**[Lok Sabha Election Analysis]**

**A PROJECT REPORT**

***Submitted by***

**[PATEL AYUSH PARASKUMAR]**

**[210670107089]**

***In partial fulfillment for the award of the degree of***

**BACHELOR OF ENGINEERING**

***In***

***[Computer Engineering]***

**SAL Institute of Technology and Engineering Research**

**Bhadaj Circle, Ahmedabad,Gujarat (Affiliated with GTU)**



******

**Gujarat Technological University, Ahmedabad**

**Academic Year**

**(2024-2025)**



******

**SAL Institute of Technology and Engineering Research**

**Bhadaj Circle, Ahmedabad, Gujarat (Affiliated with GTU)**

**CERTIFICATE**

This is to certify that the project report submitted along with the project entitled **Internship** has been carried out by **Ayush Patel** under my guidance in partial fulfillment for the degree of Bachelor of Engineering in **Computer Engineering,** 7th Semester of Gujarat Technological University, Ahmadabad during the academic year 2024-25.

**Prof. Ajeet Patel Dr. Nimisha Patel**

Internal Guide Head of the CE & CSE Department SALITER

**[Industry Letter Head]**

**Date: 24/06/2024**

**TO WHOM IT MAY CONCERN**

This is to certify that **Ayush Paraskumar Patel,** a student of **Sal Institute of Technology and Engineering Research** has successfully completed his internship in the field of **Data Analysis from 29thJune to 12thJuly (Total number of Weeks: 2) under the guidance of Rahul kirpekar.**

His internship activities include **Lok Sabha Election Analysis**

During the period of her/ his internship program with us, he / she had been exposed to different processes and was found diligent, hardworking and inquisitive.

We wish him every success in his life and career.

For **Grownited Private Limited**

Authorized Signature with Industry Stamp

******

**SAL Institute of Technology & Engineering Research**

**Bhadaj Circle, Ahmedabad, Gujarat (Affiliated with GTU)**

**DECLARATION**

**CERTIFICATE**

We hereby declare that the Internship / Project report submitted along with the Internship entitled Internship submitted in partial fulfillment for the degree of Bachelor of Engineering in **Computer Engineering** to Gujarat Technological University, Ahmedabad, is a bonafide record of original project work carried out by **me at Grownited Private Limited** **/ Sal Institute of Technology and Engineering Research** under the supervision of **Rahul Kirpekar & Prof. Ajeet Patel** and that no part of this report has been directly copied from any students’ reports or taken from any other source, without providing due reference.

**Name of the Student (Enrollment No) Sign of Student**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Acknowledgement**

I/We wish to express our sincere gratitude to our External guide **Mr.** **Rahul Kirpekar** for continuously guiding me at the company and answering all my doubts with patience. I/We would also like to thank my/our Internal Guide **Prof. Ajeet Patel** for helping us through our internship by giving us the necessary suggestions and advice along with their valuable co-ordination in completing this internship.

We also thank our parents, friends and all the members of the family for their precious support and encouragement which they had provided in completion of our work. In addition to that, we would also like to mention the company personals who gave us the permission to use and experience the valuable resources required for the internship.

Thus, In conclusion to the above said, we once again thank the staff members of **Grownited Private Limited** for their valuable support in completion of the project.

Thank You

Ayush Patel

**CONTENTS**

|  |  |  |  |
| --- | --- | --- | --- |
| Title Page | | | I |
| Certificates | | | II |
| Acknowledgement | | | V |
| Content | | | VI |
| List of Figures | | | VII |
|  | | |  |
| 1 |  | Introduction | 1 |
|  | 1.1 | Project Summery /Introduction of [Summer Internship/Mini Project Name/Specification/Seminar] | 1 |
|  | 1.2 | Aim and Objectives | 2 |
|  | 1.3 | Tools & Technologies | 3 |
| 2 |  | Implementation | 7 |
|  | 2.1 | Functional | 7 |
|  | 2.2 | Non-Functional | 8 |
| 3 |  | Outcomes | 9 |
|  | 3.1 | Conclusion | 9 |
|  | 3.2 | Future Enhancement | 9 |
|  | 3.3 | Progress report with result Pictures | 10 |
| 4 |  | Bibliography (Include only website links or book) | 24 |

**List of Figures**

|  |  |  |
| --- | --- | --- |
| **SR NO.** | **FIGURES NAMES** | **PAGE NO.** |
| 1.1  1.2  1.3  1.4  1.5  1.6  3.1.1  3.2.1  3.2.2  3.2.3  3.2.4  3.2.5  3.2.6  3.2.7  3.2.8  3.2.9  3.3.1  3.3.2  3.3.3  3.3.4  3.3.5  3.4.1  3.4.2  3.4.3  3.4.4  3.4.5  3.4.6  3.4.7  3.4.8  3.4.9  3.4.10 | Power Bi Symbol  Google Colaboratory Symbol  Excel Symbol  Power Query Symbol  Jira Symbol  GitHub Symbol  Main Dashboard  Voter turnout  Party Performance Trend  Winning Magin  Distribution of Winning Margin  Electors Vs Margin  Electors Vs Turnout  heatmap Turnout & Electors  heatmap Turnout, Margin & Margin%  Margin percentage  Voter turnout comparison  Distribution of votes  Distribution of Margin percentage  Winning margin by State  Winning rate by State  Average Voter Turnout over the Year  Turnout vs Margin for Party  Year wise Margin Trend for Party  Year wise Turnout Trend for Party  Average Margin by Caste  Pie chart (Total votes by State)  Correlations Between Turnout & Margin %  heatmap (Votes, Margin % & Turnout)  Vote Margins Across Different States  Turnout By Caste | 3  3  4  5  5  6  10  11  11  12  12  13  13  14  14  15  16  17  17  18  18  19  19  20  20  21  21  22  22  23  23 |

**Chapter-1 Introduction**

* 1. **Project Summary of Lok Sabha Election Analysis**
     1. **Project Profile**

The Lok Sabha Elections Analysis project examines India's parliamentary elections from 2004 to 2024 using Power BI and Python libraries for data analysis and visualization. It aims to provide insights into voter turnout, candidate demographics, and political party performance.

* + 1. **Project Purpose**

The purpose of this project is to conduct an in-depth analysis of the Lok Sabha elections in India over the past two decades, from 2004 to 2024. By leveraging advanced data analysis and visualization tools, the project aims to uncover trends, patterns, and insights that can provide a deeper understanding of voter behaviour, candidate demographics, and political party performance. This understanding is crucial for various stakeholders, including political strategists, researchers, and policymakers, to make informed decisions and strategies for future elections.

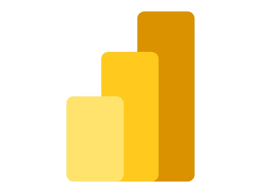
* + 1. **Project Expected Outcome**

The project will deliver insights into voter turnout, candidate demographics, and political party performance in Lok Sabha elections over the past 20 years. By analyzing regional voting patterns and demographic impacts, it will aid political strategists, researchers, and policymakers in understanding voter behavior and shaping future electoral strategies.

