36 )
37
38 # Create x-values for the curve(top60)
39 x <- seq(min(h, na.rm = TRUE), max(h, na.rm = TRUE), length = 200)
40
41 # Normal density values
42 curve_y <- dnorm(x, mean = mean(h, na.rm = TRUE), sd = sd(h, na.rm = TRUE))
43
44 # Scale the curve to the histogram frequency scale
45 curve_y_scaled <- curve_y * max(h_hist$counts) / max(curve_y)
46
47 # Add the curve
48 lines(x, curve_y_scaled, col = red , lwd = 2)
49
50 dev.off()
51
52 png( assets/ScatterPlots/totalScatter.png , width = 800, height = 600)
53
54 plot(df$GDPperCapita, df$HappScore,
55       xlab = "Productivity",
56       ylab = "Happiness Score",
57       main = "Scatterplot of Happiness Score and GDP",
```

```

58     pch = 19,
59     xlim = c(0, 2),
60     ylim = c(0, 8)
61 )
62
63 abline(lm(HappScore ~ GDPperCapita, data = df), col = red, lwd = 2)
64
65 dev.off()
66
67 png( assets/HistogramPlots/top60Histogram.png , width = 800, height = 600)
68
69 # Histogram using frequency (default)
70 h1_hist <- hist(h1,
71   breaks = 15,
72   col = lightblue ,
73   main = Histogram of Happiness Score (Frequency) for high-productivity
74   countries ,
75   xlab = Happiness Score ,
76   ylab = Frequency ,
77   ylim = c(0, 10.0),
78   xlim = c(2.0, 8.0)
79 )
80
81 # Create x-values for the curve(top60)
82 x1 <- seq(min(h1, na.rm = TRUE), max(h1, na.rm = TRUE), length = 200)
83
84 # Normal density values
85 curve_y1 <- dnorm(x1, mean = mean(h1, na.rm = TRUE), sd = sd(h1, na.rm = TRUE))
86
87 # Scale the curve to the histogram frequency scale
88 curve_y1_scaled <- curve_y1 * max(h1_hist$counts) / max(curve_y1)
89
90 # Add the curve
91 lines(x1, curve_y1_scaled, col = red, lwd = 2)
92
93 dev.off()
94
95 png( assets/ScatterPlots/top60Scatter.png , width = 800, height = 600)
96
97 plot(top60$GDPperCapita, top60$HappScore,
98   xlab = Productivity ,
99   ylab = Happiness Score ,
100  main = Scatterplot of Happiness Score and GDP for high-productivity
101  countries ,
102  pch = 19,
103  xlim = c(0, 2),
104  ylim = c(0, 8)
105 )
106
107 abline(lm(HappScore ~ GDPperCapita, data = top60), col = red, lwd = 2)
108
109 dev.off()
110
111 png( assets/HistogramPlots/bottom60Histogram.png , width = 800, height = 600)
112
113 h2_hist <- hist(h2,
114   breaks = 15,
115   col = lightblue ,
116   main = Histogram of Happiness Score (Frequency) for low-productivity
117   countries ,
118   xlab = Happiness Score ,
119   ylab = Frequency ,
120   ylim = c(0, 10.0),
121 )

```

```

118     xlim = c(2.0, 8.0)
119 )
120
121 # Create x-values for the curve(bottom60)
122 x2 <- seq(min(h2, na.rm = TRUE), max(h2, na.rm = TRUE), length = 200)
123
124 # Normal density values
125 curve_y2 <- dnorm(x2, mean = mean(h2, na.rm = TRUE), sd = sd(h2, na.rm = TRUE))
126
127 # Scale the curve to the histogram frequency scale
128 curve_y2_scaled <- curve_y2 * max(h2_hist$counts) / max(curve_y2)
129
130 # Add the curve
131 lines(x2, curve_y2_scaled, col = magenta, lwd = 2)
132
133 dev.off()
134
135 png( assets/ScatterPlots/bottom60Scatter.png , width = 800, height = 600)
136
137 plot(bottom60$GDPperCapita, bottom60$HappScore,
138       xlab = Productivity ,
139       ylab = Happiness Score ,
140       main = Scatterplot of Happiness Score and GDP for low-productivity
141             countries ,
142       pch = 19,
143       xlim = c(0, 2),
144       ylim = c(0, 8)
145 )
146 abline(lm(HappScore ~ GDPperCapita, data = bottom60), col = red , lwd = 2)
147
148 dev.off()
149
150 # Correlation Analysis
151 # Correlation for the entire dataset(in-terms of GDP)
152 cor_total <- cor.test(df$HappScore, df$GDPperCapita, method = spearman )
153 print( --- Correlation: Total Dataset (Pearson) --- )
154 print(cor_total)
155
156 # Correlation for the top 60 countries
157 cor_top60 <- cor.test(top60$HappScore, top60$GDPperCapita, method = spearman )
158 print( --- Correlation: Top 60 Countries (Spearman) --- )
159 print(cor_top60)
160
161 # Correlation for the bottom 60 countries(in-terms of GDP)
162 cor_bottom60 <- cor.test(bottom60$HappScore, bottom60$GDPperCapita, method =
163                           spearman )
163 print( --- Correlation: Bottom 60 Countries (Spearman) --- )
164 print(cor_bottom60)

```

Listing 1: R code used for analysis and visualisation

## Appendix B: GitHub log output

The following pages contain the complete GitHub commit history for the project, including author names, timestamps, and commit messages.

