Collection and Preparation of Data

Immigration Data: *Annual crude rate of net migration for each country in the European Union*

To measure immigration, I used annual crude rate of net migration to Europe ((number of immigrants – migrants) / population). I decided to use ratio values rather than actual numbers to eliminate biases such as global population increase that affects both number of immigrants and number of votes of each party. I used "Population change - Demographic balance and crude rates at national level" data provided by official site of EU, Eurostat. After accesing to annual crude rate of net migration to European countries, I mannually filtered the European Union countries and eliminated statistical abbrevations by using Excel tools on the spreadsheet.

(https://ec.europa.eu/eurostat/databrowser/view/demo_gind/default/table?lang=en)

Nationalism Data: *National Election vote shares for each EU country's nationalist party*

To measure nationalism, I used each country's most recent and most popular nationalist party's vote shares from two recent elections. First I searched the most recent and most popular nationalist party of all 27 countries in the European Union. After adding the party abbrevations for each country to my spreadsheet, I listed the last national election that each party participated in along with the previous national election. Then I listed the vote shares of each party for both elections. When searching for parties and vote shares, I used ParlGov website (database created for social research purpose) and official election authority websites of the respective countries, and I double-checked the data.

(https://parlgov.fly.dev/)

Attribute Descriptions:

prev_year: year of the national election prior to the most recent national election that each party participated in

prev_vote: vote share of the attributed party in the national election prior to the most recent national election

last_year: year of the last (most recent) national election that each party participated in

last_vote: vote share of the attributed party in the most recent national election

---

Explanatory Data Analysis

Descriptive Statistics:

The analysis began with basic descriptive statistics on the vote share dataset. This helped summarize the central tendencies (mean, median) and variability (standard deviation) of vote shares in both the previous and most recent elections. These statistics provided an initial look into how much vote shares have shifted and whether there was a general increase in nationalist sentiment across countries.

Data Preparation and Transformation:

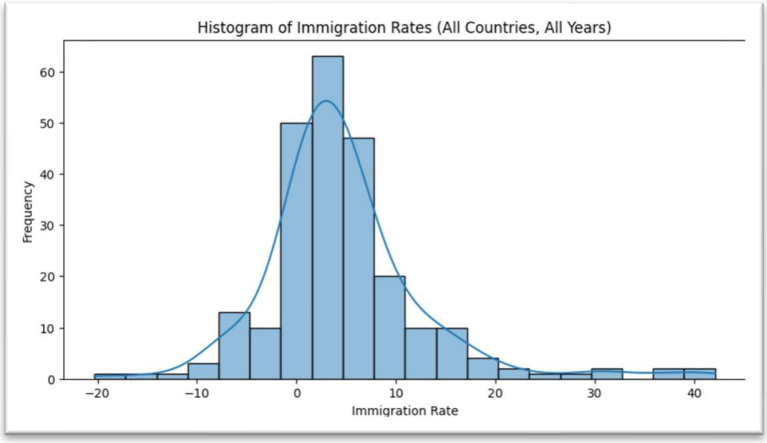
For proper time-series and comparative analysis, the immigration dataset was transformed from a wide format to a long format, enabling easier plotting and year-wise calculations. The vote share dataset was also enhanced by calculating a new variable—vote share change—representing the difference between the most recent and previous elections. This was done for each country, giving a clearer picture of which nations experienced growth or decline in nationalist party support.

To quantify immigration change over time, the average yearly trend was calculated per country. This value serves as a proxy for how rapidly immigration has increased or decreased in recent years.

Exploratory Visualizations

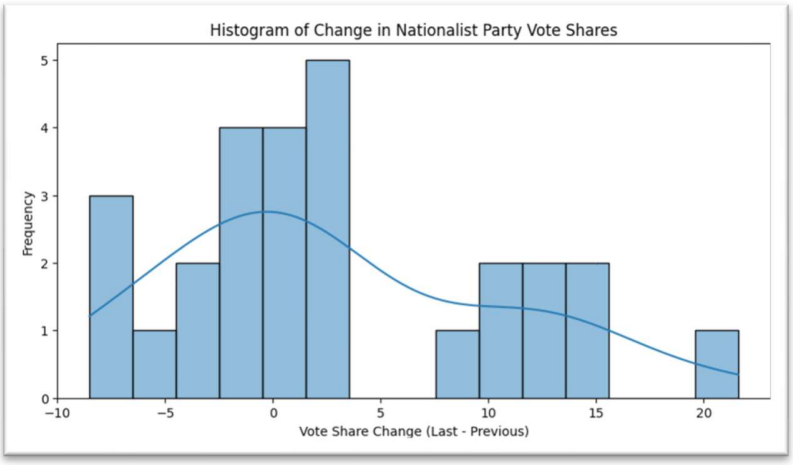
Histograms:

A set of histograms was created to explore the distribution of both immigration rates and vote share changes. The histogram for immigration rates showed that while most countries had moderate values, a few countries experienced either very high or very low migration trends, indicating significant demographic movement. The histogram for vote share change showed that some countries had notable increases in nationalist support, while others saw declines or stability.



Interpretation:

- Most countries cluster around modest immigration rates (0–10).
- Some countries have negative values (more emigration than immigration).
- A few countries show unusually high positive rates, indicating strong inbound migration.

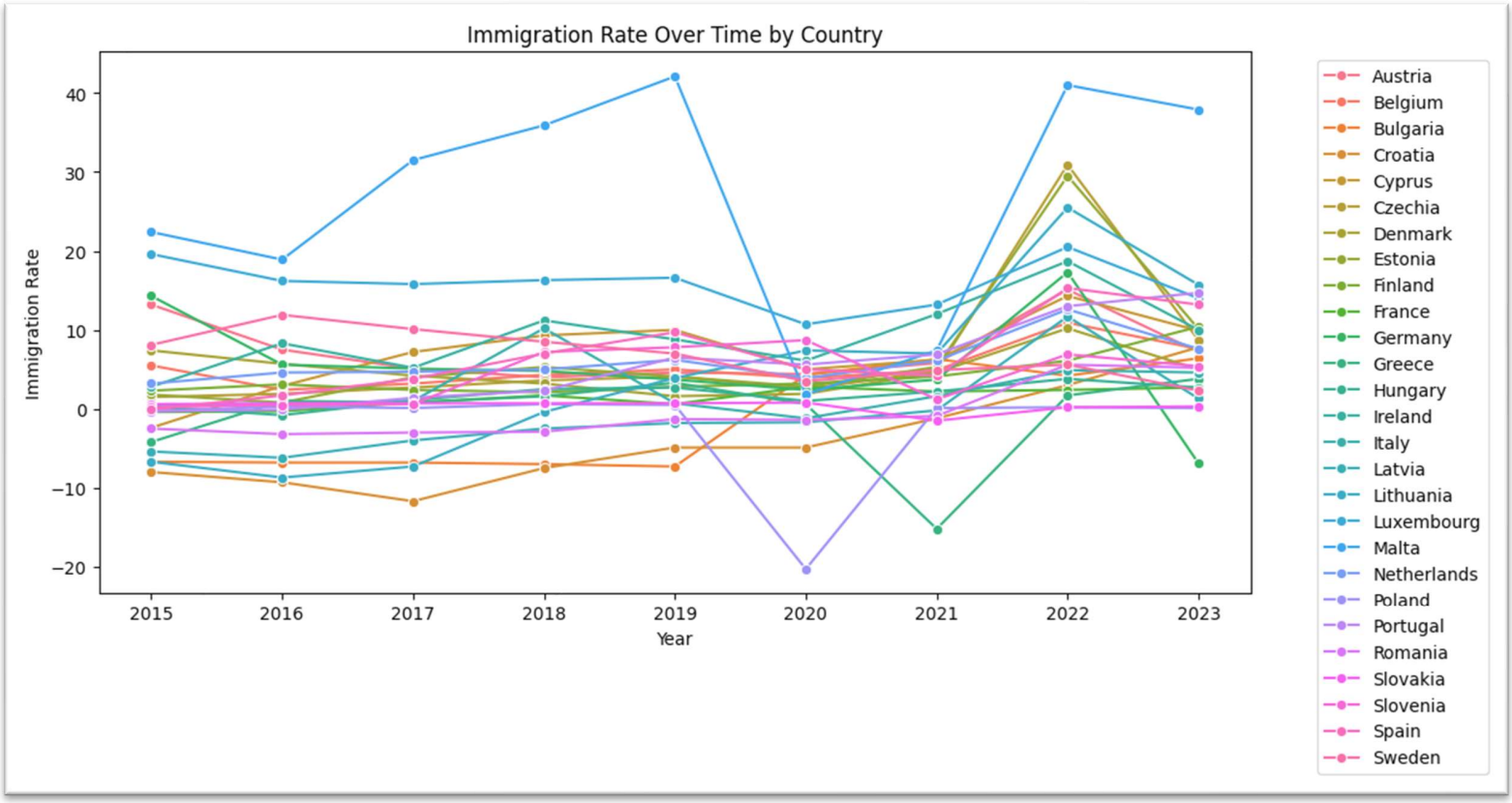


Interpretation:

- Many countries experienced modest increases or stability in vote shares.
- Several countries saw notable spikes in nationalist support.
- A few experienced a decline in nationalist vote shares, showing variation across Europe.

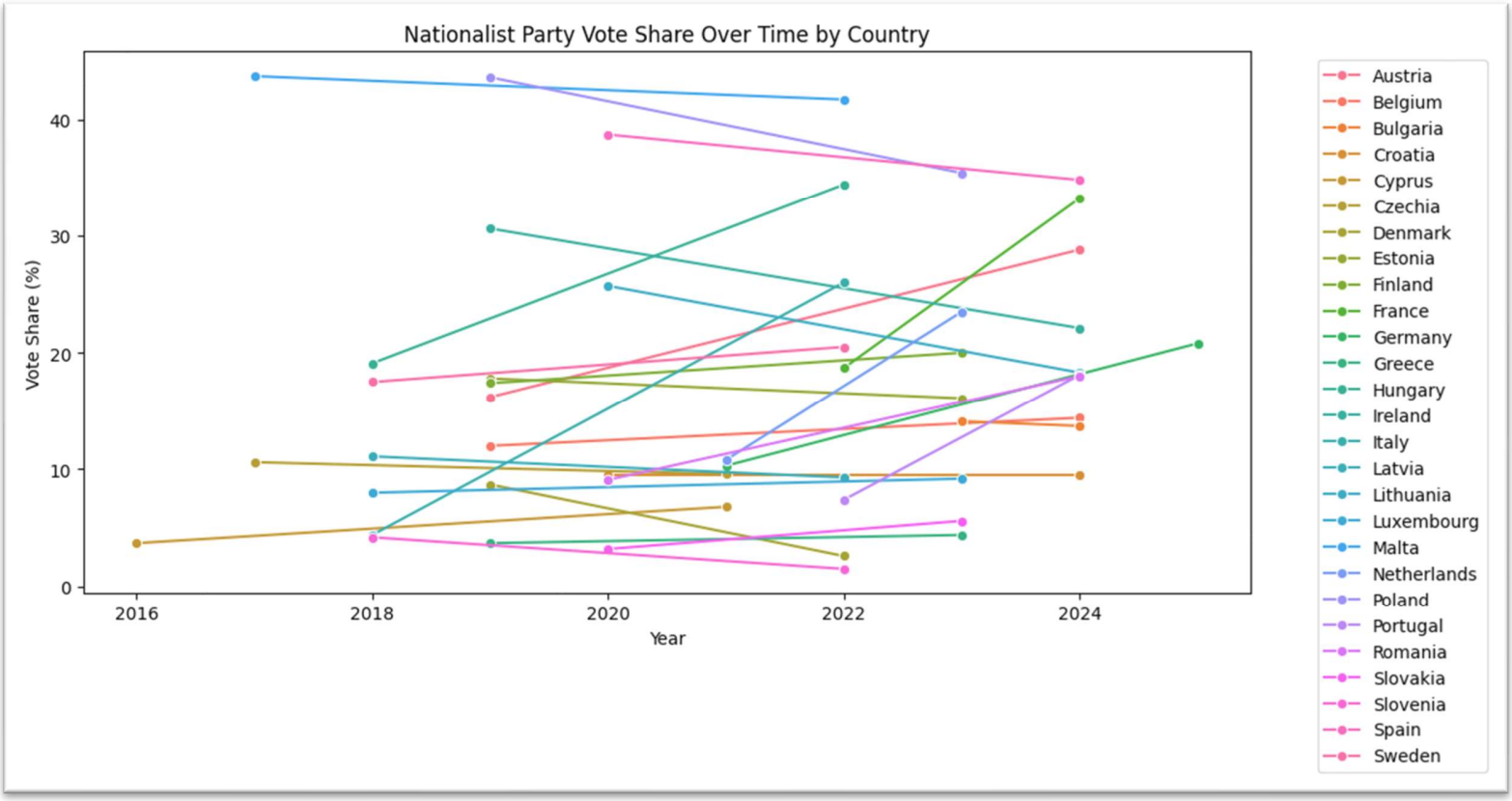
Time Series:

The time-series analysis was crucial for contextualizing the hypothesis. It helped visually identify countries where both immigration and nationalist votes increased in parallel. Although time-series data alone doesn't prove causality, observing these parallel trends in multiple cases supports the rationale behind further hypothesis testing. For example, in countries like Hungary or Poland, visual analysis showed a concurrent rise in both variables, hinting at a possible relationship that could be statistically tested. Time-series plots not only revealed the temporal dynamics of immigration and political sentiment but also guided which countries might be driving the overall trends observed in correlation and regression analysis.



Interpretation:

- Immigration patterns vary significantly by country.
- Some show steady growth (e.g., Germany, Spain), while others fluctuate or decline.
- Parallel spikes in recent years in several countries hint at recent demographic shifts

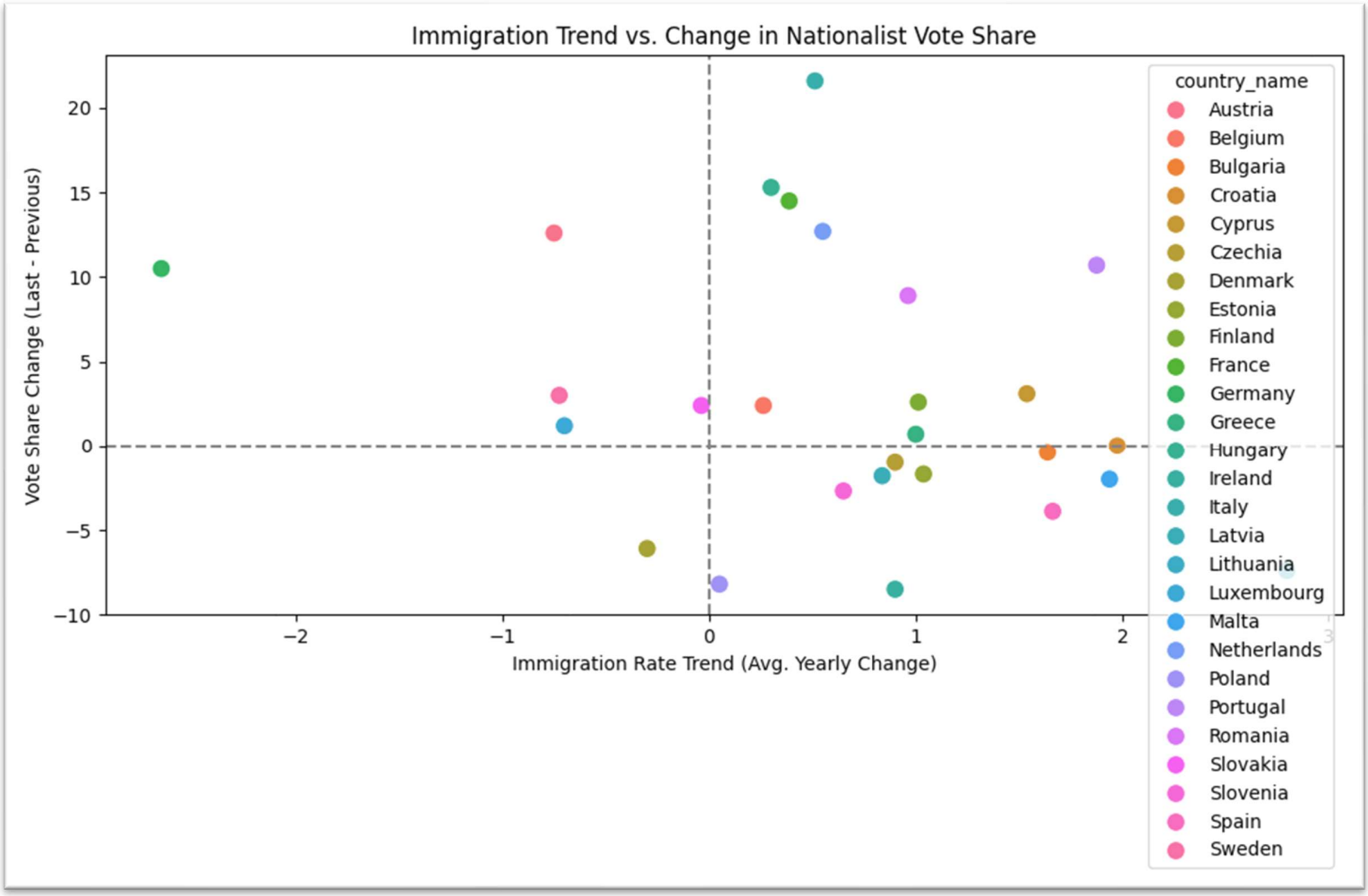


Interpretation:

- Many countries show increases in nationalist party support over the last election cycle.
- A few countries like Poland and Hungary exhibit steep rises, aligning with the hypothesis.
- Others (e.g., Sweden, Malta) demonstrate relative stability or decline.
-

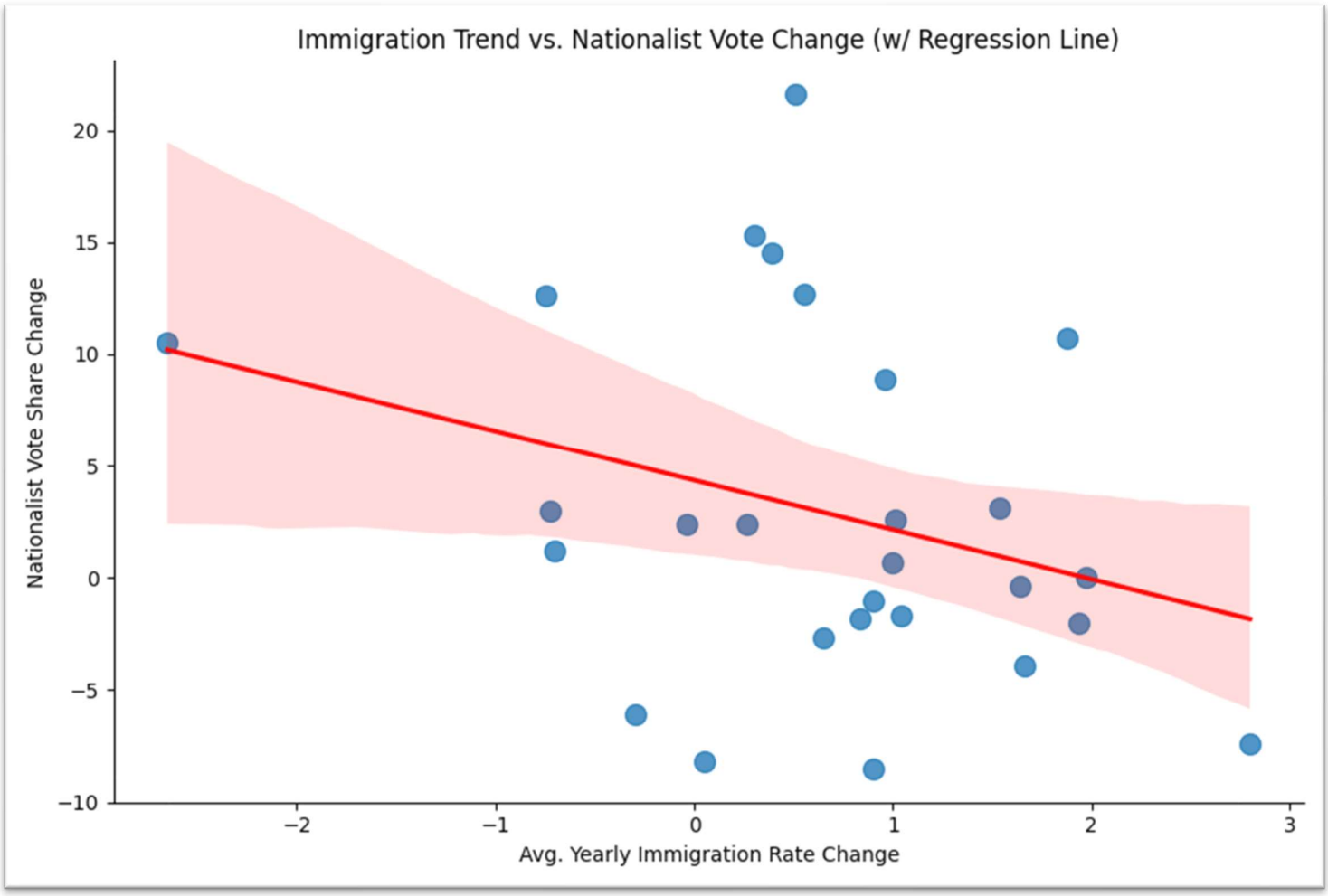
Scatterplot:

Scatterplots were used to visually assess the relationship between the average immigration trend and the change in nationalist party vote shares. Each point represented a country, with its horizontal position determined by the immigration trend and the vertical position by vote share change. The distribution of points suggested a potential—but not visually overwhelming—pattern of positive association. To enhance interpretability, the scatterplot was supplemented with a trend line and marginal histograms in a joint plot. This allowed for a more holistic look at the spread and potential linear relationship between the variables.



Interpretation:

- Many countries show increases in nationalist party support over the last election cycle.
- A few countries like Poland and Hungary exhibit steep rises, aligning with the hypothesis.
- Others (e.g., Sweden, Malta) demonstrate relative stability or decline.

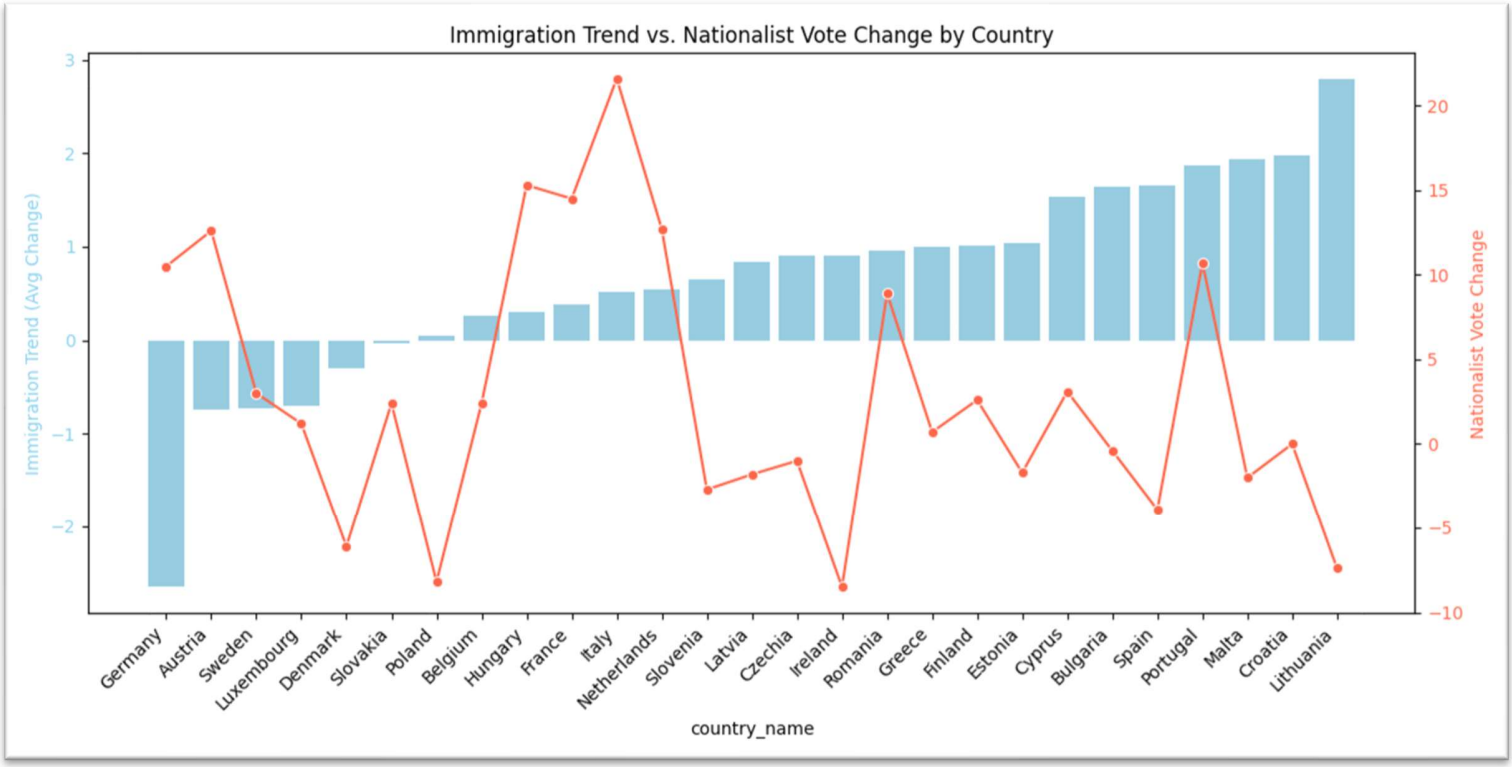


Interpretation:

- The regression line indicates a **slightly negative** trend.
- The confidence band (shaded area) is wide, suggesting uncertainty and low predictive power.
- The result visually questions the direct causality hypothesis.

Bar-line chart:

To overcome the complexity and clutter of traditional line plots with overlapping lines, more focused visuals were created. A grouped bar-line chart was developed where immigration trends were shown as bars and vote changes were shown as a line chart overlaid on the same x-axis of countries. This dual representation helped contextualize both metrics simultaneously for each country.



Interpretation:

- Bars show average yearly immigration change; lines show nationalist vote share change.
- Some countries like Italy, France, and Hungary stand out with both high immigration and vote increases.
- Others (e.g., Germany, Sweden) show inverse relationships.
- Visually effective at comparing both dimensions side by side, highlighting outliers and inconsistencies.

Hypothesis Testing

Pearson Correlation Analysis Interpretation:

The Pearson correlation test yielded a **correlation coefficient of -0.31** and a **p-value of 0.116**. This indicates a weak negative relationship, and importantly, the result is not statistically significant at the conventional 0.05 level.