* 1. **Aim and Objectives**
     1. **Aim of Project**
* Examine trends and changes in voter participation across different elections and regions.
* Evaluate the demographics of winning candidates, including age, gender, education, and occupation.
* Measure the performance of major political parties over the last 20 years.
* Understand the relationship between demographic factors and electoral results, focusing on the size and composition of the electorate.
* Use advanced data analysis and visualization tools to present clear and actionable insights for stakeholders.
* Offer valuable information to political strategists, researchers, and policymakers to refine their approaches and strategies for future elections
  + 1. **Objectives**
* To analyze voter turnout and participation trends from 2004 to 2024.
* To examine the demographic profiles of winning candidates, focusing on age and gender.
* To assess the performance of major political parties and the emergence of regional parties.
* To explore regional voting patterns and the urban-rural divide.
* To evaluate the impact of electoral reforms and technology on the election process

2

* 1. **Tools & Technologies**
     1. ***Power Bi***



# Fig 1.1 Power Bi Symbol

Power BI is a suite of business analytics tools that deliver insights throughout your organization. It helps in connecting to hundreds of data sources, simplifying data preparation, and driving ad hoc analysis. With Power BI, users can produce beautiful reports and publish them for their organization to consume on the web and across mobile devices.

* + 1. ***Google Colab***



# Fig 1.2 Google Colaboratory Symbol

Google Colab is an online Jupyter notebook environment hosted by Google. It allows users to write and execute Python code directly in their web browser, making it accessible from any device with an internet connection. Colab provides the computational resources, such as CPUs and GPUs, required to run code efficiently, which is especially beneficial for tasks involving large datasets and complex computations.

* + 1. ***Python Libraries***
* ***Pandas*** is a powerful, open-source data manipulation and analysis library for Python. It provides data structures and functions needed to manipulate numerical tables and time series data.
* ***Matplotlib*** is a comprehensive library for creating static, animated, and interactive visualizations in Python. It is highly customizable and widely used for creating publication-quality plots and graphs.
* ***Seaborn*** is a Python data visualization library based on Matplotlib that provides a high-level interface for drawing attractive and informative statistical graphic
* ***BeautifulSoup*** is a Python library used for web scraping purposes to pull the data out of HTML and XML files. It creates parse trees from page source codes that can be used to extract data easily.
  + 1. ***Excel***



# Fig 1.3 Excel Symbol

Microsoft Excel is a spreadsheet program that allows users to organize, format, and calculate data using a system of rows and columns. It is a versatile tool that can be used for various purposes, including data analysis, financial modelling, project management, and more. Excel's interface consists of a grid of cells, each of which can contain numbers, text, formulas, or functions.

* + 1. ***Power Query***



# Fig 1.4 Power Query Symbol

Power Query is a data transformation and data preparation engine. It allows users to extract data from various sources, transform it to meet their needs, and load it into Excel or Power BI for further analysis. Power Query provides an intuitive, userfriendly ] interface for data manipulation, making it accessible to both technical and non-technical users.

* + 1. ***Jira***



# Fig 1.5 Jira Symbol

Jira is a comprehensive and flexible project management tool that supports various methodologies and use cases. Its powerful features for issue tracking, agile project management, collaboration, and reporting make it an essential tool for software development teams and organizations managing complex projects. Whether you are tracking bugs, planning sprints, or managing business processes, Jira provides the tools and capabilities needed to streamline your workflows and achieve your project goals.

* + 1. ***GitHub***



Fig 1.6 GitHub Symbol

GitHub is an essential tool for modern software development, providing powerful version control, collaboration, and project management capabilities. Its wide range of features, active community, and seamless integrations make it an invaluable resource for developers, teams, and organizations working on projects of all sizes and complexities. Whether you are developing proprietary software, contributing to open-source projects, or managing a complex development pipeline, GitHub offers the tools and support needed to succeed.

**Chapter-2 Implementation**

1. **Functional**

**2.1.1 Data Collection Phase:**

* Gathered data from reliable sources such as the Election Commission of India website, Wikipedia and India votes including details on candidates, votes, parties, and constituencies. Used Power Query for web scraping and preprocessing of the raw data from the websites/sources.

**2.1.2 Data Cleaning and Preprocessing Phase:**

* Cleaned and preprocessed the data for analysis using Python and Pandas; such as handling missing values, normalization, removing unnecessary column, etc.

**2.1.3**  **Data Analysis Phase:**

* **Using NumPy and Pandas we worked on tasks like Analysing voter turnout trends and Winning Margin analysis or Impact of Electorate size on Outcomes or State wise performance of Parties or Turnout and winning candidate analysis or Margin percentage analysis.**

**2.1.4 Visualization Phase:**

* The analyzed data is visualized using Power BI, focusing on creating interactive dashboards and visualizations that effectively communicate the insights derived from the data.

**2.1.5 Reporting Phase:**

* The final phase involves compiling the analysis and visualization results into a comprehensive report that includes detailed findings, interpretations, and recommendations based on the analyzed data.
  1. **Non-Functional**

**2.2.1 Performance**

* Ensure the system can handle large datasets efficiently.
* Optimize data processing and visualization generation for quick response times.

**2.2.2 Scalability**

* Design the system to accommodate future data expansions and additional election cycles.
* Ensure the system can scale to handle increased data volume and user load.

**2.2.3 Usability**

* Ensure a user-friendly and intuitive interface.
* Provide clear navigation and accessible features.

**2.2.4 Reliability**

* Ensure the system is reliable and available with minimal downtime.
* Implement backup and recovery mechanisms to protect against data loss.

**2.2.3 Maintainability**

* Develop the system with modular and well-documented code for easy maintenance.3
* Ensure the system can be updated and improved over time without significantdisruptions.

**2.2.3 Compatibility**

* Ensure compatibility with various devices and platforms.
* Design the system to work seamlessly with different web browsers and operating systems.

**Chapter-3 Outcomes**

**3.1 Conclusion**

The Lok Sabha Elections Analysis project has provided a thorough examination of India's parliamentary elections from 2004 to 2024, uncovering valuable insights into voter turnout, candidate demographics, and political party performance. Utilizing advanced data analysis and visualization tools, the project has shed light on regional voting patterns and demographic impacts, offering crucial information for political strategists, researchers, and policymakers. The findings from this project enable a deeper understanding of electoral trends and behaviours, aiding in the development of informed strategies for future elections.

**3.2 Future Enhancement**

Looking ahead, the project can be enhanced by incorporating real-time data analysis for more immediate insights into ongoing elections. Expanding the dataset to include more granular data, such as constituency-level details, could provide even richer insights. Additionally, integrating machine learning algorithms to predict election outcomes and voter behaviour patterns could significantly augment the project's analytical capabilities. Collaborating with social media platforms to analyze sentiment and its impact on elections could also offer a new dimension to the analysis, making it even more comprehensive and impactful.

**3.3 Progress Report with Result Pictures**

* + 1. **Dashboard**

****

Fig 3.1.1 Main Dashboard

* ***Description***

This dashboard provides an analysis of Indian election data, featuring various visualizations. Key metrics include vote counts, elector counts, and winning candidates by year, state, party, and caste. The visualizations include bar charts, pie charts, and a map to illustrate geographic distribution. The data spans from 2004 to recent elections, highlighting trends in voter turnout and party performance. Additionally, the dashboard shows the total votes and electors by state, offering insights into regional variations. It also breaks down the count of winning candidates by caste, showing the representation across different demographics. Overall, this comprehensive analysis aids in understanding the electoral landscape and voter behaviour in India.

* + 1. **Python Codes (Colab Files)**
* ***Pandas:***
* **Analyzing voter turnout trends:** How does the voter turnout percentage vary across different states in the 2024 Lok Sabha election?

