

Lok Sabha Election - 2024 Analysis

By Rashad Mohammad, 4th Year CSE (AIML) Student, SRM University

Introduction

This project focuses on analyzing data extracted from the official website of the Election Commission of India at <https://results.eci.gov.in/>. The dataset pertains to the Lok Sabha election results, offering a comprehensive view of the electoral outcomes across India.

Using Python for data analysis and Power BI for visualization, this project aims to uncover significant insights from the election data. Analyzing these results is crucial for understanding voter preferences, evaluating political party performance, and identifying demographic and regional voting patterns.

By examining which parties performed best and worst in specific constituencies, this analysis aims to reveal underlying socio-political dynamics, including regional influences and demographic preferences.

Dataset Overview

The dataset used for this analysis is sourced from the official website of the Election Commission of India (<https://results.eci.gov.in/>). It contains comprehensive results of the 2024 Lok Sabha elections across various constituencies in India.

Structure and Key Columns:

The dataset is structured with the following key columns:

Constituency: Name of the electoral constituency.

Leading Candidate: Name of the candidate leading in votes.

Leading Party: Political party of the leading candidate.

Trailing Candidate: Name of the candidate trailing in votes.

Trailing Party: Political party of the trailing candidate.

Margin: Number of votes by which the leading candidate is ahead of the trailing candidate.

Status: Election status (e.g., contested, uncontested).

Votes: Total number of votes received by the leading and trailing candidates.

Preprocessing Steps

Before analysis, the dataset underwent the following preprocessing steps:

Handling Missing Values: Checked for and handled missing values, particularly noting instances where missing values corresponded to uncontested election statuses, which were retained.

Data Cleaning: Ensured data integrity by removing any duplicate entries and verifying consistency across columns such as party affiliations and electoral results.

Data Conversion: Converted relevant columns (e.g., 'Margin') to numeric format for accurate numerical analysis and visualization.

Categorical Analysis: Conducted exploratory data analysis (EDA) to understand the distribution of categorical variables like leading parties and election statuses

For the dataset, we encountered NaN values specifically in the constituency of Surat, which were retained due to the 'UNCONSENTED' status of the election results.

Exploratory Data Analysis

Descriptive Statistics:

Descriptive statistics are numerical measures that summarize and describe the main features of a dataset. They provide insights into the central tendency, variability, and distribution of data. Common descriptive statistics include measures like mean, median, mode, range, standard deviation, and percentiles. These statistics help analysts and researchers to understand the basic characteristics of the data, identify outliers, and make initial interpretations.

Mean: Average value of a numerical dataset.

Median: Middle value that separates the higher half from the lower half of a dataset.

Standard Deviation: Measure of the amount of variation or dispersion of a set of values.