As such, the analysis does not provide sufficient evidence to reject the null hypothesis. There is no statistically reliable indication that increases in immigration are linearly associated with increases in nationalist support.

This outcome suggests that nationalist voting behavior cannot be directly inferred from immigration trends alone.

Linear Regression Analysis Interpretation

To further quantify the relationship, an ordinary least squares (OLS) regression was conducted. The resulting **coefficient for immigration trend was -2.21**, and the associated **p-value was again 0.116**.

The **R^2 value of 0.096** suggests that immigration trends explain less than 10% of the variance in vote share changes, indicating very limited explanatory power.

These results reinforce the earlier conclusion: there is no significant linear relationship between the rate of immigration change and the growth in nationalist vote share.

Results and Discussion:

The hypothesis that increased immigration correlates with increased nationalist voting was not supported by the data. Both correlation and regression analyses failed to demonstrate a statistically significant relationship.

This project aimed to examine whether increasing immigration in European Union (EU) countries is associated with a rise in support for nationalist parties. The hypothesis was motivated by widely observed political narratives and recent developments across the continent. In contemporary discourse, immigration is frequently positioned as a driver of nationalist sentiment, particularly during election periods.

The assumption underlying the hypothesis was intuitive: higher immigration rates might correlate with increased nationalist vote shares. However, the data provided a more nuanced perspective.

Although the hypothesis was not supported, the findings are both insightful and instructive. It is plausible that immigration may influence nationalism in some contexts, but its impact is not uniform or easily isolated.

Other social, political, and economic factors — such as unemployment, media influence, populist leadership, or historical context — may play a more critical role in shaping nationalist sentiment.

The lack of statistical significance in this case may also reflect the limitations of using a simple linear model to capture what is likely a complex and multifaceted relationship.