Eurovision: Do countries favor their neighbors while voting?

Kai Lüdemann
Matrikelnummer 6963776
kai.luedemann

Valentin Schmidt
Matrikelnummer 5552070
valentin.schmidt

Keyu Wang
Matrikelnummer 6999777
keyu.wang

Pablo Santos Blázquez
Matrikelnummer 6958794
pablo.santos-blazquez[@student.uni-tuebingen.de]

Abstract

Voting patterns in the Eurovision Song Contest may be biased toward neighboring countries. Although intended to promote European unity, relationships between neighboring countries may still be prominent in voting. This report investigates the impact of geographic distance on voting behavior while accounting for the effects of gross domestic product (GDP) per capita and music preferences. We used clustering, linear regression, and linear mixed models to analyze this question. Results show that votes decrease with distance and increase with shared artists, indicating a minor but consistent geographical and cultural bias. Our findings suggest that countries favor their neighbors in Eurovision voting.

1 Introduction

The Eurovision Song Contest is an annual international music competition that fosters unity and cultural exchange. However, concerns about voting biases persist, where countries may favor neighbors or regional allies, raising questions about the contest's fairness. This has attracted attention, with studies [1, 2, 3] investigating various political and cultural factors. However, previous studies focused on earlier contest editions, mostly before 2005. Since 2016, jury and televoting results have been reported separately, doubling the points each country distributes [4]. To address this problem, this study incorporates recent data to assess how this new data has influenced voting patterns. To better control for confounding variables, we examine shared musical tastes using Spotify data and consider GDP per capita, infrequently examined factors in Eurovision. Our analysis extends beyond general trends to explore country-specific biases, showing how geography, economy, and culture influence voting behaviors.

2 Methodology

2.1 Data Collection and Preprocessing

We obtained voting records up to 2023 from the Eurovision Dataset repository [5], focusing exclusively on votes cast during the grand final due to different voting dynamics. Geographical distance data was derived from the Natural Earth dataset [6], with distances between capitals calculated using a projected coordinate system. Capital cities and country abbreviations for each participant were obtained manually [7]. To measure cultural similarities, we collected the top 30 most popular artists and genres for each participating country using the Spotify API [8]. We applied the Dice-Sørensen coefficient [9] to quantify similarity based on the proportion of shared artists/genres, which accommodates differing list sizes. Finally, GDP per capita data was sourced from the World Development Indicators dataset by the World Bank Group [10], with minor discrepancies between the dataset and actual Eurovision participants. For preprocessing, data collection errors, one-time participants, and defunct countries were removed. Australian entries were also excluded as they would be extreme outliers for distance. Distance and GDP per capita were standardized using z-score normalization to compare effect sizes.

2.2 Clustering

We applied clustering to assess whether countries with similar features exhibit similar voting patterns. We used both linear k-means clustering and non-linear spectral clustering.

We constructed matrices for geographic distance, GDP differences, music similarity, and voting behavior. Clustering each of these matrices by country and computing the similarity of these clustering distributions would result in non-aligned labels and inaccurate results. To address this, we first clustered the geographic, economic, and music similarity matrices separately, then used the labels to categorize voting matrices. We calculated silhouette scores and visualized the categories via principal components analysis (PCA). These outcomes were compared against random clustering for validation.

2.3 Linear Regression

We used linear regression to quantify the impact of various factors on Eurovision voting. We performed a regression on total points, and using data from 2016 onward, regressions on jury and televoting points. To isolate the effect of geographic distance, we first ran the regressions using only distance. We then applied a two-step approach: first including all factors and then excluding distance to evaluate its contribution to voting behavior. We repeated analyses using the median instead of the mean to determine if the effect is due to outliers. We used significance level .05 for all statistical tests.

2.4 Linear Mixed Model

We applied a linear mixed-effects model as an alternative to multiple linear regressions (2.3), leveraging the hierarchical structure of the data. Countries casting votes were treated as groups j, with linear coefficients y_{ij} similar to those of linear regression. We used a random intercept u_{ij} to account for national biases. No random slopes were used, resulting in the following model:

$$\text{Total votes}_{ij} = (y_{00}+u_{0j})+y_{10}\cdot\text{Dist.}+y_{20}\cdot\text{GDP diff.}+y_{30}\cdot\text{Genre sim.}+y_{40}\cdot\text{Artist sim.} \quad (1)$$

We used maximum likelihood estimation which allows model evaluation with the Akaike Information Criterion (AIC) [11]. We also ran a model excluding distance to assess the impact of distance.

3 Results

3.1 Clustering

We set the number of clusters $k = 3, 4, 5, 6$. Only the results of k-means ($k = 4$) were reported, while the full clustering diagrams and scores are available in the notebook.