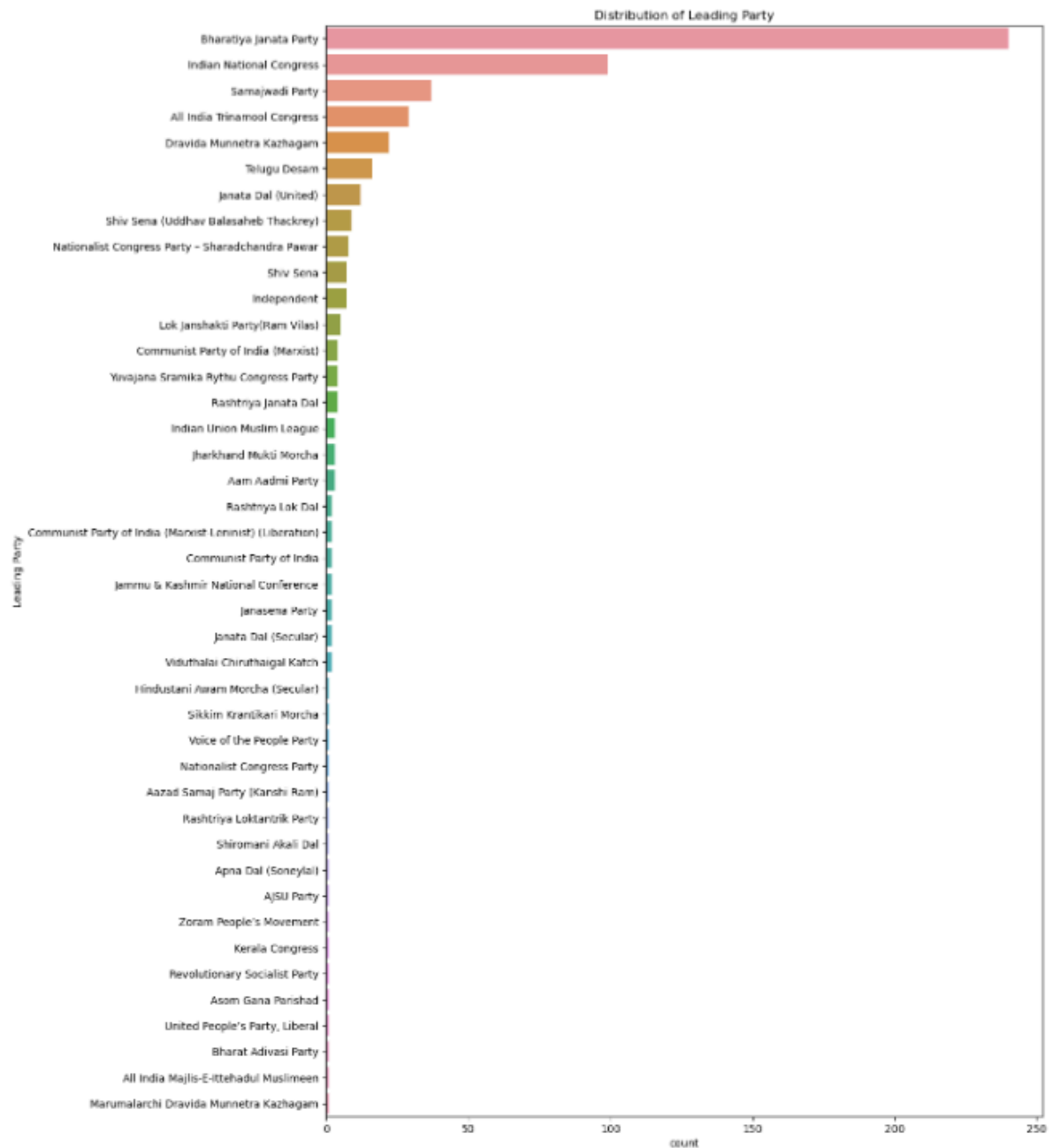
Range: Difference between the maximum and minimum values in a dataset.

Percentiles: Values below which a given percentage of observations in a group of observations falls.

	Const. No.
count	543.000000
mean	18.178637
std	16.427840
min	1.000000
25%	6.000000
50%	14.000000
75%	25.000000
max	80.000000

Categorical Analysis:

Categorical analysis involves examining categorical variables within a dataset to understand the distribution and relationships among different categories. Categorical variables are qualitative variables that can take on a limited, fixed number of possible values



Findings:

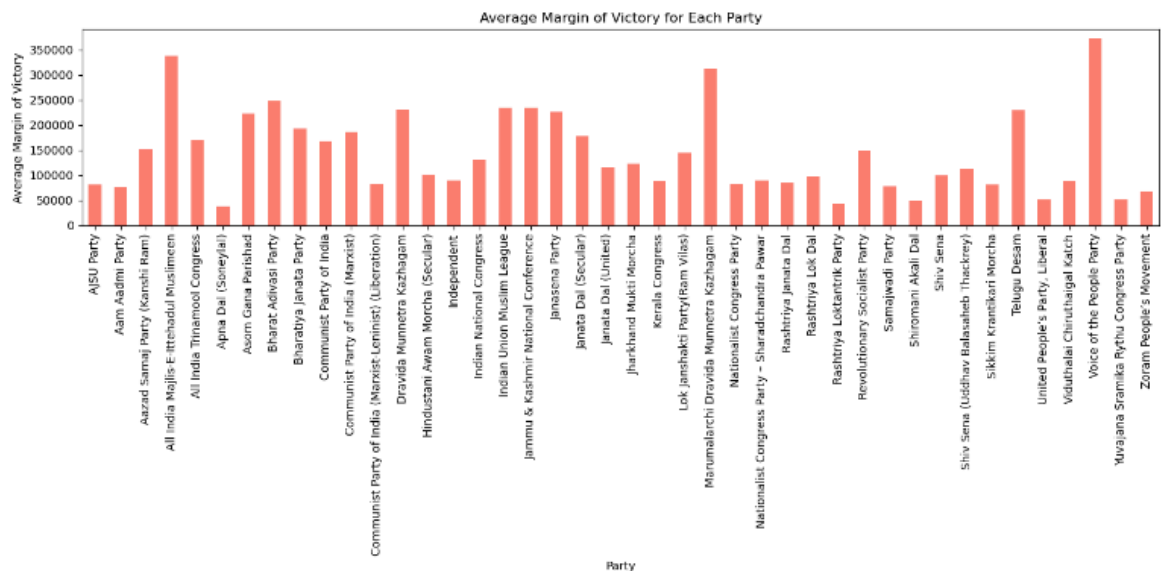
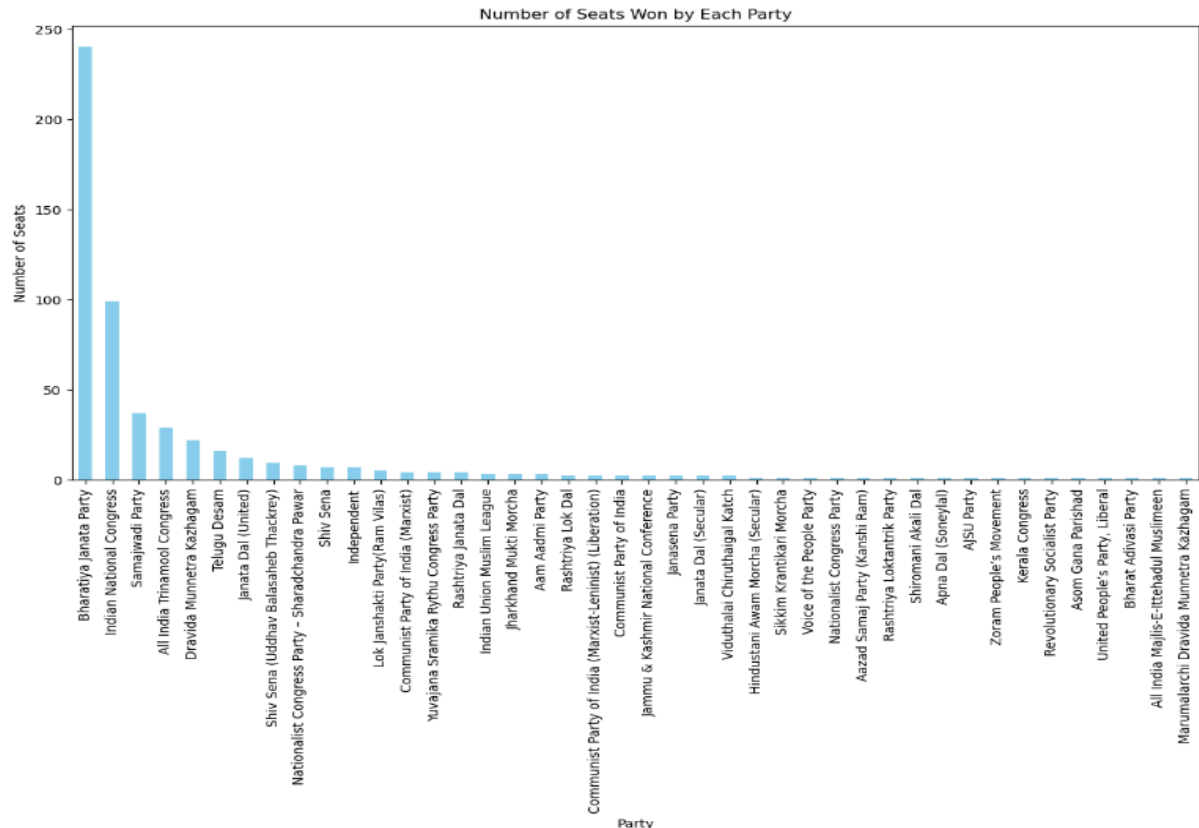
Toughest Competition

The toughest competition in the 2024 Lok Sabha elections was observed in several constituencies:

Mumbai North West had the smallest margin of just 48 votes between the leading and trailing candidates.

Other closely contested constituencies include Attingal (684 votes), Jajpur (1587 votes), and Jaipur Rural (1615 votes).

Number of Seats Won By Each Party & Average Margin of Victory for Each Party:



Top Regions with Highest Avg Margins for Top 3 Parties

Voice of the People Party: Won by 371,910 votes in Shillong.

