

Visualization-Driven Analysis of the World Happiness Dataset

prepared by: Aleesha Bukhari, Wajeeha Khalid, Ayesha Bint Imran
Department of Data Science and Artificial Intelligence, FAST-NUCES

Abstract—This report focuses entirely on visualization-based analysis of the World Happiness dataset consisting of 158 countries and 12 variables. The goal is to explore how happiness varies globally by using advanced plots such as correlation matrices, bubble charts, ridgeline plots, lollipop charts, pairplots, choropleth maps, and Sankey diagrams. Instead of statistical modelling, this report emphasizes the interpretability, insights, and comparative analysis derived through graphical techniques.

Index Terms—Data Visualization, Happiness Score, Correlation, Ridgeline Plot, Bubble Chart, Pairplot, Choropleth Map, Sankey Diagram

I. INTRODUCTION

The World Happiness dataset collects multiple quantitative indicators such as Economy (GDP per capita), Family, Health, Freedom, Trust, and Generosity. Visualizations serve as the most intuitive way to identify global patterns, outliers, inequalities, and factor interactions.

This report compiles a set of complementary visualization methods designed to uncover how happiness varies across nations, what factors drive it, and how regions differ in wellbeing.

II. CORRELATION MATRIX VISUALIZATION

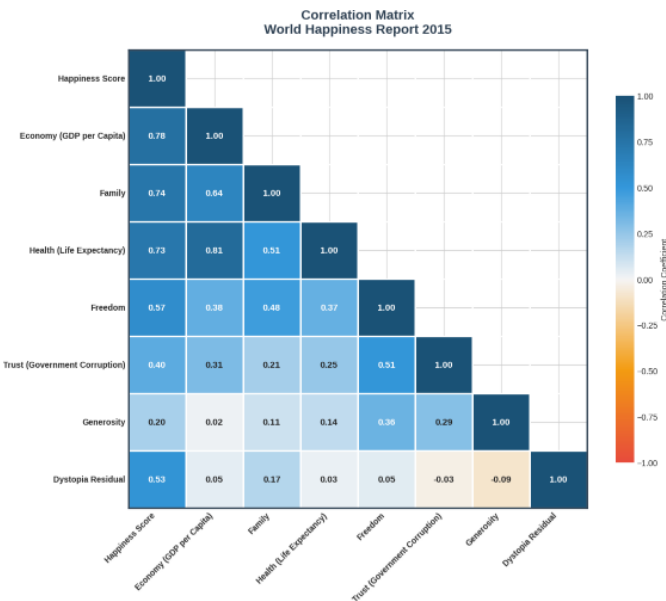


Fig. 1. Correlation Matrix of Happiness Factors

Interpretation

Economy, Family, and Health show the strongest correlations with Happiness. Trust and Generosity show weak influence. This confirms that money, social support, and life expectancy strongly drive wellbeing.

III. HORIZONTAL BAR CHART VISUALIZATION

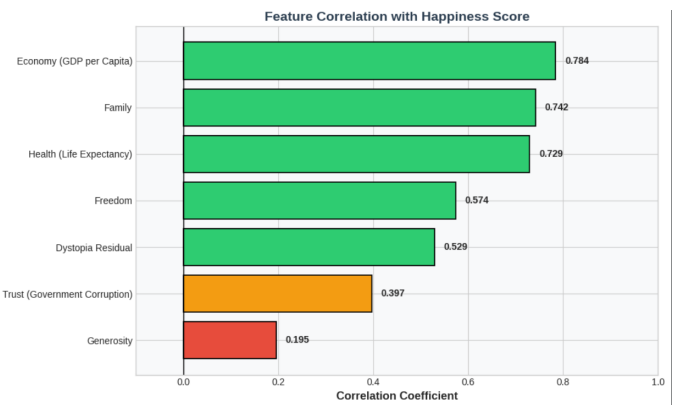


Fig. 2. Importance of Each Factor Relative to Happiness

Interpretation

The chart ranks factors by correlation strength: Economy (0.78), Family (0.74), Health (0.73) dominate, while Generosity contributes the least.