All silhouette scores for voting matrices using k-means or spectral clustering were under 0.01, similar to results of random clustering, indicating no strong correlation between voting patterns and distance, GDP, or music preferences. The scatter plot for music tastes (Fig. 1) reveals a linear pattern, suggesting that regression may be more suitable for this data. Clustering is effective when there are distinct separations. However, our data do not show clear separations but are rather uniformly distributed across different principal component areas, leading to suboptimal clustering performance.

3.2 Overall Linear Regression

We first examined the relationship between voting scores and distance. The Pearson correlation between mean total votes and distance was -0.2 (Fig. 2c), with jury and televoting correlations of

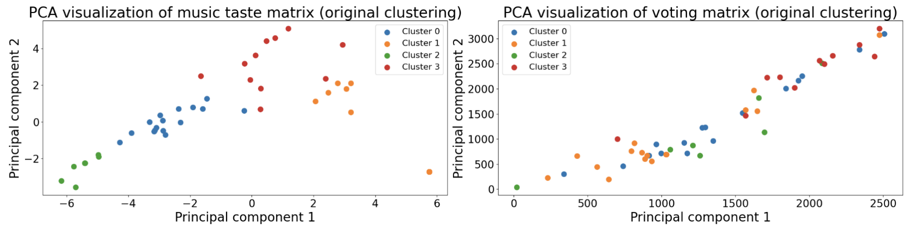
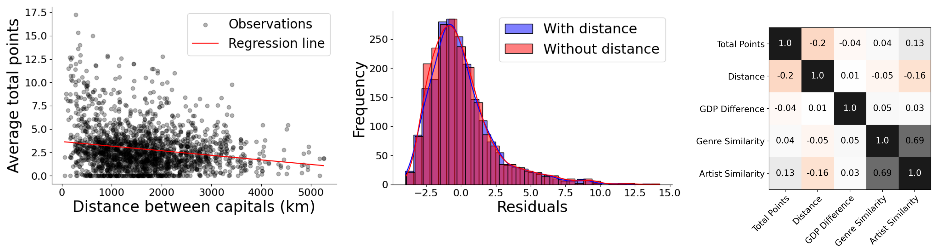


Figure 1: Clustering example for artist similarity (k-means, $k = 4$)



(a) Observed values vs. distance (b) Residuals distribution (c) Correlation of means

Figure 2: Linear regression results

-0.09 and -0.23, respectively. Linear regression models indicated a significant negative relationship ($p < .001$). To test whether this effect stems from cultural or economic factors, we compared models with and without distance using a partial F-test. In all cases, distance remained a significant predictor ($p < .001$ for total points and televoting; $p = .002$ for jury votes), indicating the effect persists despite other factors (Fig. 2a). For total points, scores decreased by 0.45 (± 0.11) per standard deviation of distance (919 km), with a stronger effect in televoting (-0.55 ± 0.12) than jury voting (-0.18 ± 0.11), suggesting a relationship between distance and Eurovision votes.

Model evaluation showed limited predictive power, with only 5.5% of variance explained. Residual analysis (Fig. 2b) revealed non-normal residuals for both models. A Q-Q-plot (Fig. 3b) shows similarly non-normal residuals. To check for heteroscedasticity, we applied a Breusch–Pagan test [12] and found significant results ($p < .001$).

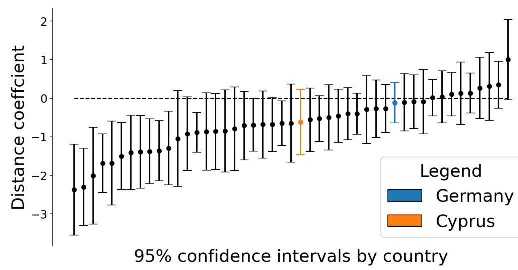
3.3 Country-wise Linear Regression

We examined the relationship between distance and voting at the country level. Although the overall correlation was negative, the effects of the individual countries varied, including some positive relationships (Fig. 3a). After applying Bonferroni Correction, only four countries showed significant negative effects when regressing only on distance. When including all variables, only two countries retained significance for distance. Televoting results were similar, but no country showed a significant effect in jury voting.

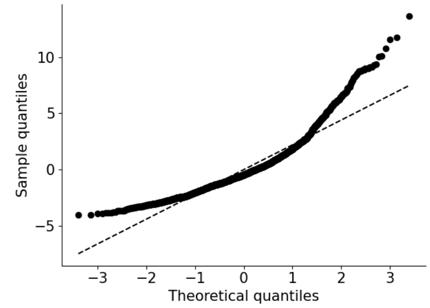
3.4 Linear Mixed Models

After applying linear mixed models, the results showed that distance had a significant negative effect on total points awarded ($-0.504, p < .001$), consistent with previous findings. artist similarity had the largest positive effect ($1.152, p < .001$), while GDP difference had a significant small effect ($0.082, p = .003$). However, genre similarity was not significant ($p = .529$). To assess the importance of distance, we ran a model excluding it and another one only regressing on distance. Both resulted in a higher AIC, making them significantly less likely (probability $< .001$) to minimize information loss compared to the full model. Models for televotes and jury votes followed similar trends.

