Exploring the Insights within the Global Happiness Index

Toh Shao Ern, Kuick Si Qi, Wong Khin Foong, Sebestian Fernandez, Yu Haotong (Group 12)

Keywords — Happiness Index, Typst, Quarto

I. Introduction

i. Objective

To critically assess and enhance a static visualization of global happiness levels, identifying its strengths and weaknesses, and proposing improvements for a more interactive and insightful presentation.

ii. Global Happiness Score

The Life Ladder score, also known as the World Happiness Score, is derived from the Gallup World Poll data. The score is typically calculated based on several key variables that reflect the well-being and quality of life in different countries. Here's a general outline of the formula and components involved in calculating the World Happiness Score:

Component	Description
Life Evaluations (Ladder Question)	Respondents rate their current life from 0 (worst) to 10 (best).
GDP per Capita	Natural logarithm of GDP per capita (PPP terms).
Social Support	Binary response to: "Do you have relatives or friends you can count on for help if needed?"
Healthy Life Expectancy	Life expectancy at birth.
Freedom to Make Life Choices	Binary response to: "Are you satisfied with your freedom to choose what you do with your life?"
Generosity	Donations to charity in the past month, adjusted for GDP per capita
Perceptions of Corruption	Perceived corruption in government and business.
Positive and Negative Affect	Average responses to questions about recent experiences of positive and negative emotions (e.g., laughter, enjoyment, sadness, and anger).

The World Happiness Score is calculated using a regression model that combines these variables to predict the happiness score for each country. The formula for the World Happiness Score can be represented as follows:

$$\begin{aligned}\text{Happiness Score} = \beta_0 &+ \beta_1 \times (\text{Log GDP per capita}) \\ &+ \beta_2 \times (\text{Social Support}) + \beta_3 \times (\text{Healthy Life Expectancy}) \\ &+ \beta_4 \times (\text{Freedom to Make Life Choices}) + \beta_5 \times (\text{Generosity}) \\ &+ \beta_6 \times (\text{Perceptions of Corruption}) + \beta_7 \times (\text{Positive Affect}) - \beta_8 \times (\text{Negative Affect})\end{aligned} \tag{1}$$

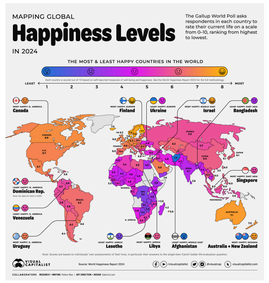
Where β_0 to β_8 are coefficients derived from regression analysis based on historical data.

iii. Visualcapitalist's Global Happiness Levels Report

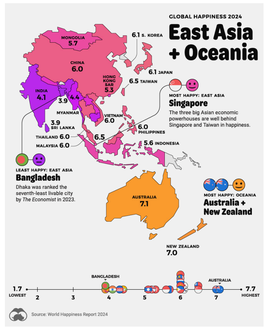
Happiness levels across the globe are a complex interplay of economic, social, psychological, and environmental factors that contribute to an individual's sense of well-being. To quantitatively assess and compare these levels internationally, the **World Happiness Report** (WHR) serves as a vital tool. Initiated by the **United Nations Sustainable Development Solutions Network**, the WHR annually ranks countries based on the self-reported happiness of their citizens.

The report utilizes data primarily sourced from the **Gallup World Poll** (GWP), a sub-dataset of the World Happiness Report, which asks respondents to evaluate their current lives on a scale from 0 to 10, known as the **Cantril Ladder**. This evaluation is supported by key variables that are believed to contribute to life satisfaction: GDP per capita, social support, healthy life expectancy, personal freedom, generosity, and perceptions of corruption.

The following Figure 1 from Visualcapitalist presents the global happiness levels in 2024, as reported in the **World Happiness Report**. The static world map provides a visual representation of the happiness scores across different countries, with the color gradient indicating the relative happiness levels. The visualization aims to highlight the disparities in happiness levels worldwide and the factors that contribute to these variations. However, through our critical analysis, we identified several strengths and weaknesses in the original visualization that we aim to improve as detailed below.