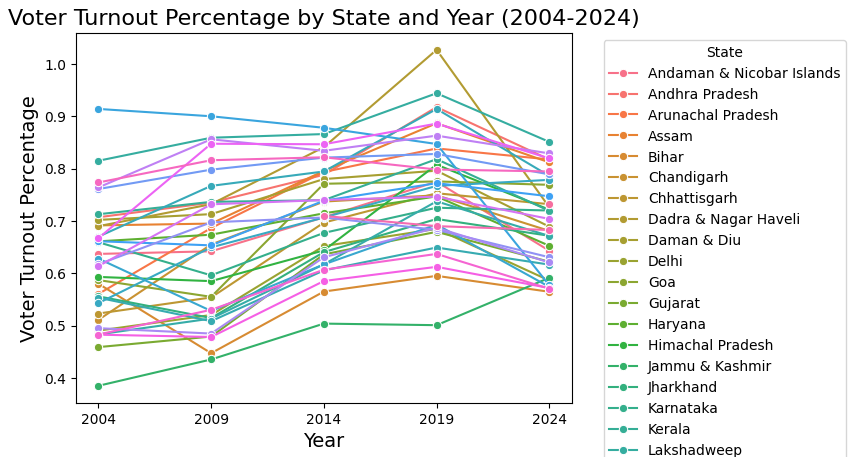
****

Fig 3.2.1 Voter turnout

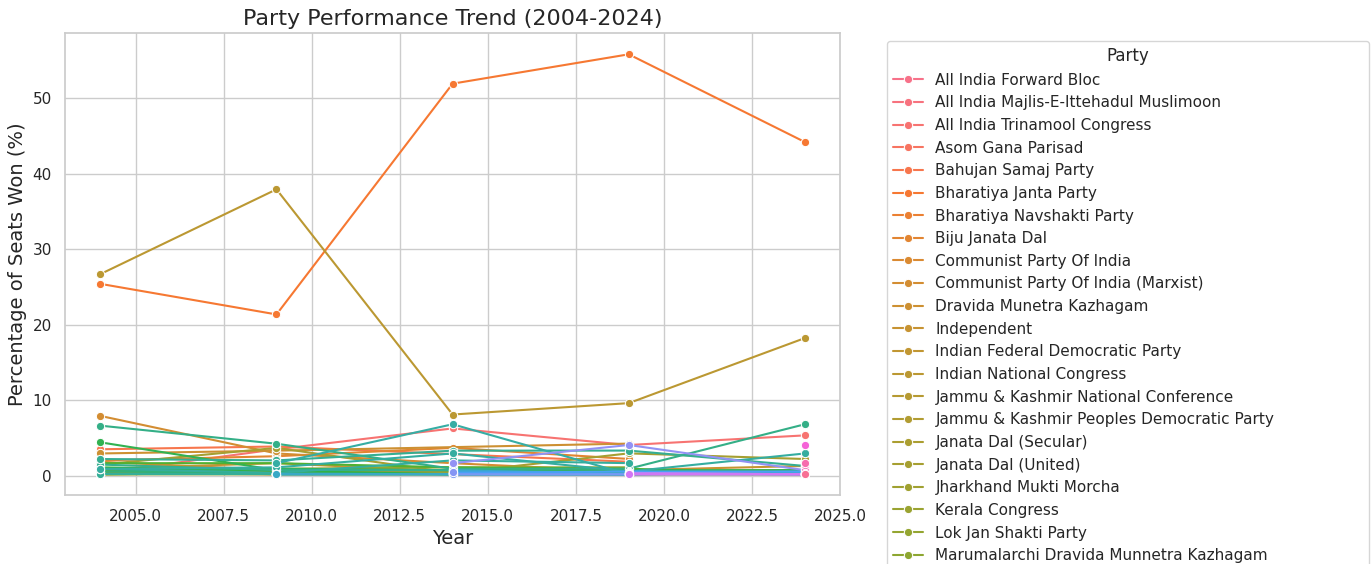


Fig 3.2.2 Party Performance Trend

* **Winning Margin analysis:** What is the distribution of winning margins across all constituencies, and how does it compare between different states?

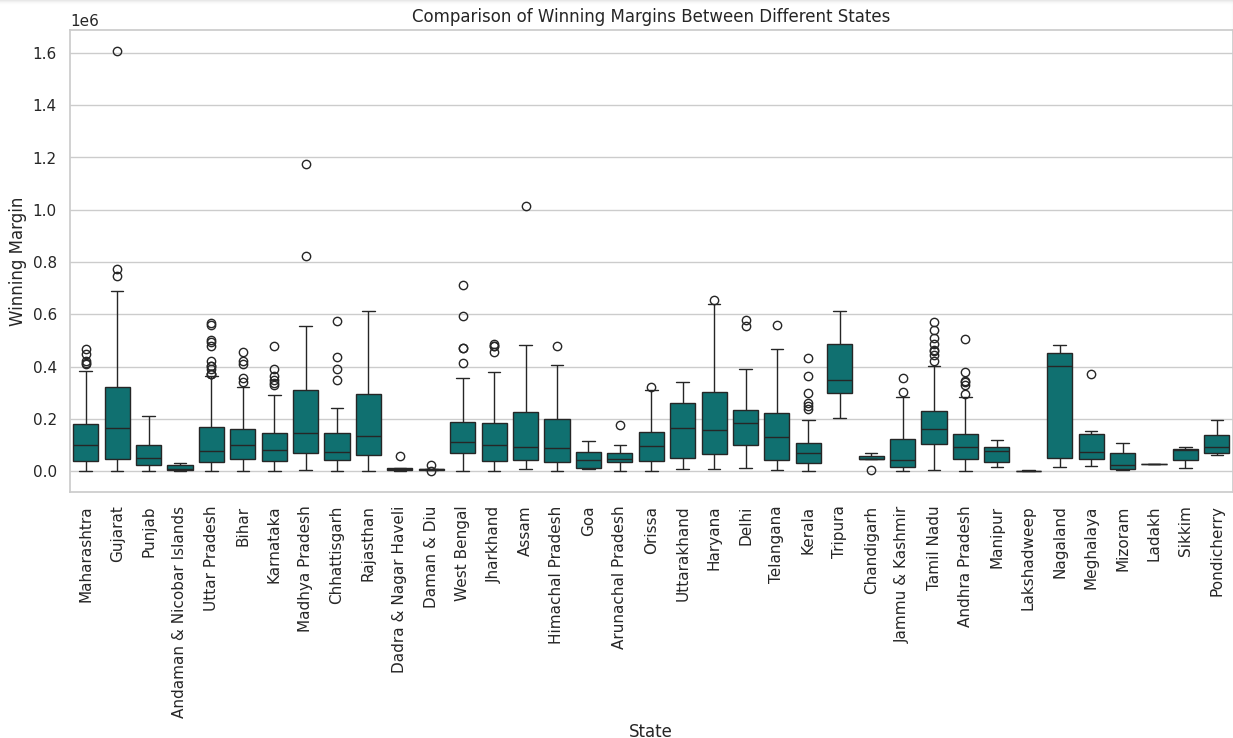
****

Fig 3.2.3 Winning Magin

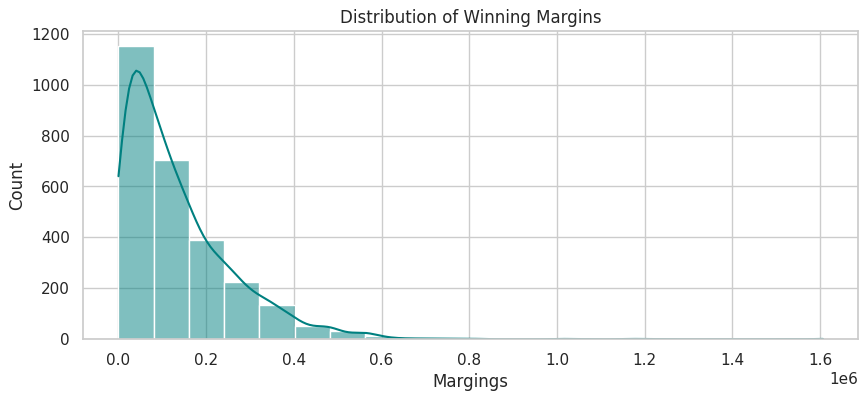


Fig 3.2.4 Distribution of Winning Margin

**Impact of Electorate size on Outcomes:** Is there a correlation between the number of electors in a constituency and the voter turnout percentage?