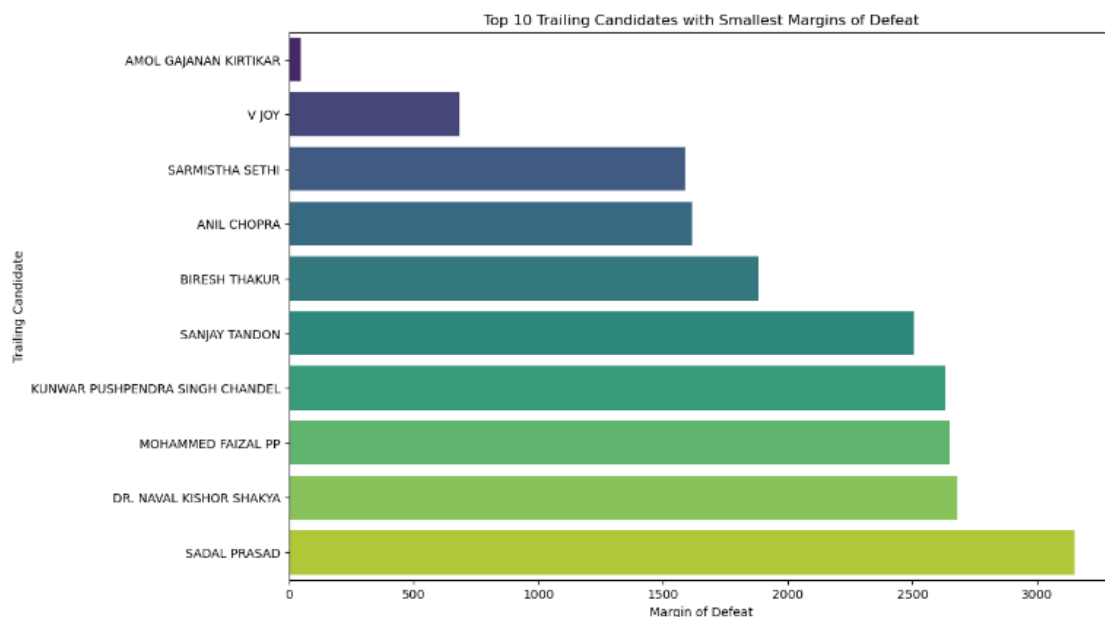
All India Majlis-E-Ittehadul Muslimeen: Secured victory by 338,087 votes in Hyderabad.

Marumalarchi Dravida Munnetra Kazhagam: Achieved a margin of 313,094 votes in Tiruchirappalli.

Constituency with the highest margin

Indore: Bharatiya Janata Party won with a significant margin of 1,175,092 votes, indicating strong local support. People were overwhelmingly supportive of the Bharatiya Janata Party, leaving little chance for other parties to win.

Trailing Candidates



Identified as potential future contenders due to their strong performances despite narrow margins of defeat.

Potential Future Contenders:

Trailing candidates identified in this analysis have shown strong performances despite their losses, indicating they could be significant contenders in future elections. Their narrow margins of defeat suggest a robust voter base and high potential for future success. These candidates' competitive results highlight their potential influence in upcoming elections.

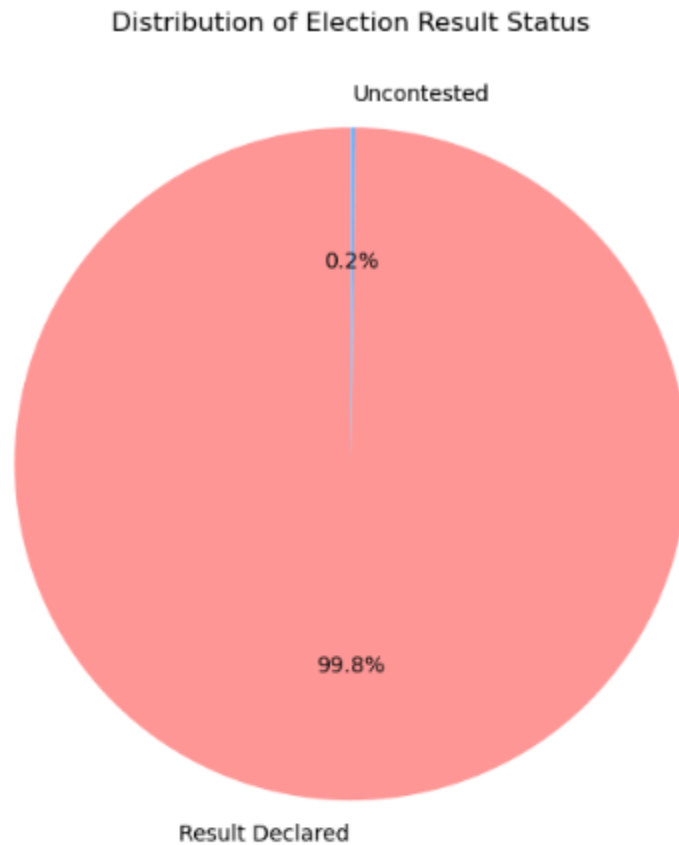
Party Dynamics:

The analysis reveals intriguing dynamics within and between political parties. Intra-party competitions, such as those within Shiv Sena, demonstrate competitive races among candidates from the same party. Furthermore, close contests between major parties like the Bharatiya Janata Party (BJP) and the Indian National Congress (INC) underscore intense

rivalries and shifting voter preferences, providing insights into the evolving political landscape.

Status Distribution

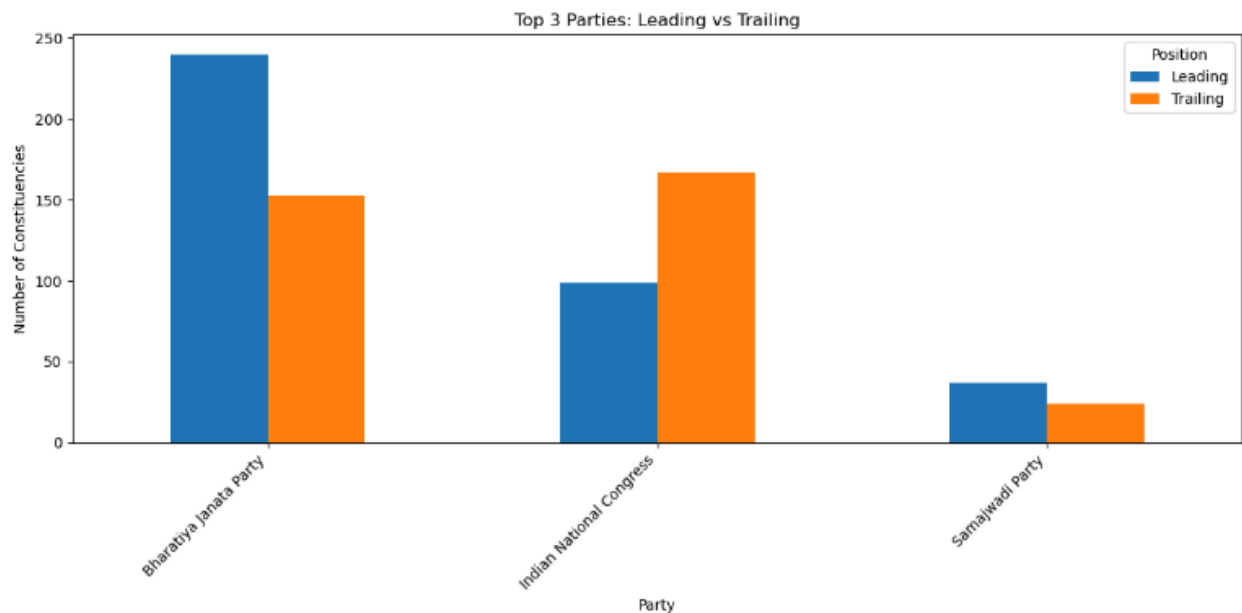
The constituency of Surat, representing 0.2% of the results, remained uncontested.



```
Status Distribution:  
Result Declared    542  
Uncontested         1  
Name: Status, dtype: int64
```

The 0.2% shown in the graph represents Surat, where the status is uncontested.

Political Performance



This analysis illustrates the performance of the top 3 parties, detailing where they led and trailed in constituencies. Specifically, the BJP secured victory in 240 constituencies.

Party Performance by Margin Ranges

To evaluate electoral performance, we categorized election outcomes into distinct margin ranges:

Very Close: Elections won by a very narrow margin, indicating strong competition and potential volatility in future elections.

