

Visual Analytics of Global Happiness



INTRODUCTION & MOTIVATION

- Happiness is a holistic measure of societal well-being, shaped by a combination of economic and non-economic factors. Understanding how these elements interact is essential but often complex.
- This project explores the key drivers of happiness and introduces an interactive tool to analyze global trends among countries/ regions and support predictive/ future forecasting.

DATASET & DATA PREPATION

- Our dataset comprises historical happiness scores across 173 countries, covering the years 2014 to 2024.
- The data is sourced from the **World Happiness Report** and includes key indicators such as **ladder score, economy score, social support, life expectancy, freedom, generosity, perceptions of corruption, and a calculated residual score**. Additional columns include country, year, global rank, and region.
- For countries with missing values, imputation was performed using the mean of the previous three years for the same country and variable.

SHINY APP OVERVIEW

R SHINY PACKAGES



- Tidyverse, sf, tmap, dplyr, spatstat, spdep, lubridate, leaflet, plotly, ggplot2, DT, plm etc.

MODULES COVERED

1 GEOSPATIAL/ ASPATIAL

- Choropleth Map & Proportional Map
- LISA cluster map
- Aspatial Map

2 FORECAST & PANEL

- Forecasting using Exponential smoothing(ETS)
- Line graph for Trends
- Panel analysis