IV. RIDGELINE PLOT VISUALIZATION

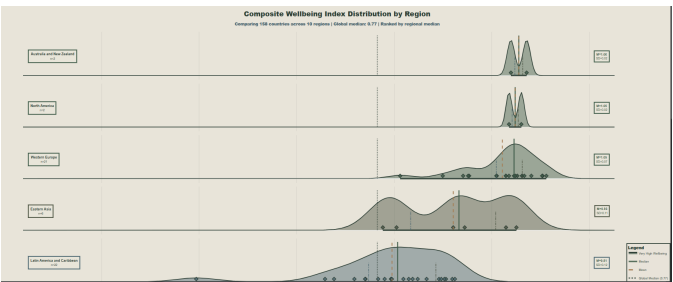


Fig. 3. Ridgeline Plot of Happiness Distributions by Region

Interpretation

Western Europe and Oceania show the highest and most consistent scores. Sub-Saharan Africa has the lowest distribution with wider spread, highlighting uneven wellbeing.

V. LOLLIPOP CHART (DIVERGING BAR CHART)

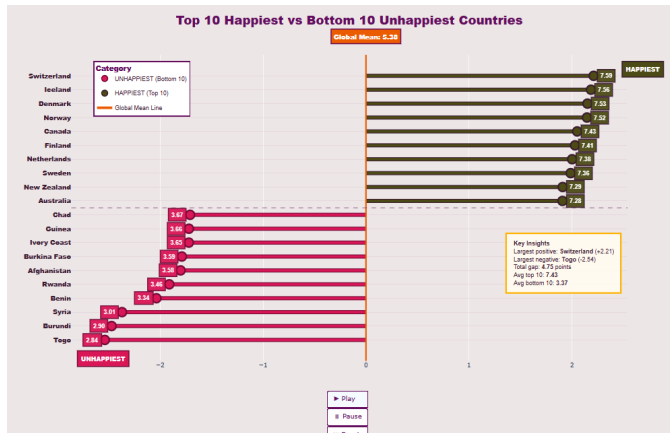


Fig. 4. Top 10 vs Bottom 10 Countries Relative to Global Mean

Interpretation

This chart visually contrasts the happiest and unhappiest countries. Switzerland is +2.21 above global mean, Togo is -2.54 below it. This demonstrates massive inequality.

VI. 3D SCATTERPLOT VISUALIZATION

3D Composite Wellbeing vs Economic Dependency vs Happiness

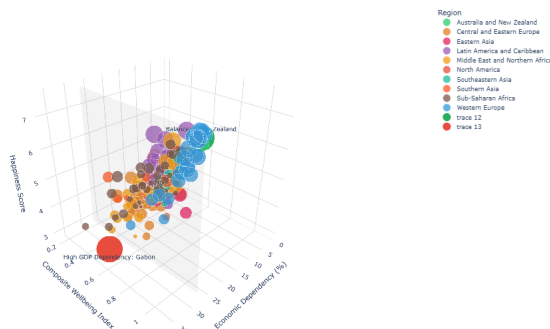


Fig. 5. 3D Scatter: Economic Dependency, Wellbeing, and Happiness

Interpretation

A strong upward trend exists between Composite Wellbeing and Happiness. Economic Dependency does not correlate strongly, but reveals outliers (e.g., Gabon and New Zealand).

VII. PAIRPLOT VISUALIZATION

Interpretation

Confirms statistical relationships: Economy vs Health is extremely strong (0.81). Happiness is strongly predictable using Economy, Health, and Family.