(a) Global Overview World map of Happiness Levels in 2024



(b) Map of Happiness Levels of East Asia + Oceania in 2024

Figure 1: Visualcapitalist Global Happiness Levels Report

iv. Weakness

The original visualization has several weaknesses that limit its effectiveness in conveying the information to the audience. These weaknesses include:

1. **Lack of Interactivity**: The static nature of the world map limits the audience's ability to interact with the data and explore specific countries or regions in more detail. An interactive visualization would allow users to hover over countries to see detailed information and compare happiness scores across different years.
2. **Limited Context**: While the color gradient provides a general overview of happiness levels, the visualization lacks additional context or explanations of the factors contributing to these scores. Including annotations or tool-tips that explain the key variables influencing happiness levels would enhance the audience's understanding of the data.
3. **Inconsistent Labeling**:

- The labeling of happiness scores and name of countries are inconsistent in placement or are not displayed, making it challenging for the audience to quickly identify and compare scores across countries.
- A standardized labeling format would improve the readability and clarity of the visualization. In addition to this, there are countries' happiness scores that are not labeled due to the lack of space for example Cambodia and Laos, which may lead to confusion for the audience.
- Furthermore, the lack of pointers and inconsistent formatting severely affect readability for countries such Philippines and Indonesia have confusing unclear placements as to which country the label belongs to.

4. **Limited Historical Comparison**: The visualization only presents the happiness scores for the year 2024, without providing any historical comparison or trend analysis. Including data from previous years or a time slider feature would allow users to track changes in happiness levels over time and identify emerging patterns or trends.
5. **Limited Data Insights**: The visualization focuses solely on the happiness scores of countries without providing additional insights or analysis of the data. Including annotations, data summaries, or comparisons with other relevant metrics would enrich the audience's understanding of the factors influencing happiness levels.
6. **Color Blindness**: The color gradient used in the visualization may not be accessible to color-blind users, as certain color combinations may be difficult to distinguish. Using color schemes that are accessible to users with color vision deficiencies would improve the inclusivity of the visualization.

Figure 2: Linear scale of Happiness score of each country by region

7. **Linear Scale**: The linear scale at the bottom of the region map is poorly visualized and does not provide a distinct separation and vivid depiction between the countries. This further worsened by the clustering and overlapping of icons, makes it difficult for the audience to interpret and compare the happiness scores of each country within the region.
8. **Incorrect Date for the Visualization**: The visualization is titled **"A Map of Global Happiness by Country in 2024"** but the data is from 2023. This inconsistency may lead to confusion among the audience and affect the credibility of the visualization.

II. Methods

i. Data Sources

World Happiness Report (WHR) 2021-2023: Contains happiness scores and key variables for countries worldwide. Gallup World Poll (GWP): Provides self-reported happiness evaluations on a scale from 0 to 10.

ii. Data Understanding and Cleaning

iii. Load Data

```
df <- read_excel("data/DataForTable2.1.xls")
```

iv. List and Identify Missing Countries

```
data(World)
countries <- select(World, iso_a3, name)
countries <- countries$name
```

```
# how many countries are missing in the dataset use Country name
missing_countries <- setdiff(countries, unique(df$"Country name"))
missing_countries
```

```
 [1] "Antarctica"             "Fr. S. Antarctic Lands" "Bahamas"
 [4] "Bosnia and Herz."       "Brunei"                 "Central African Rep."
 [7] "Cote d'Ivoire"          "Dem. Rep. Congo"        "Congo"
[10] "N. Cyprus"              "Czech Rep."             "Dominican Rep."
[13] "Eritrea"                "Fiji"                   "Falkland Is."
[16] "Guinea-Bissau"          "Eq. Guinea"             "Greenland"
[19] "Korea"                  "Lao PDR"                "Macedonia"
[22] "New Caledonia"          "Papua New Guinea"       "Puerto Rico"
[25] "Dem. Rep. Korea"        "Palestine"              "W. Sahara"
[28] "S. Sudan"               "Solomon Is."            "Somaliland"
[31] "Swaziland"              "Timor-Leste"            "Turkey"
[34] "Taiwan"                 "Vanuatu"
```

We can see that there are 34 countries missing from the dataset.

v. Data Cleaning and Transformation

As we will be only using data from 2021 - 2023, we will filter out the data from other years. We will also check for missing data and fill in the missing values with the mean of the available data from each country.