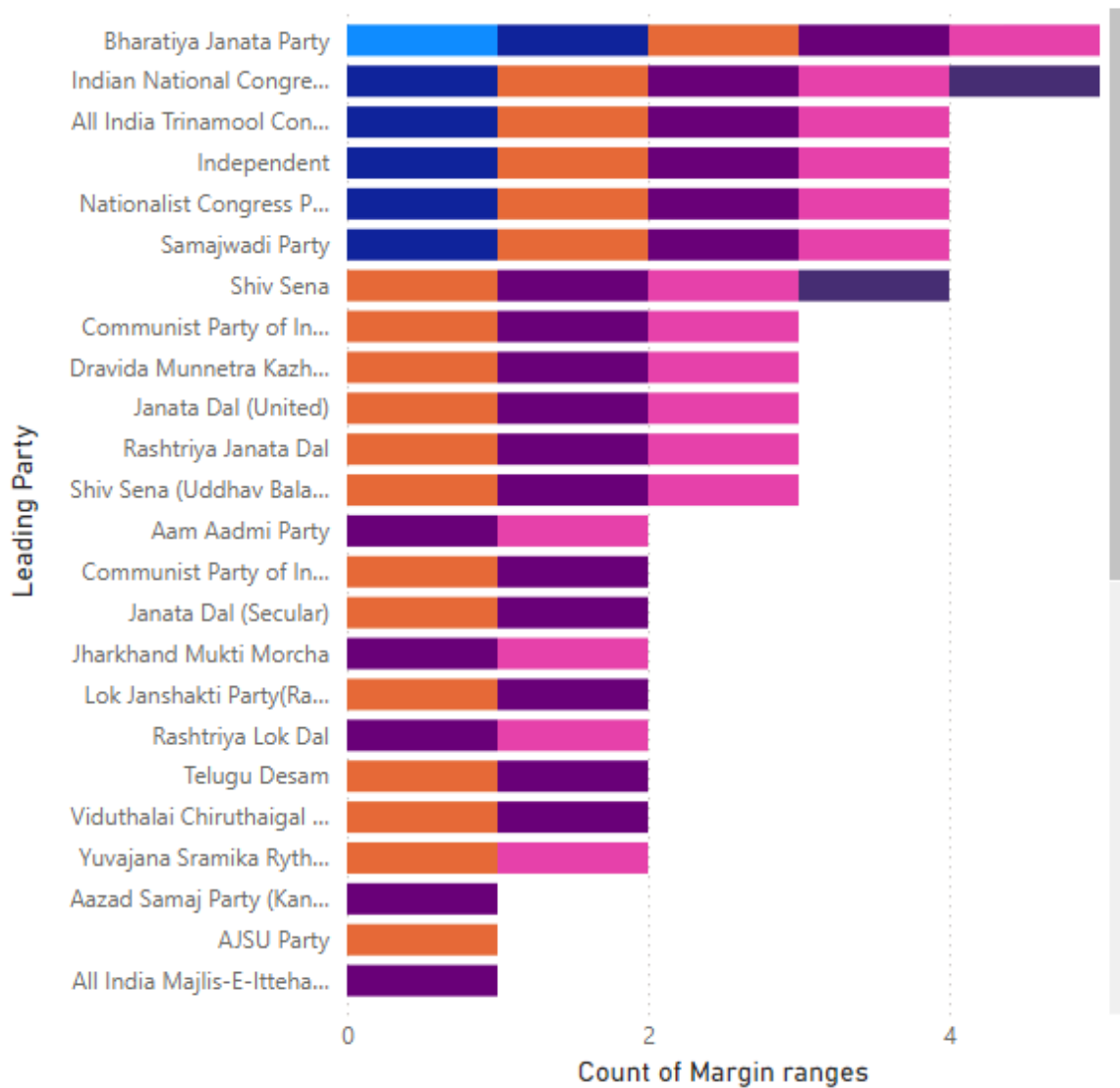
Close to Moderate: Contests won by a relatively close margin, reflecting competitive races with a clearer advantage.

Moderate to Comfortable: Elections where the margin is moderate, suggesting a safer margin but still competitive.

Comfortable to Landslide: Reflects elections won by a comfortable to a very large margin, indicating significant voter support and potential dominance in that constituency.

Parties Performance by Margin Ranges

Margin ranges (Blank) Close Comfortable Landslide Moderate Very Close



These are some of my findings, and I will provide the Python notebook and Power BI dashboards showcasing these insights on my GitHub link: [LINK](https://github.com/rashadmd/Lok-Sabha-Election-2024-Analysis) (<https://github.com/rashadmd/Lok-Sabha-Election-2024-Analysis>)