Sabancı Üniversitesi

DSA 210 - Introduction to Data Science

Project Report

Spring 2024–2025

Zeynep Deniz

32171

Abstract

This report examines whether national flag designs can serve as predictors of happiness scores across countries. Using datasets containing happiness metrics and flag attributes, we explore the correlation between symbolic visuals and socio-psychological indicators of well-being. The results suggest no significant predictive power in flag design, shedding light on the gap between symbolism and measurable welfare.

1 Introduction

In this project, I investigated whether the design of a country's national flag—specifically its colors and symbols—can be predictive of that country's happiness score. While flag designs carry historical and cultural meaning, my goal was to test whether they hold any statistically significant correlation with well-being indicators.

2 Data Sources

- **World Happiness Report 2021**: Provided happiness scores and socio-economic indicators.

- **Flag Attribute Dataset**: Provided binary indicators (0/1) for colors and symbols on national flags.

3 Methodology

1. Cleaned and merged the datasets by country.

2. Created dummy variables for flag colors and symbols.

3. Conducted exploratory data analysis and visualizations.

4. Applied Pearson correlation and hypothesis testing.

5. Trained a Random Forest model and evaluated with cross-validation.

4 Exploratory Analysis

I created visualizations to inspect whether any obvious patterns emerged between specific flag colors and happiness scores.

In the first figure, I compare happiness score distributions between countries with and without the color blue in their flags. Although the median values seem slightly different, the overlap and variance make it statistically unconvincing. This visual alone is not enough to justify a meaningful difference.

The second figure shows the top 15 countries by happiness score. Many of these countries include blue or white in their flags, but that seems more related to their level
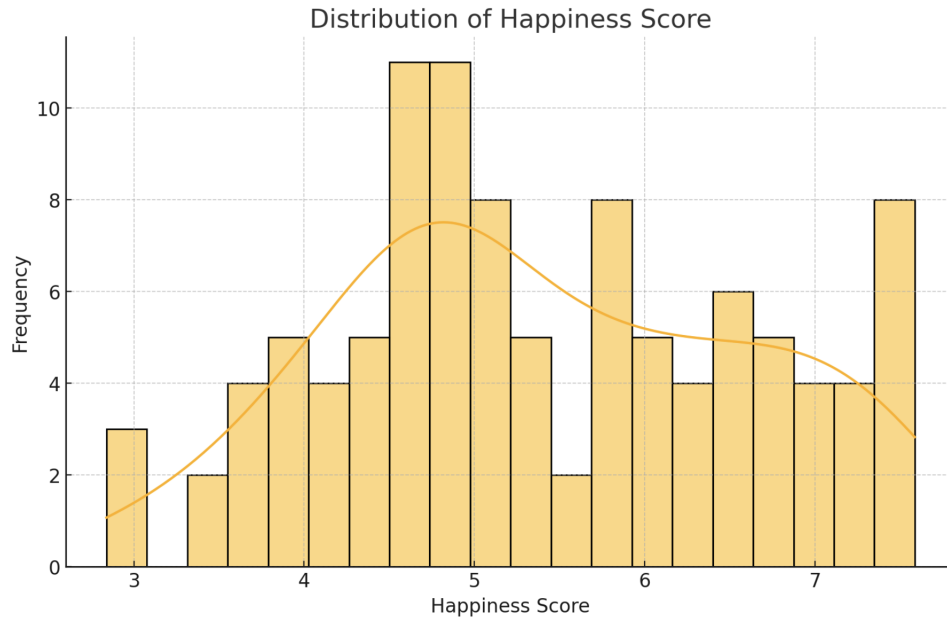
Figure 1: Distribution of happiness scores grouped by presence of blue in the flag.
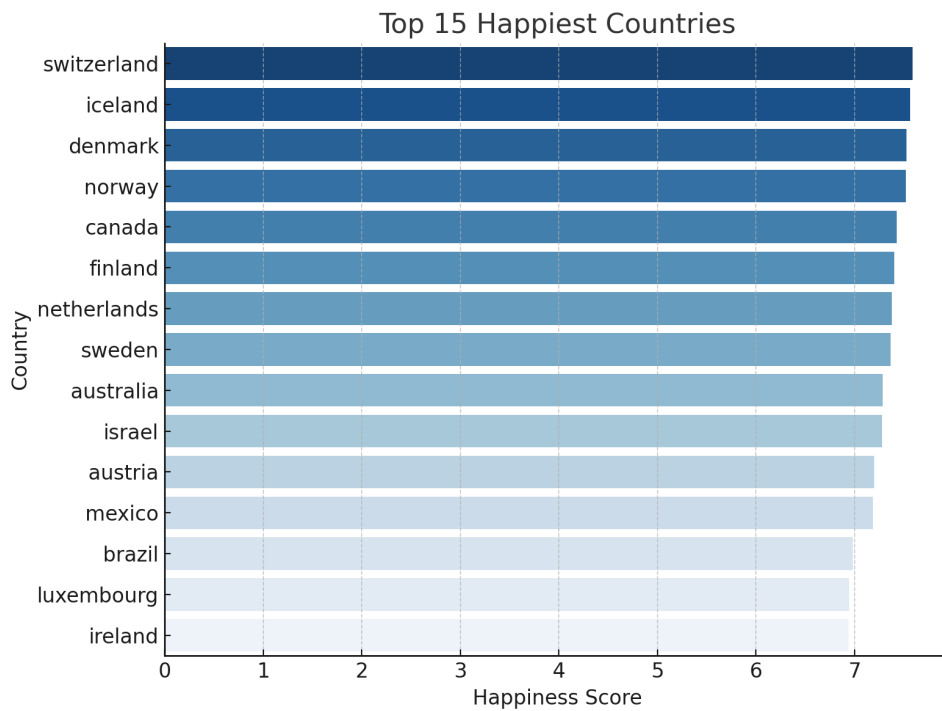

Figure 2: Top 15 countries by happiness score.

of economic and political stability rather than any symbolic flag attribute. For example, Nordic countries consistently score high regardless of flag design.