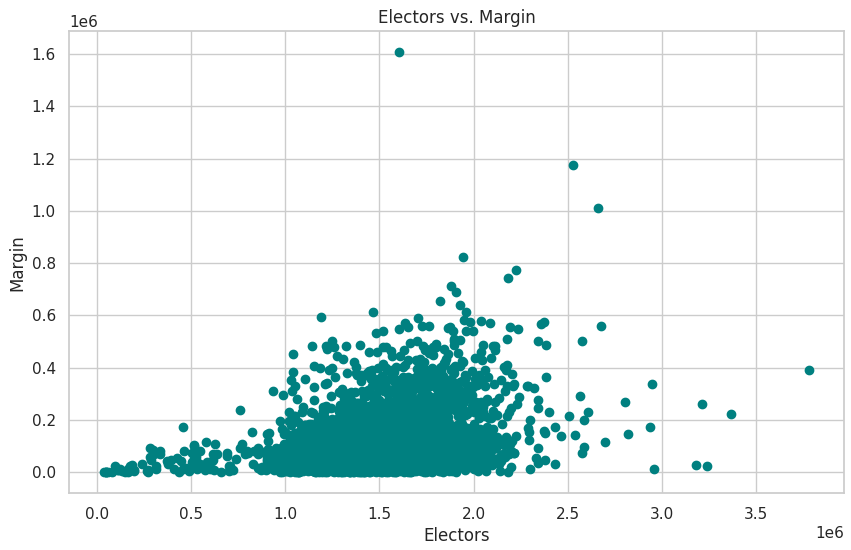
****

Fig 3.2.5 Electors Vs Margin

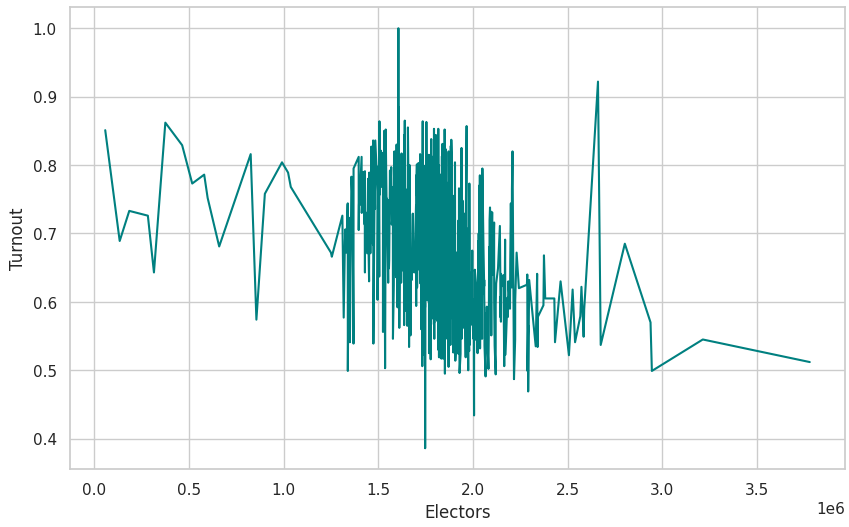


Fig 3.2.6 Electors vs Turnout

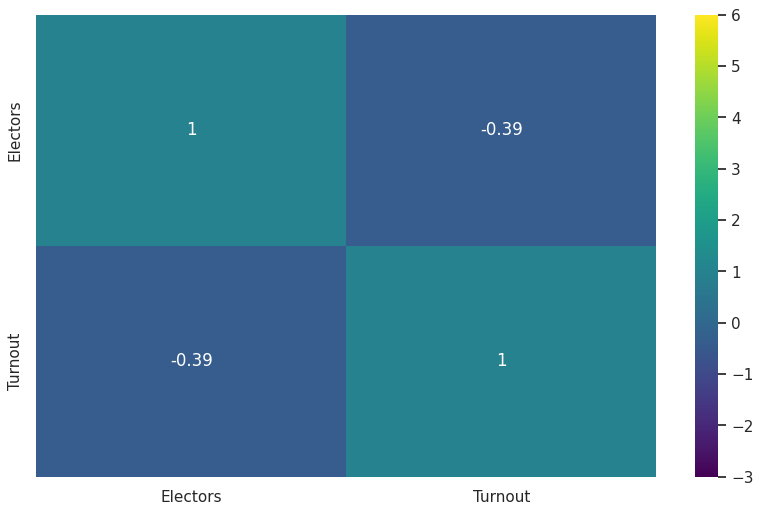


Fig 3.2.7 heatmap Turnout & Electors

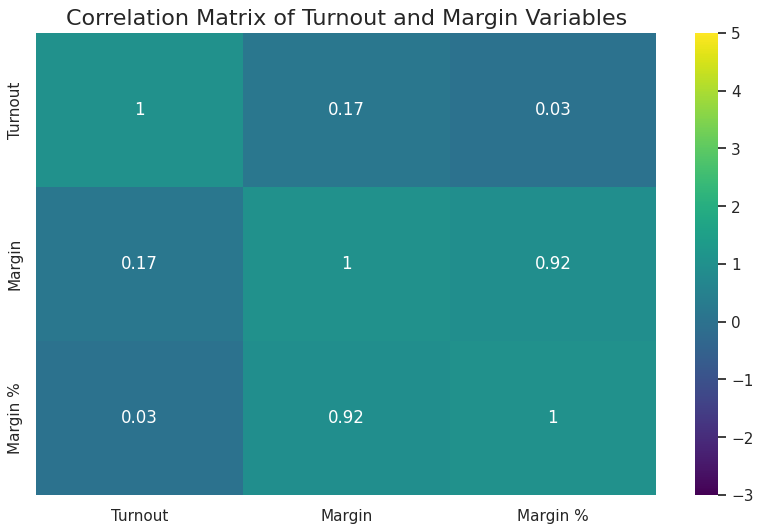


Fig 3.2.8 heatmap Turnout, Margin & Margin%

* **Margin percentage analysis:** What is the average margin percentage for winning candidates in constituencies with different majority castes?