1) *Filter data for years 2021 - 2023:* 2) *Renaming headers and keeping only the necessary columns:* 3) *Include Missing Countries with NA Scores:*

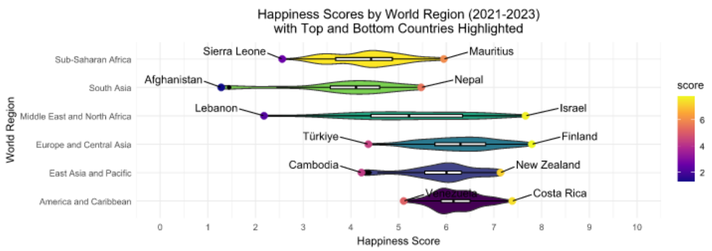
vi. Data Visualization

1) *Create violin plot:*

```
# Create the violin plot with boxplots
base_plot <- ggplot(df_combined_region, aes(y = `World Region according to the World Bank`,
x = score, fill = `World Region according to the World Bank`)) +
  geom_violin(trim = TRUE) +  # Create horizontal violin plots
  geom_boxplot(width = 0.1, fill = "white", color = "black", orientation = "y") +  # Overlay
boxplots on the violins horizontally
  scale_x_continuous(limits = c(0, 10), breaks = seq(0, 10, by = 1)) +  # Set the x-axis
```

```
limits and breaks
  scale_fill_viridis(discrete = TRUE, guide = FALSE) +  # Use viridis color scale for discrete data and remove legend
   labs(title = "Happiness Scores by World Region (2021-2023)", y = "World Region", x = "Happiness Score") +
  theme_minimal() +  # Use a minimal theme
  theme(
    plot.title = element_text(hjust = 0.5)  # Center the plot title
  )

# Highlight the top and bottom of each region with text labels of the respective country
highlight_data <- df_combined_region %>%
  group_by(`World Region according to the World Bank`) %>%
  filter(score == max(score) | score == min(score))

# Add the highlighted points and text labels to the base plot
final_plot <- base_plot +
  geom_point(data = highlight_data, aes(y = `World Region according to the World Bank`, x = score, color = score), size = 3) +
  geom_text_repel(data = highlight_data, aes(y = `World Region according to the World Bank`, x = score, label = country),
                  nudge_x = ifelse(highlight_data$score >= mean(highlight_data$score), 1, -1), nudge_y = 0.3, direction = "y") +
  scale_color_viridis_c(option = "plasma", na.value = "grey90") +
  labs(title = "Happiness Scores by World Region (2021-2023)\nwith Top and Bottom Countries Highlighted")

# Print the final plot
print(final_plot)
```

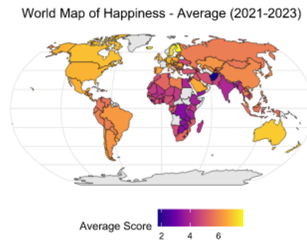
```
Warning: The `guide` argument in `scale_*()` cannot be `FALSE`. This was deprecated in
ggplot2 3.3.4.
ℹ Please use "none" instead.
```



2) *World Map:*

```
# Create bins for the scores with 0.5 intervals
bins <- seq(0, 10, by = 0.5)
labels <- paste0(bins[-length(bins)], "-", bins[-1])
df_combined <- df_combined %>%
  mutate(score_bin = cut(score, breaks = bins, labels = labels, include.lowest = TRUE, right = FALSE))

# Check for any remaining missing values
missing_scores <- df_combined %>% filter(is.na(score))
if (nrow(missing_scores) > 0) {
  warning("There are still countries with missing scores.")
}

# Inspect the data for missing scores
# print(missing_scores)

# Get world data
world <- ne_countries(scale = "small", returnclass = "sf")

# Compute the average score for each country over the years
df_avg <- df_combined %>%
  group_by(country_code) %>%
  summarize(avg_score = mean(score, na.rm = TRUE))

# Merge world data with the average scores
world_data_avg <- left_join(world, df_avg, by = c("iso_a3" = "country_code"))

# Check for missing data after the join
missing_world_avg <- world_data_avg %>% filter(is.na(avg_score))

# print(missing_world_avg)
```

```
# Plot the average data
ggplot() +
  geom_sf(data = world_data_avg, aes(fill = avg_score), color = "black") +
  scale_fill_viridis_c(option = "plasma", na.value = "grey90", name = "Average Score") +
  theme_minimal() +
  theme(legend.position = "bottom") +
  labs(title = "World Map of Happiness - Average (2021-2023)", fill = "Average Score") +
  theme(plot.title = element_text(hjust = 0.5)) +
  coord_sf(crs = "+proj=robin")
```