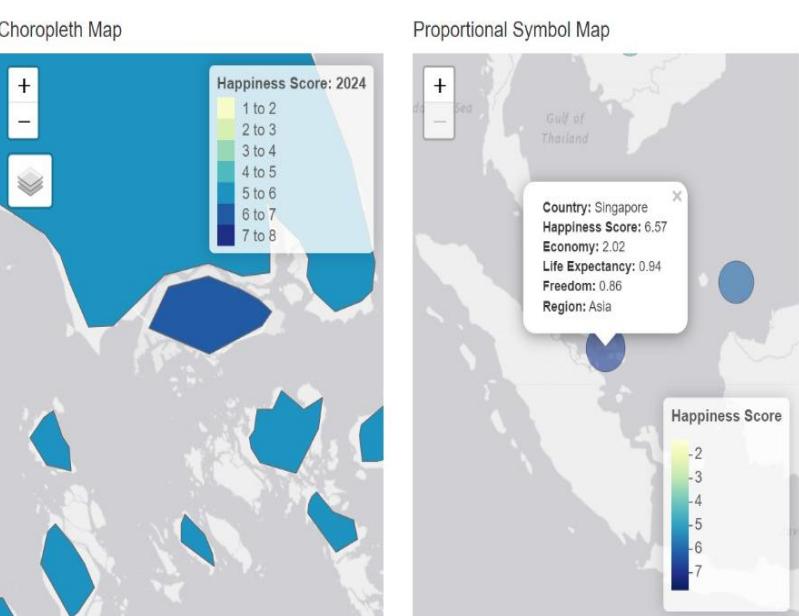
3 CLUSTERING

- Time Series Clustering
- Multivariate Clustering

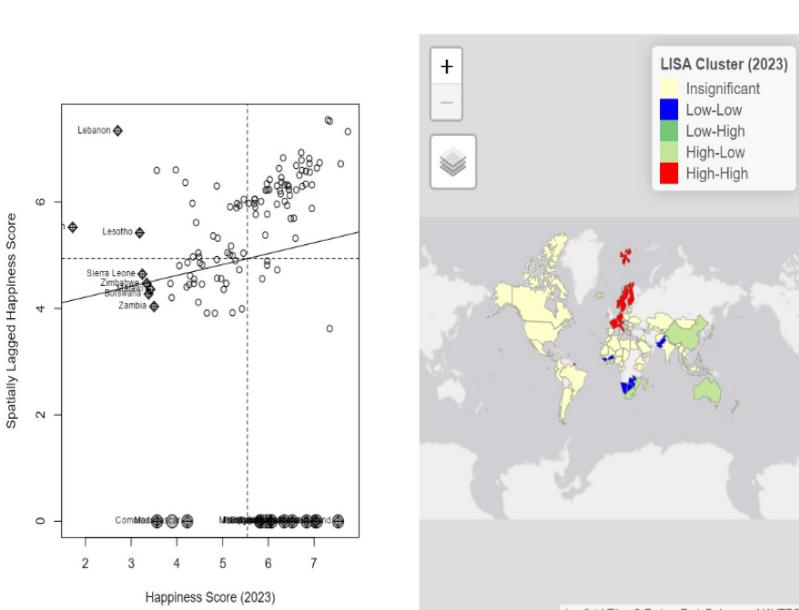


MODULE 1 – GEOSPATIAL/ ASPATIAL

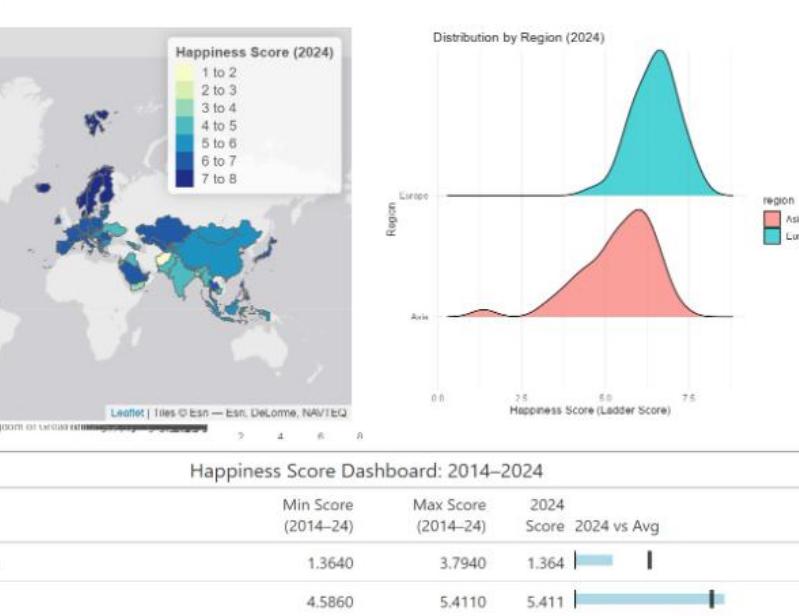
World Happiness Explorer



Spatial Clustering Analysis



World Happiness Explorer



CORE FUNCTIONALITIES

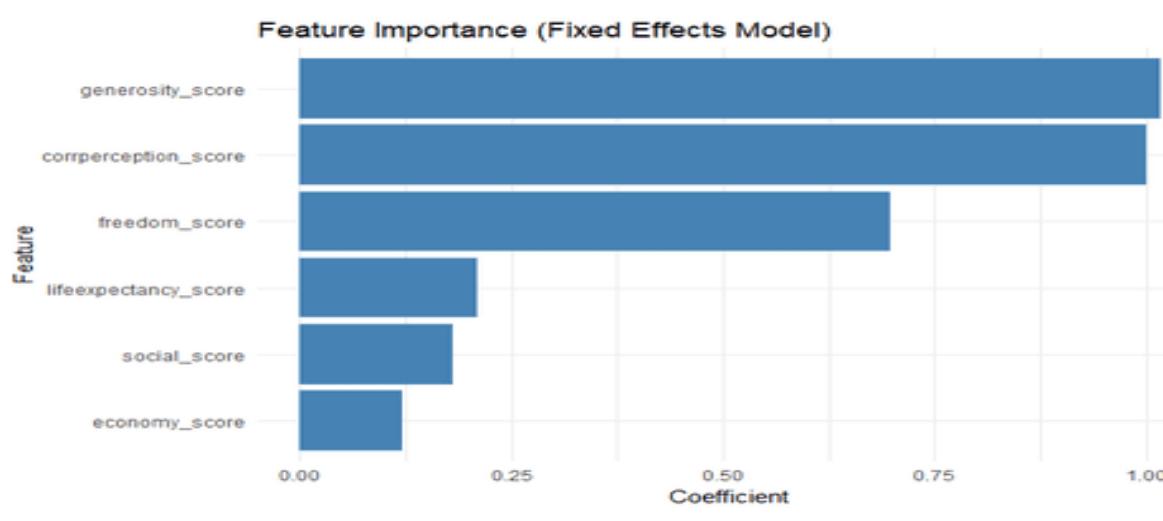
- Interactive map:** User can explore global happiness score on both Choropleth Map and Proportional Symbol Map across year, region and countries
- Spatial Clustering (Moran's I & LISA Map):** Identifies global patterns and local clusters of happiness by measuring spatial autocorrelation and highlighting significant groupings (e.g., high-high, low-low, outliers).
- Region & Country Comparison:** Enables filtering by region and visualizing spatial trends, with density plots to compare happiness distributions across regions.
- Time Series Dashboard:** Interactive table with spark bars to track each country's happiness trends, extremes, and comparisons over time.

EXPLORATORY MODEL USING PANEL DATA

- This model explores the key factors influencing happiness across countries from 2014-2024 using panel data techniques.
- By applying fixed and random effects models, we aim to uncover the relationships between various social, economic, and political factors, while considering both cross-sectional and time-series variations.