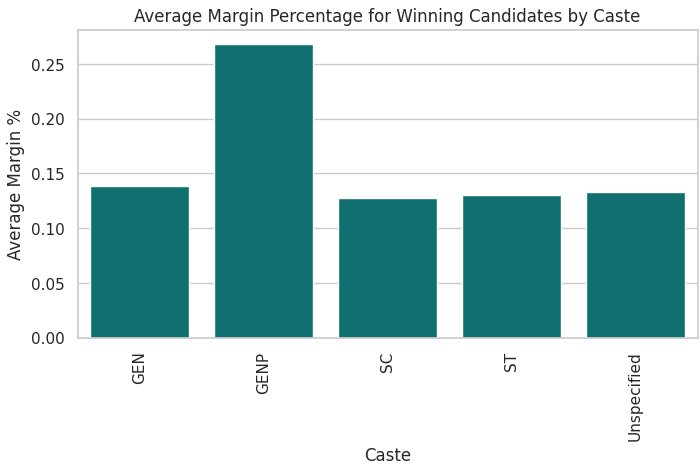
****

Fig 3.2.9 Margin percentage

**3.3.3 Statistical Methods and Hypothesis Testing**

In our analysis of the Lok Sabha election data, we employed various statistical methods to interpret and understand the voting patterns and trends. Using Python libraries such as pandas, matplotlib, and seaborn, we conducted exploratory data analysis (EDA) to visualize the distribution of votes, turnout rates, and party performance over the years.

***Central Tendency***: Calculate mean, median, and mode for voter turnout, votes, and margins.

***Dispersion***: Calculate standard deviation, variance, range, and interquartile range for turnout and margins.

***Distribution***: Plot histograms and box plots to understand the distribution of votes, turnout, and margins.

***Correlation***: Calculate correlation coefficients to understand the relationships between variables, such as turnout and margin.