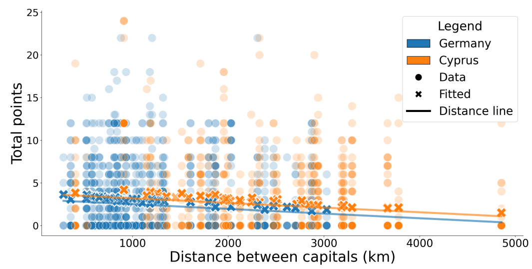
Despite these insights, the model does not fit well and underestimates total points awarded (Fig. 3c). Additionally, the Q-Q plot (Fig. 3d) suggests residuals deviate from normality.



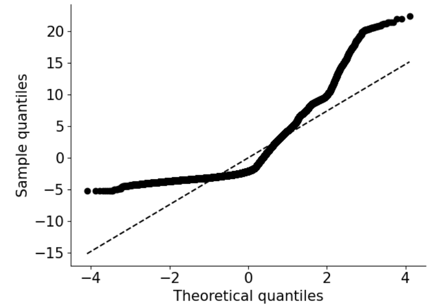
(a) Country-wise coefficients



(b) Q-Q plot for the overall model



(c) Actual vs. predicted votes for Germany and Cyprus



(d) Q-Q plot residuals

Figure 3: Linear regression coefficients and model validation (above). Linear effects results (below).

4 Discussion

One key limitation is that distance may confound with cultural or political factors. Although we controlled for some variables, the choice of included factors impacts the effect attributed to distance. Another issue is model assumptions. Eurovision votes are discrete, bounded, and not independent since it is a ranking-based system, violating linear regression assumptions as can be seen in the diagnostic plots (Figs. 3b, 3d). This does not necessarily invalidate results. Changes in Eurovision's voting system may have led to underestimation. We used recent music data, which may not historically reflect preference similarity. Additionally, using capital-to-capital distances may not fully capture population interactions. Similarly, religion was initially included as a variable, showing consistent results in the mixed model, but it was dropped due to data inconsistencies. Further, this report only captures a subset of the analysis, making it susceptible to reporting bias.

Further studies might examine additional factors, such as voting in intergovernmental organizations or immigrant populations. One could also model nonlinear relationships or changes over time to provide deeper insight into long-term voting trends.

We found a small but significant negative relationship between distance and votes, which persists even when accounting for music preferences, as observed in 3.2 and 3.4. Across all types of voting, the mean votes decrease with distance while increasing with shared artists, though individual country effects are rarely significant. These findings indicate that geographic proximity influences Eurovision voting, but other factors, such as cultural and political ties, likely play a larger role. Despite these patterns, Eurovision remains an unpredictable and dynamic contest, continuing to celebrate musical diversity across Europe.

5 Statement of Contributions

Kai led initial coordination, collected Eurovision data, and handled preprocessing. Valentin gathered distance and economic data and ran the linear mixed model. Pablo acquired music data, conducted the linear regression analysis, and later coordinated. Keyu ran the clustering analysis. Kai and Pablo analyzed regression results, while all members contributed to writing, editing, and visualization.

References

[1] Alexander V. Mantzaris, Raphael Rein, and Nicholas Hopkins. Examining collusion and voting biases between countries during the eurovision song contest since 1957. *Journal of Artificial Societies and Social Simulation*, 21(4):6, 2018.

[2] Derek Gatherer. Comparison of eurovision song contest simulation with actual results reveals shifting patterns of collusive voting alliances. *Journal of Artificial Societies and Social Simulation*, 9(2):1, 2006.

[3] Victor Ginsburgh and Abdul G. Noury. The eurovision song contest: Is voting political or cultural? *European Journal of Political Economy*, 24(1):41–52, 2008.

[4] European Broadcasting Union. Eurovision song contest voting rules, 2024. Accessed: 2025-01-31.

[5] Marleen Spijkervet. Eurovision dataset, 2021. Accessed: 08-11-2024.

[6] Natural Earth. Natural earth data, 2024. Accessed: 21-11-2024.

[7] International Organization for Standardization. Online browsing platform (OBP). Accessed: 05-02-2025.

[8] Spotify Developers. Spotify web API documentation, 2024. Accessed: 10-01-2025.

[9] Thorvald Sørensen. A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on danish commons. *Kongelige Danske Videnskabernes Selskab*, 5(4):1–34, 1948.

[10] World Bank Group. World development indicators, 2024. Accessed: 25-01-2025.

[11] Hirotugu Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19:716–723, 1974.

[12] T. S. Breusch and A. R. Pagan. A simple test for heteroscedasticity and random coefficient variation. *Econometrica*, 47(5):1287–1294, 1979.