MODEL SELECTION – FIXED EFFECTS (FE) MODEL

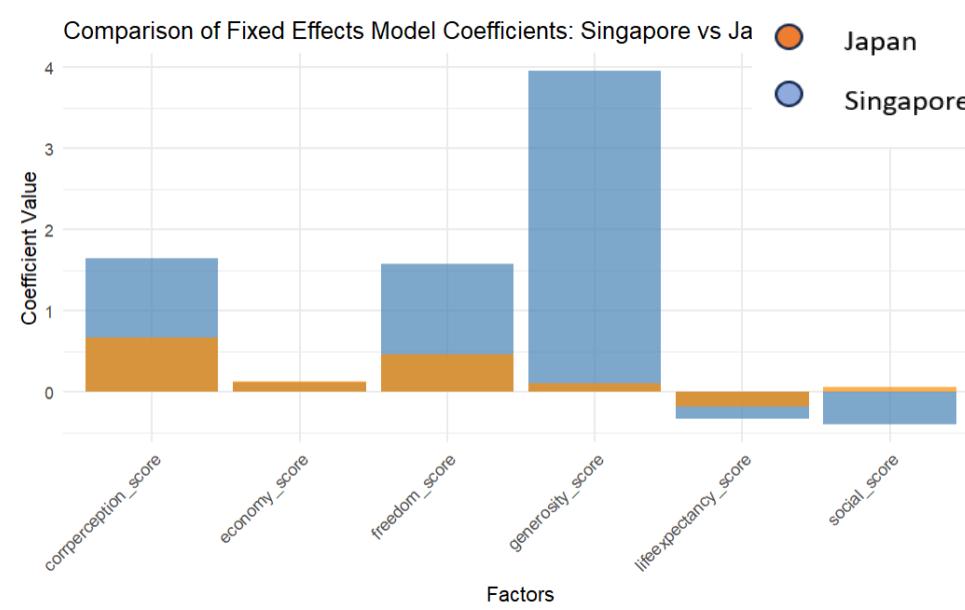
- The **Fixed Effects (FE) model** was selected after performing a **Hausman test**, which compared it against the **Random Effects (RE) model**. The Hausman test revealed a significant difference between the two models ($p\text{-value} < 2.22\text{e-}16$), indicating that the **Fixed Effects (FE) model** is more appropriate for controlling for unobserved individual heterogeneity



- Top Influencing Factors
- 1 Generosity
 - 2 Corruption Perception
 - 3 Freedom
 - 4 Life Expectancy
 - 5 Social Support

- The dashboard allows users to easily compare the impact of various factors (such as economy, freedom, and generosity) on happiness scores across different countries and regions. It provides an interactive interface where users can adjust input values for these factors and see how the predicted happiness score changes.

COMPARATIVE ANALYSIS



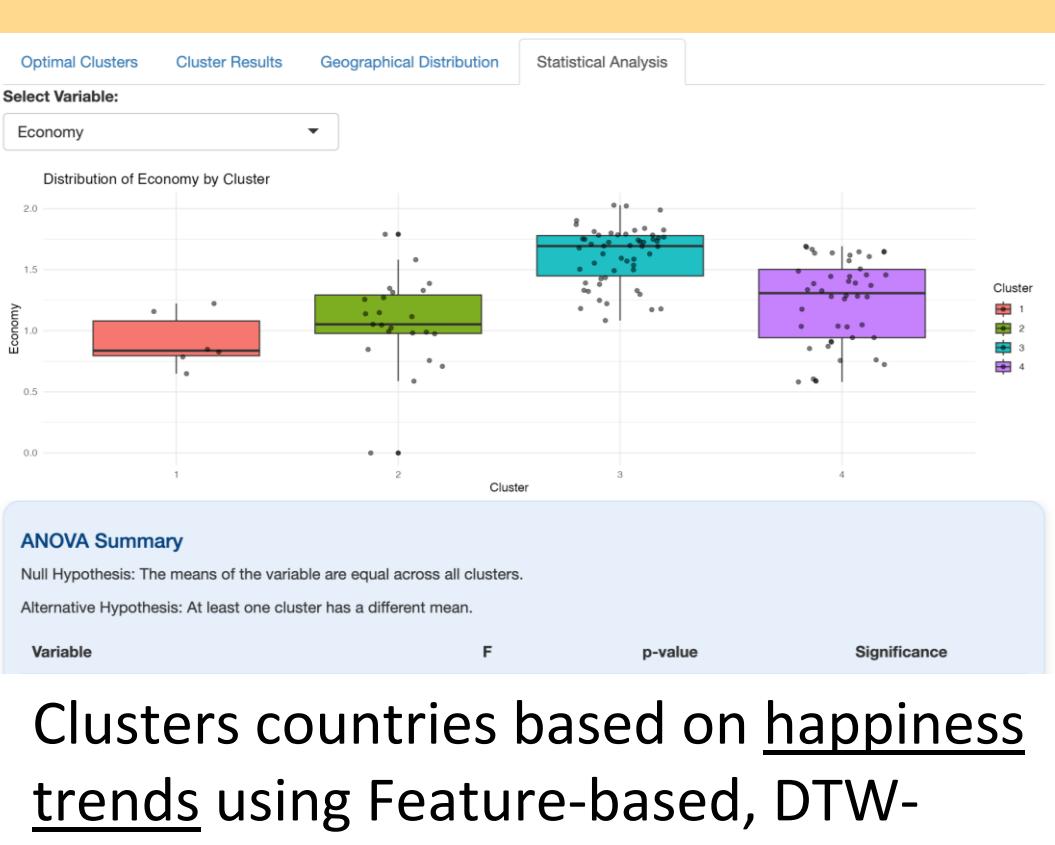
Freedom and **generosity** appear to be significantly more influential on happiness in Singapore than in Japan. Conversely, while both countries show **negative** effects from **life expectancy** on happiness. Social score also plays a contrasting role: it negatively impacts happiness in Singapore, whereas it has almost no effect in Japan.