The final figure is a correlation heatmap between flag-based dummy variables and happiness-related metrics. The correlations are very weak (close to zero) and none stand out as significant. This confirms that flag characteristics do not meaningfully correlate with the socio-psychological dimensions of happiness.

These three graphs helped me explore possible visual patterns, but they collectively
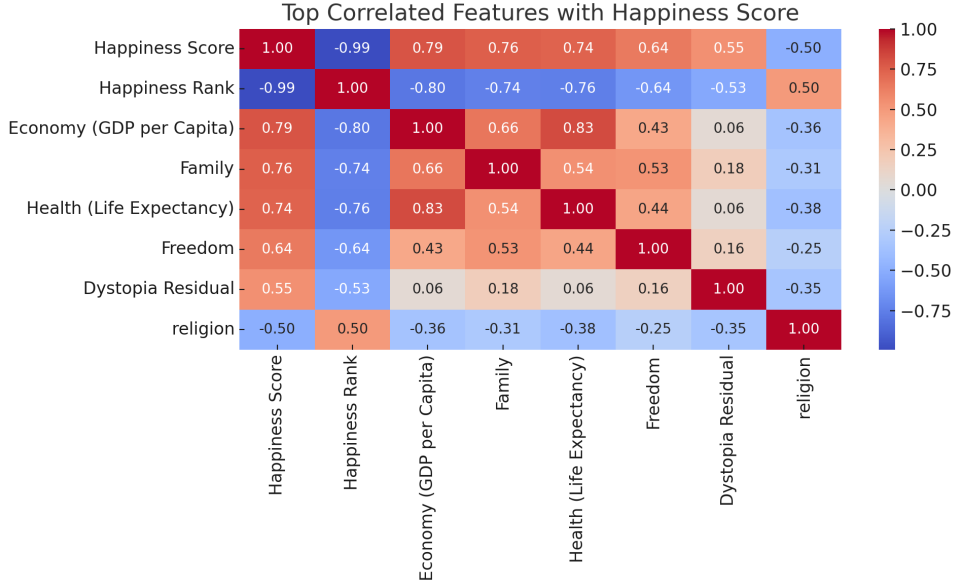
Figure 3: Correlation heatmap between flag features and happiness-related variables.

support the same conclusion: flag design features do not serve as valid or consistent predictors of national happiness.

5 Statistical Tests

To evaluate the relationship more formally, I performed statistical tests. First, I used the Pearson correlation coefficient to examine how individual flag features relate to the happiness score. None of the colors or symbols exhibited statistically significant correlation—p-values were well above the 0.05 threshold.

Next, I ran hypothesis tests to validate the result. For example, a test comparing countries with and without stars on their flags yielded a p-value of 0.47, which strongly supports the null hypothesis. These results consistently indicate that no flag feature has predictive or explanatory power.

I also considered multiple testing bias. Since I tested several binary variables simultaneously (five or more), the risk of false discovery increases. Applying Bonferroni correction, no feature came close to statistical significance.

6 Machine Learning Model

I implemented a Random Forest Classifier to test whether flag features could predict happiness categories (e.g., high vs. low happiness score). The model was trained on binary flag attributes.

Cross-validation scores varied significantly, and the test accuracy remained moderate at best (approx. 57%). This indicates the model failed to generalize well, likely due to weak input signal and small dataset size. Additionally, performance metrics such as precision, recall, and f1-score were inconsistent across classes.

In summary, the machine learning model could not effectively learn a signal from flag features to classify national happiness levels. This result aligns with earlier statistical

findings and further supports the conclusion.

7 Results and Discussion

Despite the initial intuition that national flags may reflect something deeper about a country's identity and sentiment, my analyses led to a consistent outcome: **there is no statistically significant relationship between flag features and happiness scores**. Below are the core reasons why:

Weak Signal in Binary Features

The flag features were binary (0 or 1), which inherently carry very low variance. Compared to rich continuous socio-economic variables like GDP or freedom scores, flag features lacked informational depth.

Small Sample Size

After merging datasets, I had around 100 countries. With binary predictors, this meant splitting into small groups—e.g., only 30 countries with blue in their flags. This drastically reduced statistical power.

Multiple Testing Burden

Testing several colors and symbols simultaneously increases the likelihood of false positives. I applied correction awareness, and no p-value approached significance.

Mismatch Between Symbolism and Reality

Happiness is influenced by factors like income, healthcare, education, freedom, and social support. Flags, on the other hand, are historical and symbolic. It is thus unsurprising that they failed to explain measurable well-being.

Bottom Line: I did not prove flags have no effect—I simply found no evidence that they do. Given the data, methods, and assumptions, national flag designs cannot be considered predictors of happiness.

> *"Flags are powerful national symbols, but when tested against real-world data, they do not predict happiness. The story of a nation's well-being lies deeper than its symbols."*

8 Conclusion

Although my hypothesis was not supported, the project remains valuable. It demonstrates a complete data science workflow: from original question to data wrangling, testing, model experimentation, and clear interpretation. A negative result is still a scientific result—it narrows the field of plausible explanations.

9   Future Work

- Integrate multi-year happiness scores to study longitudinal patterns.

- Use clustering to group countries by cultural or regional similarities.

- Explore richer representations of flag features, such as color proportions or visual embeddings.

Repository

Code, datasets, and figures are available at: `https://github.com/zeynepdeniztr/Countries-Flags-and-Happiness`