```
20fe022 - Gideon Ganah, 2025-12-10 : Merge branch 'main' of
→ https://github.com/KimtaiKogoKevin/TeamResearchPractical_ds282 Merging to
→ repo
80ba251 - Gideon Ganah, 2025-12-10 : Added some content
253a909 - KimtaiKogoKevin, 2025-12-10 : Did some revision of the documentation
→ with Gideon and Daniel and uploaded the document to github
1e259c0 - Pirategh, 2025-12-10 : Kelvin - I added an Edits file to be adopted
→ for the clean report: added Edits.docx file to repo
6aee712 - Pirategh, 2025-12-10 : yourname: updated chapter X/Y in report
faac86b - Gideon Ganah, 2025-12-08 : I prepared the innitial draft of the
→ report and have it cleaned up and well edited. kindly check for
→ corrections and improvements
34e4145 - Gideon Ganah, 2025-12-08 : I did minor corrections to the report
32e7d23 - Daniel Tule, 2025-12-08 : Minor correction in 3
49b7c1b - Daniel Tule, 2025-12-08 : Minor correction in 3.3
d0eb6af - Edosa Izevbizua, 2025-12-08 : Merge pull request #7 from
→ KimtaiKogoKevin/Samuel
f7907ab - Samuel Izevbizua, 2025-12-08 : bg research
47a3c38 - Samuel Izevbizua, 2025-12-08 : ignore all temporary files
8acc4ca - Daniel Tule, 2025-12-07 : Corrected error in visualisations. Scatter
→ plots not showing.
f42241f - Gideon Ganah, 2025-12-07 : I did some minor corrections in the
→ report
ffaa50e - Edosa Izevbizua, 2025-12-07 : Merge pull request #6 from
→ KimtaiKogoKevin/Samuel
a4ebaee - Samuel Izevbizua, 2025-12-07 : cleanup: comment out correlation
→ analysis print statements for clarity
2ba534b - Samuel Izevbizua, 2025-12-07 : chore: cleanup repository and
→ consolidate analysis scripts
28fa99a - Gideon Ganah, 2025-12-06 : I added the conclusion of our analysis to
→ the report
2fdf8ed - Daniel Tule, 2025-12-06 : Cleaning up report. Visualisations moved
→ to after table of contents for clarity.
6e9bb69 - Daniel Tule, 2025-12-06 : Description for visualisations, 3.2 and
→ 3.3, polished.
9905ee6 - Daniel Tule, 2025-12-05 : Visualisation infor and description added
→ to doc
a2bac10 - Pirategh, 2025-12-04 : Kelvin I made edits to the word count:
→ updated Chapter 5 in report template
b667716 - KimtaiKogoKevin, 2025-12-04 : Completed Introduction section of the
→ documentation
08b17e8 - Daniel Tule, 2025-12-04 : Made all visualisations have uniform
→ graduation.
ca6c9fd - Daniel Tule, 2025-12-04 : Made all visualisations have uniform
→ graduation.
590f40f - KimtaiKogoKevin, 2025-12-04 : Added the introduction plus research
→ question need to finish up on the null and alterantive hypothesis today
```

109d858 - Daniel Tule, 2025-12-04 : Realised error while documenting visualisations. Decreasing is supposed to be true, this means our visuals have changed.

e02aa4d - KimtaiKogoKevin, 2025-12-03 : Added the first paragraph of the introduction

b468331 - KimtaiKogoKevin, 2025-12-03 : Added the first paragraph of the introduction

aff9bcc - Gideon Ganah, 2025-12-03 : Minor corrections to Report

937840f - Gideon Ganah, 2025-12-03 : I did some minor corrections to the report file

8ed7102 - Gideon Ganah, 2025-12-03 : I have added the statistical analyses to the word document. that is the chapter 4

ba80604 - KimtaiKogoKevin, 2025-12-03 : Forgot to exclude DSStore added to gitignore

73ed0de - KimtaiKogoKevin, 2025-12-03 : Activated the .gitignore to remove unnecessary R files going forward that are not , such as RHistory and the Rproject user folder

1db3bcd - KimtaiKogoKevin, 2025-12-03 : added the problem area and our motivation behind the study to the repo , the document is still very messy at the moment but we aim to fine tune it soon

0dbecc19 - KimtaiKogoKevin, 2025-12-03 : Made some changes to the Introduction and a clarification with group members that this is the file we will edit for our main report and not the duplicate which I will proceed to delete

