LOK SABHA ELECTIONS ANALYSIS REPORT

By Shatakshi R

Project link:
https://github.com/shatakshii08/LokSabha-Election-WebScraping

Table of Contents

Introduction

In this report, we analyze the recently concluded Lok Sabha election results to derive meaningful insights. The analysis aims to understand the election outcomes and trends over the years, focusing on overall results, state-specific performance, and historical trends.

Web Scraping was done to collect the data using the BeautifulSoups() module in Python and stored in SQL. Data cleaning/Preprocessing and Data Analysis was also done using SQL. This data was imported as .csv file into PowerBI and insightful dashboards were created using PowerBI.

Tools Used

Various Data Analysis Tools were used. I tried to incorporate as many tools as I could to provide a detailed analysis and interpretation. The tools used are:

- Python (Jupyter Notebook)
- SQL Server
- Excel

- PowerBI

Data Collection/Web Scraping

- The data was scraped from the official Election Commission of India website (https://results.eci.gov.in) using a Jupyter Notebook.
- The data includes overall results and detailed results for 10 states.
- Python libraries like beautifulSoups() were used to scrap the data from the website.
- DataFrame function of the Pandas library were used to convert the scraped data into a neat two-dimensional table.
- The scrapped table was stored as a csv file.
- The overall Lok Sabha Elections result and the results of the following 10 highly populated states were scraped.
    - Tamil Nadu
    - Andhra Pradesh
    - Madhya Pradesh
    - Bihar
    - Gujarat
    - Karnataka
    - West Bengal
    - Rajashthan
    - Gujarat
    - Maharashthra

Data Storage

The scrapped data of the Overall Results and the ten states was stored as .csv files. The .csv files were imported into SQL. The files were stored as databases in SQL and further analysis was performed on it.

Data Preprocessing and Cleaning

The scraped data was converted into CSV files and stored in a SQL database. The following steps were taken to clean and preprocess the data:

- Identifying Duplicate Values
- Identifying Removal Values

Data Analysis

The analysis was performed using SQL queries to explore the data and derive insights.

- First, a table called ‘Historical Data’ was created for the individual party and its alliances and their vote share for the past 15 years.
- The percentage was calculated for the parties.
- The highly populated states were also analyzed using SQL queries.

Key Insights

Based on the analysis, the following 10 key insights were derived:

- **Overall Seat Distribution:** The winning party and the distribution of seats among the major parties.
- **Performance in Key States:** Analysis of the performance of major parties in Madhya Pradesh and Tamil Nadu, highlighting the changes compared to previous elections
- **Previous Years Trends:** The table with previous years trends of the alliances(NDA, I.N.D.I.A,etc) was created and analysed. How the parties performed in the last 15 years. The parties' growth and decline were derived (2014, 2019, 2024)

- The vote percentage of the alliances was derived. NDA and INDIA alliances differ only by 6% vote share in 2024 elections.
- **Government Formation:** The Party/Alliance which has more than 272 seats were analysed.
- Winning Margins: Examination of winning margins and identification of closely contested constituencies.
- **Party-wise Vote Share:** Analysis of the vote share percentage for major parties in the overall results and key states.
- The vote percentage of each party was derived as a separate column. We can infer that a regional party like Samajwadi Party alone has 6% vote share.
- In TamilNadu, we can infer that DMK and its alliances have had a clean sweep of a whopping 100%.
- Likewise, in Madhya Pradesh, BJP had an impressive clean sweep of 100% votes.
- **Regional Strongholds**: Identification of regional strongholds for major parties based on seat distribution and vote share.
- **Historical Trends:** Analysis of the historical performance of major parties over the last three elections, identifying trends and patterns. .

Data Visualization

Data Visualization was done using PowerBI. The .csv files in SQL were imported into PowerBI and three dashboards were created to visualize the insights derived from the data:

1. **Overall Election Results:**
    - A comprehensive view of the overall results, including seat distribution and vote share.
2. **State-Specific Performance:**
    - Detailed analysis of the performance of major parties in Madhya Pradesh and Tamil Nadu.
3. **Historical Election Data:**
    - Visualization of historical trends in the Lok Sabha elections for 2014, 2019, and 2024.

Conclusion

This analysis provides a detailed understanding of the Lok Sabha election results, highlighting key trends and insights. The combination of Python Web Scrapping, SQL analysis and PowerBI visualizations offers a comprehensive view of the data, enabling better interpretation and decision-making.