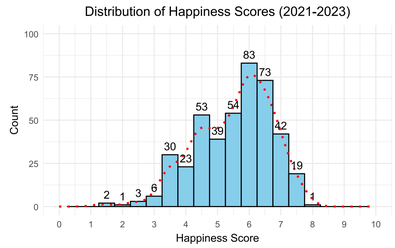
3) *Distribution Chart of Scores:*

```
ggplot(df_combined, aes(x = score)) +
  geom_histogram(binwidth = 0.5, fill = "skyblue", color = "black") +
  geom_density(aes(y = 0.5 * ..count..), color = "red", size = 1, linetype = "dotted") +
  stat_bin(binwidth = 0.5, geom = "text", aes(label = ifelse(..count.. > 0, ..count.., "")), vjust = -0.5, color = "black", position = position_stack(vjust = 1.0)) +
  scale_x_continuous(limits = c(0, 10), breaks = seq(0, 10, by = 1)) +
  scale_y_continuous(limits = c(0, 100)) +
  labs(title = "Distribution of Happiness Scores (2021-2023)",
       x = "Happiness Score",
       y = "Count") +
  theme_minimal() +
  theme(plot.title = element_text(hjust = 0.5))
```

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

```
Warning: The dot-dot notation (`..count..`) was deprecated in ggplot2 3.4.0.
ℹ Please use `after_stat(count)` instead.
```

```
Warning: Removed 2 rows containing missing values or values outside the scale range
(`geom_bar()`).
```



Overall:
The distribution suggests that while there is a central tendency around a happiness score of 6, there is considerable variability in happiness scores among countries, with more countries experiencing moderate to low levels of happiness.
Distribution Shape:
The distribution of happiness scores is approximately normal with a slight positive skew. This means most countries have happiness scores clustered around the middle range (4 to 6), but there are more countries with lower scores than higher scores.

vii. Critical Assessment and Proposed Improvements

1. Interactive Data Visualization:

- Implement hover-over information for detailed country data.
- Add filtering and sorting capabilities.
- Include historical comparison using a time slider.

2. Consistent Labeling and Annotations:

- Standardize labeling format for readability.
- Provide annotations explaining key variables influencing happiness.

3. Accessible Design:

- Ensure a user-friendly interface.
- Use color schemes accessible to color-blind users.

4. Data Insights and Analysis:

- Provide summaries and comparisons with relevant metrics.
- Enhance linear scale visualization for regional comparisons.

These methods aim to create an improved, interactive visualization of global happiness levels, offering better insights and user engagement.

III. Findings

i. New World Map

ii. Happiness Score By World Region

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim aeque doleamus animo, cum corpore dolemus, fieri.

1. Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do.
2. Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do.
3. Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim aeque doleamus animo, cum corpore dolemus, fieri tamen permagna accessio potest, si aliquod aeternum et infinitum impendere malum nobis opinemur. Quod idem licet transferre in voluptatem, ut.

Table 1: Lorem ipsum dolor sit amet

Lorem ipsum dolor sit.	Lorem ipsum.	Lorem ipsum.
Lorem ipsum dolor.	Lorem ipsum.	α
Lorem ipsum.	Lorem ipsum.	β
Lorem.	Lorem.	γ
Lorem ipsum dolor.	Lorem.	θ

Lorem Table 1 ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim aeque doleamus animo, cum corpore dolemus, fieri tamen permagna accessio potest, si aliquod aeternum et infinitum impendere malum nobis opinemur. Quod idem licet transferre in voluptatem, ut postea variari voluptas distinguique possit, augeri amplificarique non possit. At etiam Athenis, ut e patre audiebam facete et urbane Stoicos irridente, statua est in quo a nobis philosophia defensa et.

iii. Lorem ipsum dolor sit amet

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim aeque doleamus animo, cum corpore dolemus, fieri tamen permagna accessio potest, si aliquod aeternum et infinitum impendere malum nobis opinemur. Quod idem licet transferre in voluptatem, ut postea variari voluptas distinguique possit, augeri amplificarique non possit. At etiam Athenis, ut e patre.