b003f81 - Gideon Ganah, 2025-12-03 : I resolved the merge conflict and added the resolved file

ad553e6 - Gideon Ganah, 2025-12-03 : I have done the statistical tests which are updated in the combined R script in the assets directory

2a29ce5 - Daniel Tule, 2025-12-03 : Corrected some minor labeling errors in our code and generated accurate visualisations.

0e88603 - Daniel Tule, 2025-12-03 : I noticed the update, but the x-axis is still going out of bounds, so I've corrected that.

03e2a98 - Daniel Tule, 2025-12-03 : Visualisations updated as per Monday's feedback.

fae0f05 - Gideon Ganah, 2025-12-03 : I edited the combined R script to add the code for generating the histogram for the entire dataset. I also exported the PNG file to the assets directory

6085a92 - Gideon Ganah, 2025-12-03 : I did some commenting of the renamed columns in the combined R file

8a54df6 - Gideon Ganah, 2025-12-03 : I did some spell-check in chapter 5

52f4fd7 - Gideon Ganah, 2025-12-03 : Merge branch 'main' of [https://github.com/KimtaiKogoKevin/TeamResearchPractical\\_ds282](https://github.com/KimtaiKogoKevin/TeamResearchPractical_ds282) "I am trying to pull the updated repo, I have to save my local changes first"

18b0507 - Gideon Ganah, 2025-12-03 : Saved local changes before pulling

0a58ae6 - Pirategh, 2025-12-02 : Added Chapter 5: Evaluation section with group experience analysis

7b0eb53 - Samuel Izembizua, 2025-12-02 : Add .gitignore file to exclude R project files

354f3bb - Samuel Izembizua, 2025-12-02 : Add .gitignore file to exclude R project files

a4f859d - Daniel Tule, 2025-12-01 : Update Rhistory file

3958721 - Daniel Tule, 2025-12-01 : I've done the visualisations for the top and bottom 60 categories.

690b6d5 - Gideon Ganah, 2025-11-28 : I have added the report file for working  
417757e - Gideon Ganah, 2025-11-27 : I have added The R Script for the Top 50  
→ / Bottom 50 GDP Correlation analysis  
f507665 - Daniel Tule, 2025-11-27 : I've renamed some columns with long names.  
ffd7158 - KimtaiKogoKevin, 2025-11-27 : change to working with git file  
→ regarding not pushing to main because the tutors will mainly evaluate the  
→ main branch , and after consultation they say there is no need for  
→ branching, additionally I have added a folder with our power point slides  
→ for the first presentation  
4839220 - KK\_Kimutai, 2025-11-27 : Merge pull request #4 from  
→ KimtaiKogoKevin/GideonGanah  
169e809 - Gideon Ganah, 2025-11-27 : I added the scatter plot  
bbdb68a - KK\_Kimutai, 2025-11-27 : Merge pull request #3 from  
→ KimtaiKogoKevin/KogoKevin  
e28ae00 - KimtaiKogoKevin, 2025-11-27 : Created scripts directory where I have  
→ saved two scripts , one for dividing the dataset into different regions  
→ and one that draws a histogram with a curve showing that the data is  
→ normally distributed  
d7a9b91 - KK\_Kimutai, 2025-11-25 : Merge pull request #2 from  
→ KimtaiKogoKevin/KogoKevin  
3250c62 - KimtaiKogoKevin, 2025-11-25 : a function to mmap the data into  
→ different regions  
2836f16 - KK\_Kimutai, 2025-11-14 : Merge pull request #1 from  
→ KimtaiKogoKevin/KogoKevin  
b2c99c6 - KimtaiKogoKevin, 2025-11-14 : Distribution of chosen variables for  
→ research question with correction  
84fec02 - KimtaiKogoKevin, 2025-11-14 : Distribution of chosen variables for  
→ research question  
bb87bce - KimtaiKogoKevin, 2025-11-13 : change to readme.md  
129c601 - KimtaiKogoKevin, 2025-11-13 : Instructions for team members to work  
→ on their own branches , DO NOT WORK ON MAIN BRANCH  
cb70625 - KimtaiKogoKevin, 2025-11-13 : Initialization of repo with Readme  
→ file