* **Voter Turnout Comparison Across States**

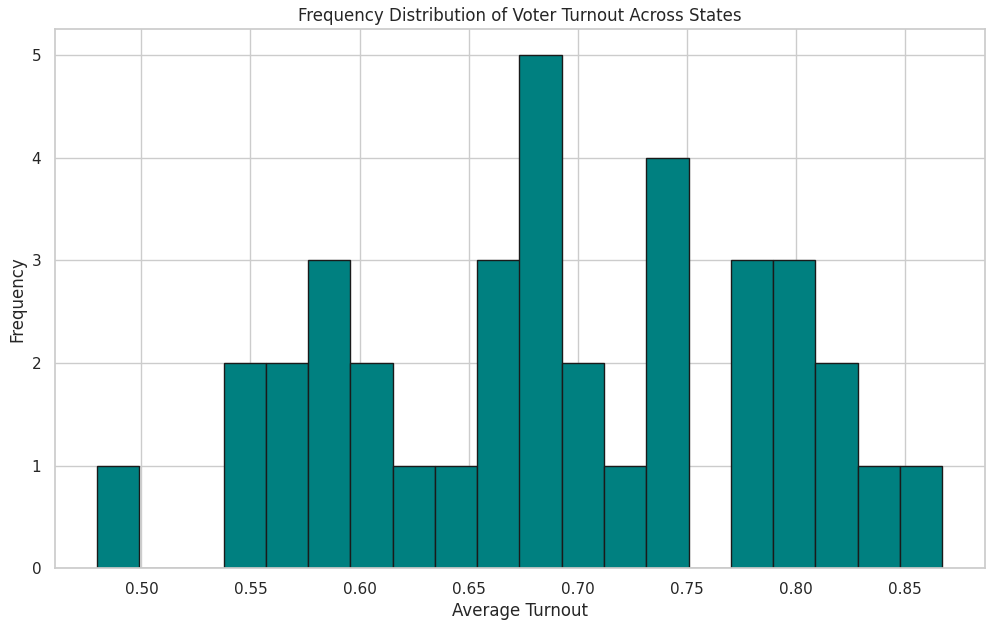
****

Fig 3.3.1 Voter turnout comparison

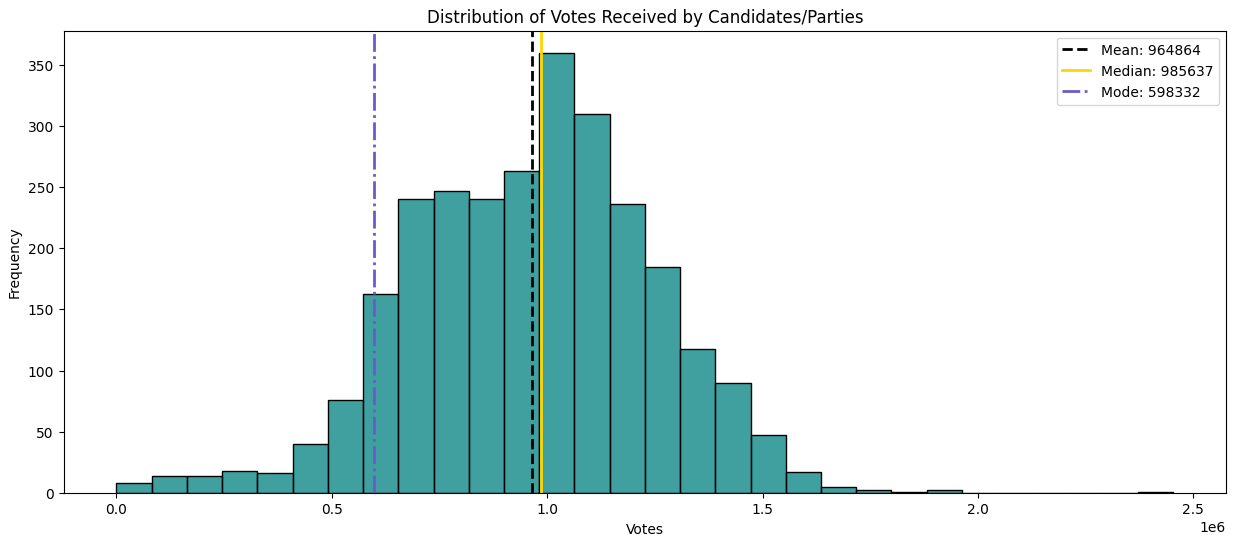
****

Fig 3.3.2 Distribution of votes

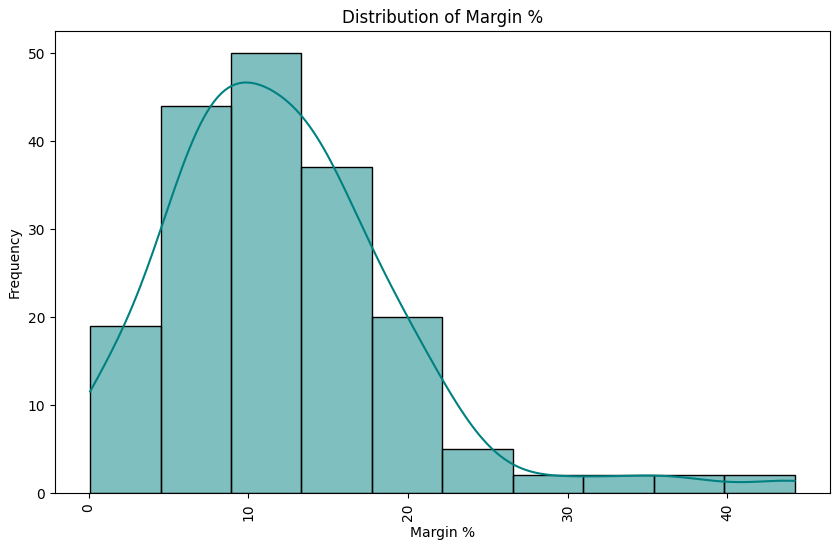
****

Fig 3.3.3 Distribution of Margin percentage

* **Winning margin by state**

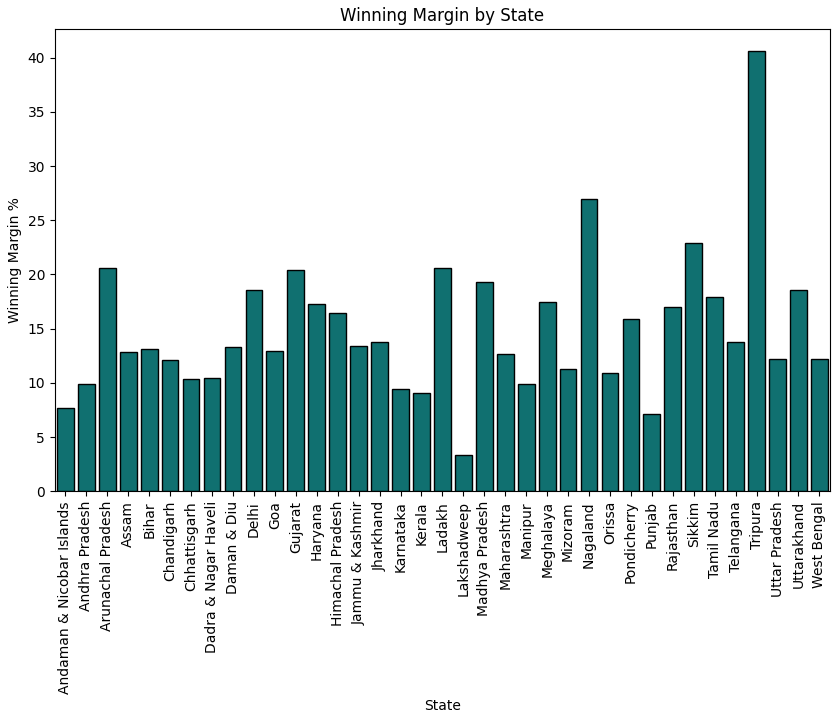
****

Fig 3.3.4 Winning margin by State

* **Winning rate by state**

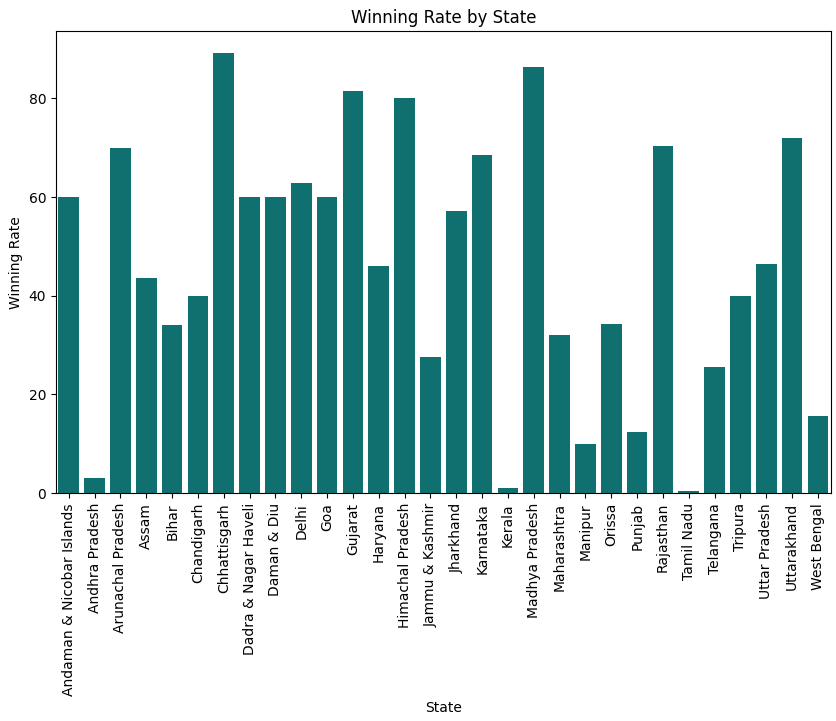


Fig 3.3.5 Winning rate by State

**3.3.4 Matplotlib and Seaborn**

**Used Matplotlib and Seaborn for visualizations like:**

* **Line Plots**: Average Voter Turnout Over the Years

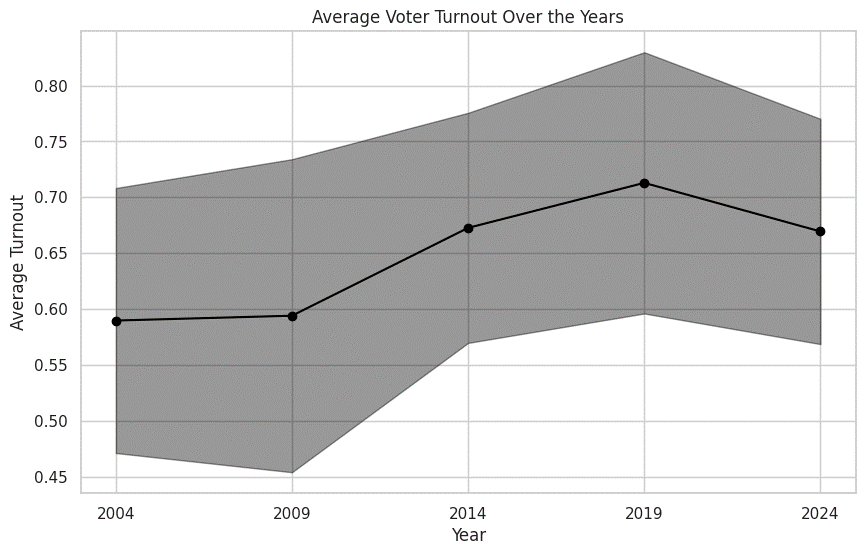


Fig 3.4.1 Average Voter Turnout over the Years

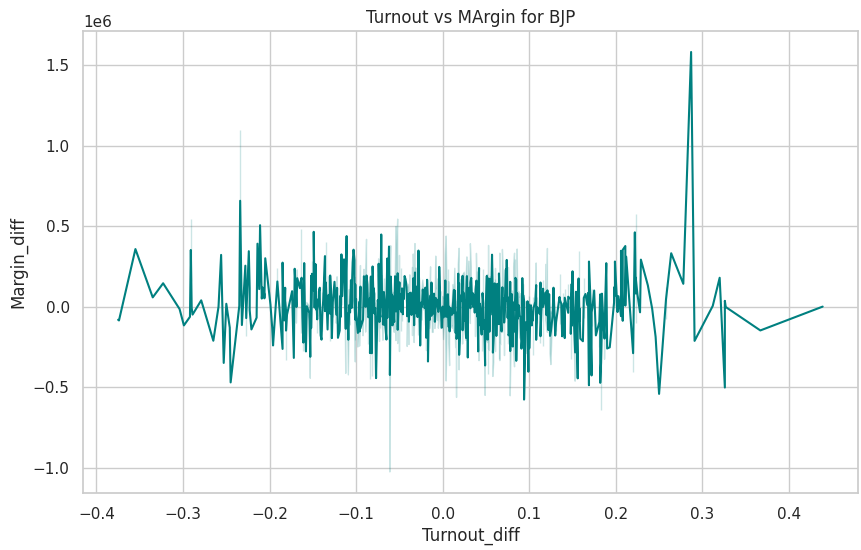


Fig 3.4.2 Turnout vs Margin for Party

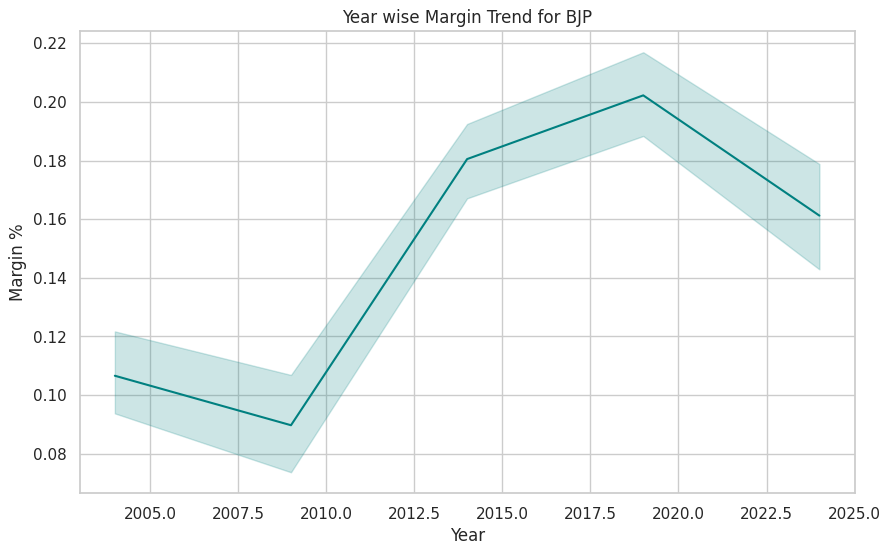


Fig 3.4.3 Year wise Margin Trend for Party

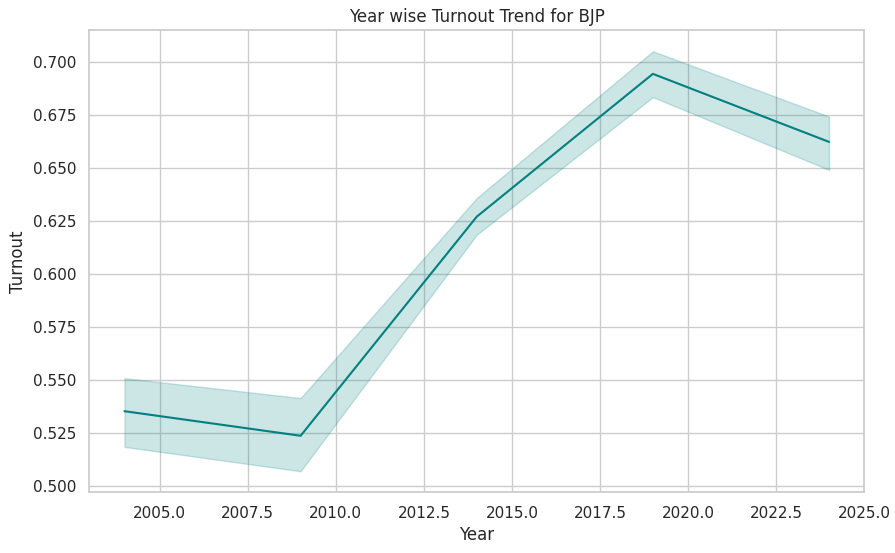