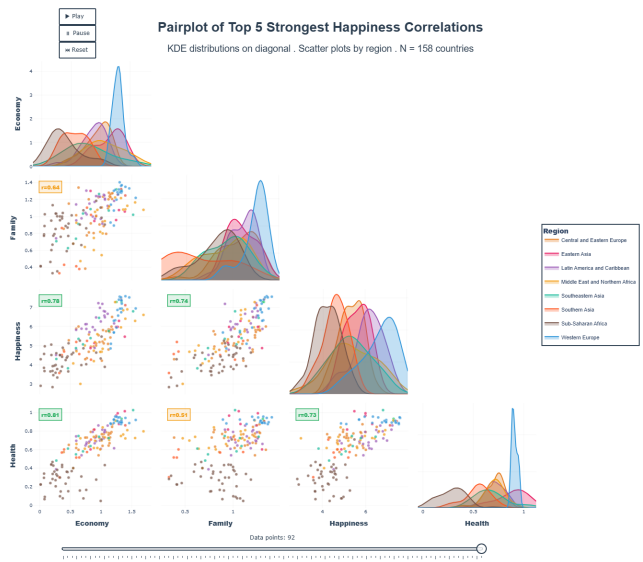


Fig. 6. Pairwise Scatter and KDE Distributions

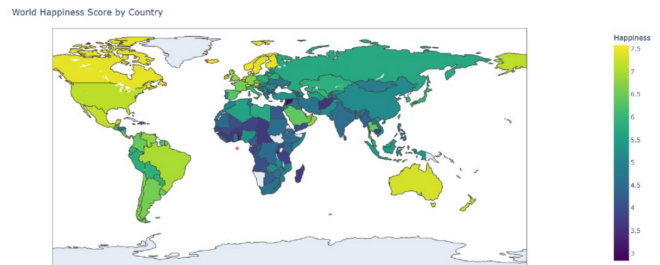


Fig. 7. Global Happiness Distribution (Choropleth Map)

VIII. CHOROPLET MAP

Interpretation

Shows regional clusters: Western Europe and North America exhibit the highest scores. Sub-Saharan Africa and conflict regions show the lowest.

IX. BUBBLE CHART VISUALIZATION

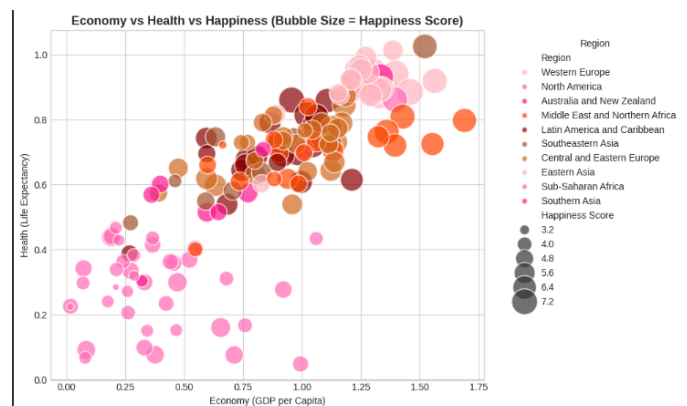


Fig. 8. Bubble Chart: GDP, Health, and Happiness

Interpretation

The diagonal trend confirms that Happiness increases with GDP and Health. Bubble size reveals how much these factors combine to produce higher happiness.

X. ICICLE CHART VISUALIZATION

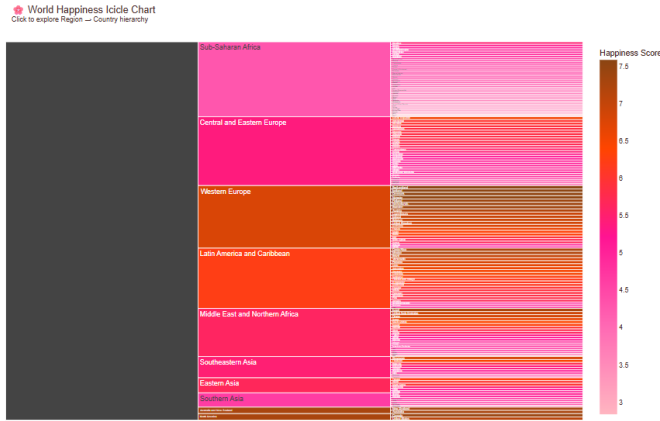


Fig. 9. Hierarchical Icicle Chart: Region → Country → Score

Interpretation

Western Europe dominates in total happiness. Each block's size reveals consistent high scores. Other regions show more variability.