MODULE 3 – CLUSTERING

Clustering patterns of happiness across countries and time:

The following two models aim to uncover yearly clusters based on explanatory variables and evolving happiness trends among countries.

TIME-SERIES



- Clusters countries based on happiness trends using Feature-based, DTW-based, and Shape-based methods.

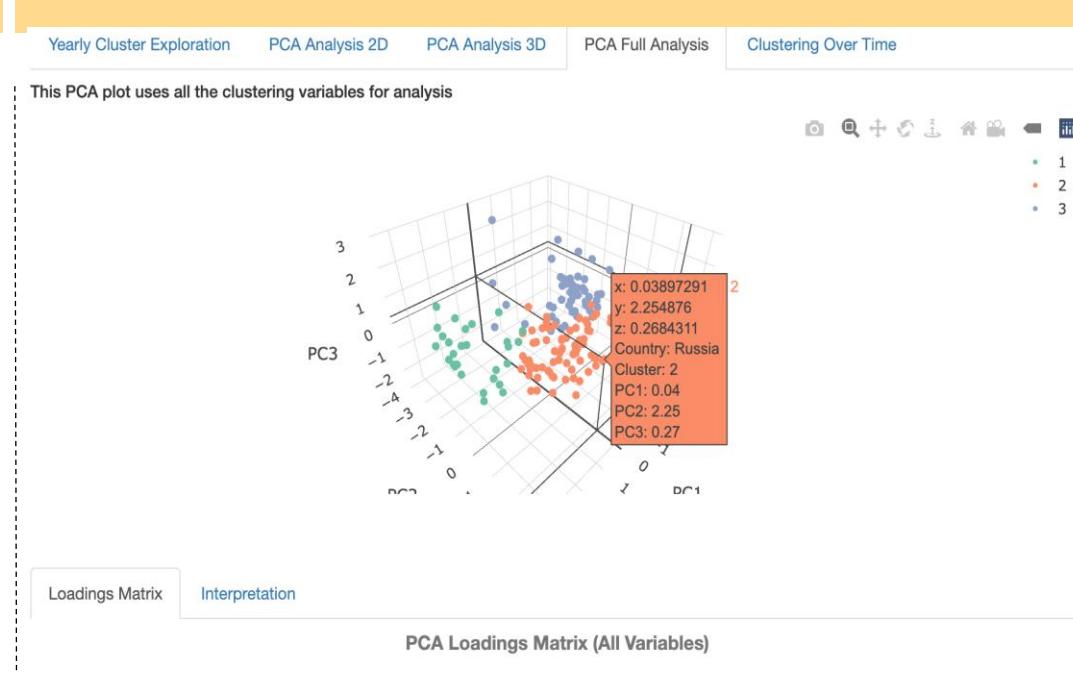
Visualizations:

- Temporal patterns of clusters
- Geographic distribution
- Differences in explanatory variables across clusters (Anova test)

Rationale

- Each method has its own strengths and limitations; no single metric would be able to identify the best model.
- Deeper insights come from interactive, step-by-step exploration within the application.

MULTIVARIATE



- Clusters countries yearly based on happiness-related factors using K-Means or Hierarchical clustering.

Visualizations:

- Interactive cluster map with cluster validation metrics and profiles
- PCA-based cluster visualization (2D & 3D) for identifying patterns
- Temporal stacked bar plot

Rationale

- Enables exploration of how countries differ across years and drivers of cluster formation.

CONCLUSION & FUTURE WORK

- Future work could explore complementary clustering techniques—such as spatially constrained hierarchical clustering and spatial point pattern analysis—to deepen spatial insights.
- Additionally, incorporating temporal-spatial clustering and predictive modeling could uncover evolving patterns over time and enable forecasts of future happiness trends based on key socio-economic indicators.

Andrea Yeo Si Han; Dhreeti Shah; Ou Yiming