Fig 3.4.4 Year wise Turnout Trend for Party

* **Bar Charts**: Compare performance of different parties across states using bar charts.

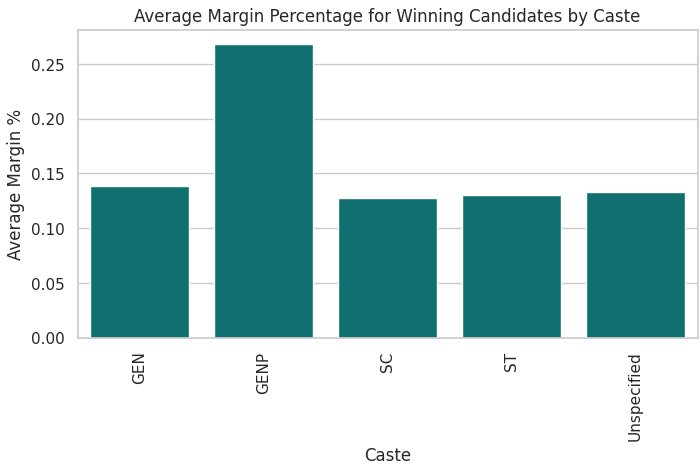


Fig 3.4.5 Average Margin by Caste

* **Pie Charts**: Visualize the distribution of seats among parties.

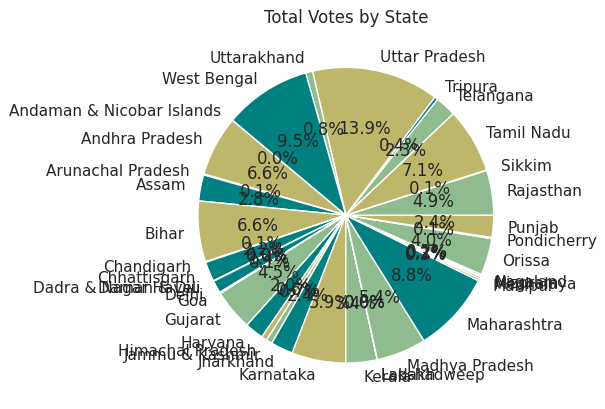


Fig 3.4.6 Pie chart (Total votes by State)

* **Scatter Plots**: Identify correlations between different variables (e.g., margin percentage and turnout).

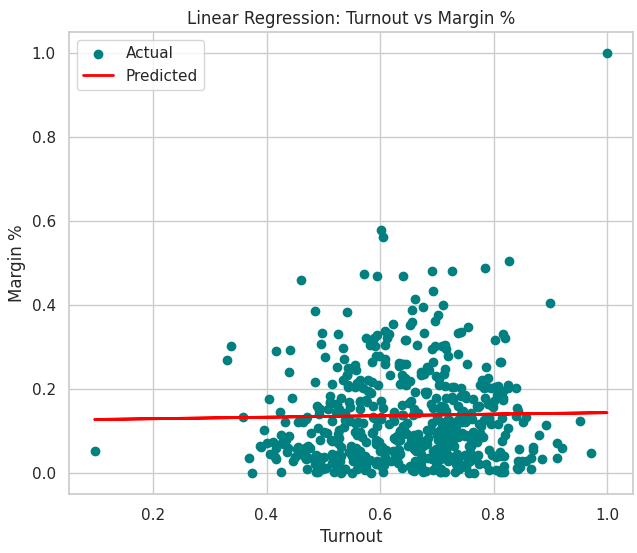


Fig 3.4.7 Correlations Between Turnout & Margin %

* **Heatmaps**: Create heatmaps to visualize correlations between multiple variables, such as turnout, margin percentage, and votes.

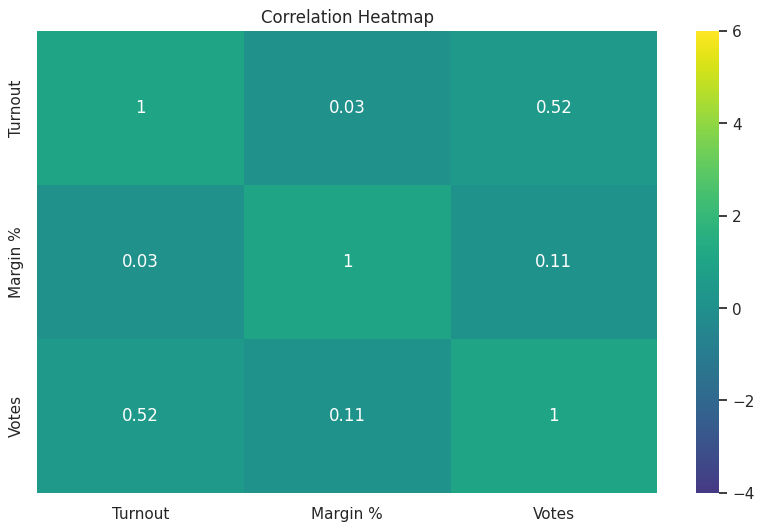


Fig 3.4.8 heatmap (Votes, Margin % & Turnout)

* **Box Plots**: Show the distribution of vote margins across different states or parties to identify outliers and overall performance.