XI. RADIAL HEATMAP VISUALIZATION

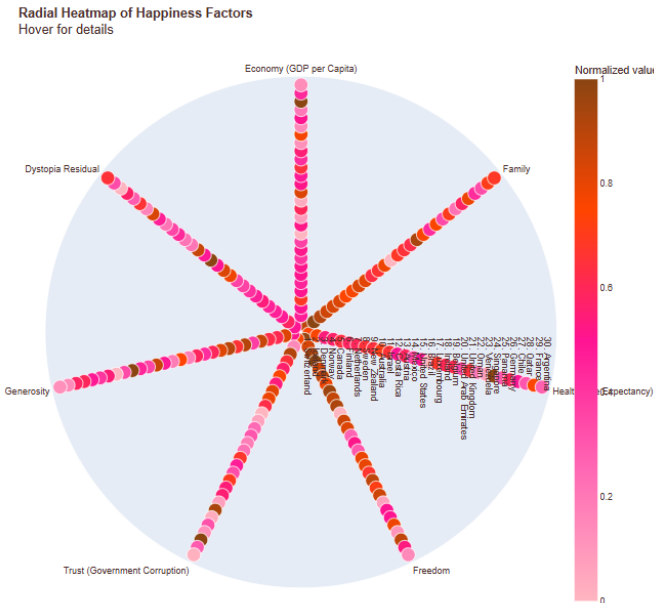


Fig. 10. Radial Heatmap of Top 30 Countries Across Factors

Interpretation

Inner rings (top countries) consistently feature darker values for Economy and Health. Trust and Generosity show large variances even among top performers.

XII. SANKEY DIAGRAM VISUALIZATION

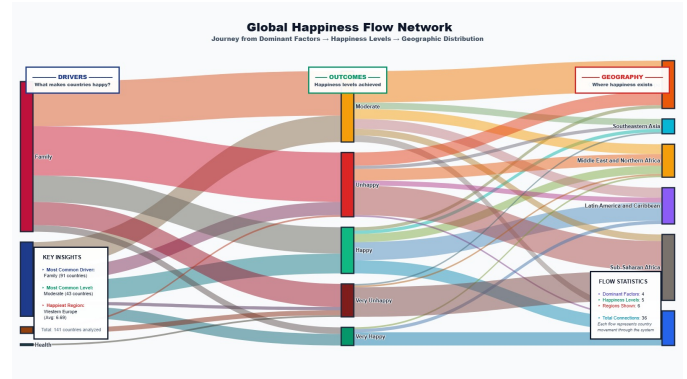


Fig. 11. Sankey Flow: Drivers → Happiness Levels → Regions

Interpretation

Family is the strongest driver (51 countries). The majority of countries fall into the “Moderate” happiness level. Western Europe receives the largest flow of high-happiness countries.

XIII. CONCLUSION

This visualization-focused report reveals global patterns in happiness through a suite of advanced graphical techniques. Across all visualizations, the most consistent insights are:

- Economy, Health, and Family are the strongest contributors.
- Western Europe and Oceania dominate high wellbeing.
- Sub-Saharan Africa contains the lowest happiness clusters.
- Outliers such as Gabon and New Zealand provide valuable context.

Visualization provides an intuitive and powerful lens for understanding global happiness dynamics.

REFERENCES

- [1] World Happiness Report Dataset, Kaggle, 2024.