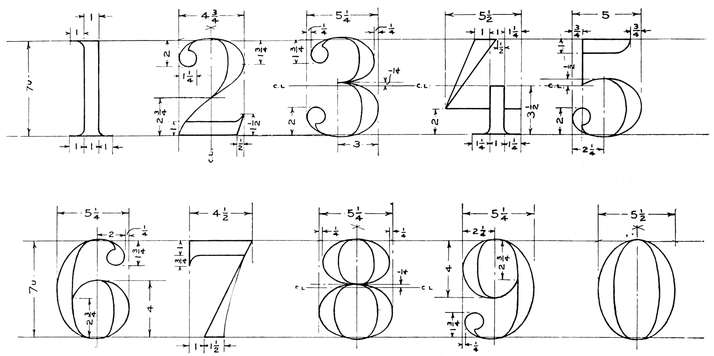


Figure 3: Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do.

IV. Findings

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim aeque doleamus animo, cum corpore dolemus, fieri tamen permagna accessio potest, si aliquod aeternum et infinitum impendere malum nobis opinemur. Quod idem licet transferre in voluptatem, ut postea variari voluptas distinguique possit, augeri amplifi- carique non possit. At etiam Athenis, ut e patre audiebam facete et urbane Stoicos irridente, statua est in quo a nobis philosophia defensa et.

- Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do.
- Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do.
- Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do.

i. Lorem ipsum etiam

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim aeque doleamus animo, cum corpore dolemus, fieri tamen permagna accessio potest, si aliquod aeternum et infinitum impendere malum nobis opinemur.

V. Conclusion

Quod idem licet transferre in voluptatem, ut postea variari voluptas distinguique possit, augeri amplificarique non possit. At etiam Athenis, ut e patre audiebam facete et urbane Stoicos irridente, statua est in quo a nobis philosophia defensa et collaudata est, cum id, quod maxime placeat, facere possimus, omnis voluptas assumenda est, om- nis dolor repellendus. Temporibus autem quibusdam et.
Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim aeque doleamus animo, cum corpore dolemus, fieri tamen permagna accessio potest, si aliquod aeternum et infinitum impendere malum nobis opinemur.

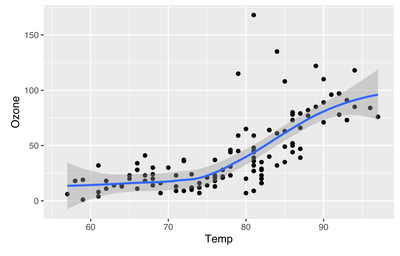


Figure 4: Temperature and ozone level

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim aeque doleamus animo, cum corpore dolemus, fieri tamen permagna accessio potest, si aliquod aeternum et infinitum impendere malum nobis opinemur. Quod idem licet transferre in voluptatem, ut postea variari voluptas distinguique possit, augeri amplificarique non possit. At etiam Athenis, ut e patre audiebam facete et urbane Stoicos irridente, statua est in quo a nobis philosophia defensa et collaudata est, cum id, quod maxime placeat, facere possimus, omnis voluptas assumenda est, om- nis dolor repellendus. Temporibus autem quibusdam et.

- Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do.
  - Lorem ipsum dolor sit amet.
  - Lorem ipsum dolor sit amet, consectetur adipiscing elit.
- Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore.
  - Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed.
  - Lorem ipsum dolor sit amet, consectetur adipiscing.

$$\sum_{(k=1)}^{n} k = \frac{n(n+1)}{2} = \frac{n^2+n}{2} \tag{2}$$

VI. Acknowledgement

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim aeque doleamus animo, cum corpore dolemus, fieri tamen permagna accessio potest, si aliquod aeternum et infinitum impendere malum nobis opinemur. Quod idem licet transferre in voluptatem, ut postea variari voluptas distinguique possit, augeri amplificarique non possit. At etiam Athenis, ut e patre audiebam facete et urbane Stoicos irridente, statua est in quo a nobis philosophia defensa et collaudata est, cum id, quod maxime placeat, facere possimus, omnis voluptas assumenda est, om- nis dolor repellendus. Temporibus autem quibusdam et.