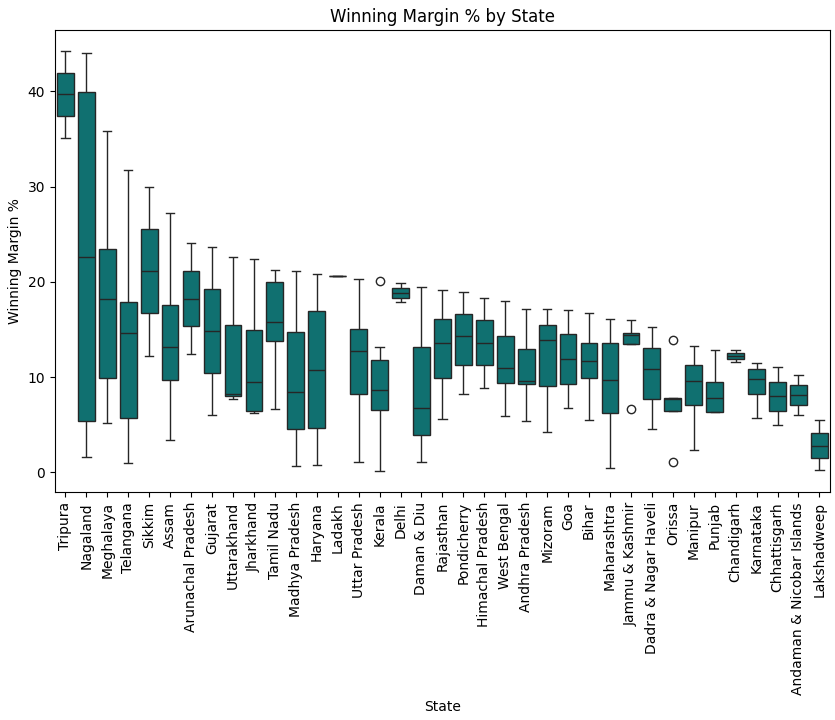


Fig 3.4.9 Vote Margins Across Different State

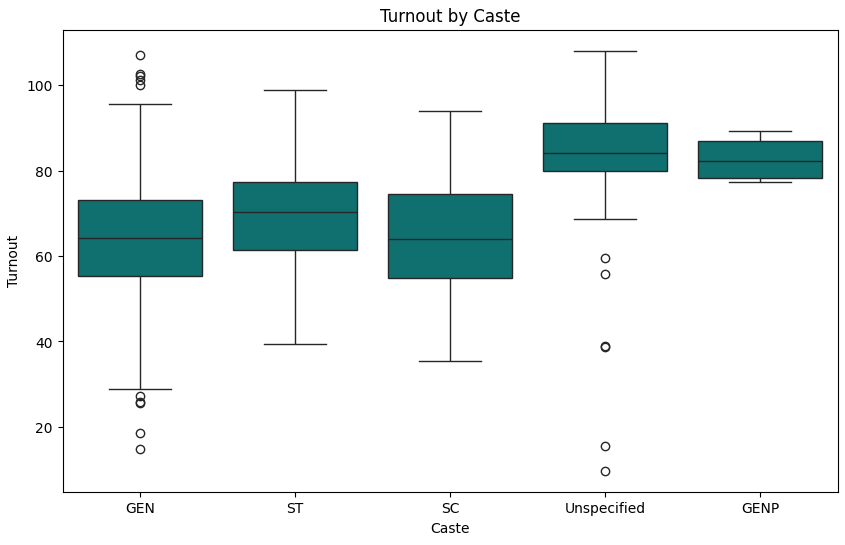
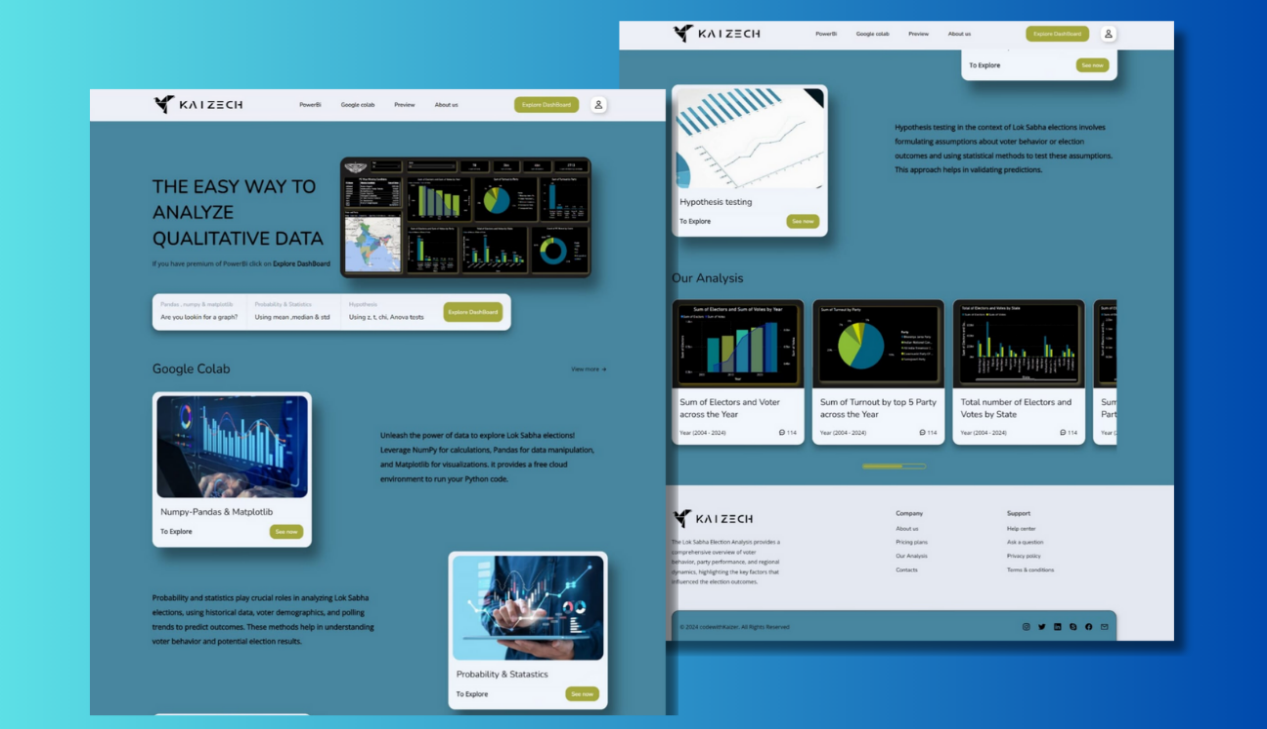


Fig 3.4.10 Turnout By Caste

* + 1. **Website Desktop**



**Chapter-4 Bibliography**

* 1. **References**

[1] <https://www.kaggle.com/>

[2] <https://www.indiavotes.com/>

[3] <https://www.eci.gov.in/statistical-reports>

[4] <https://en.wikipedia.org/wiki/2024_Indian_general_election#Surveys_and_polls>

[5] <https://www.youtube.com/watch?v=PiFAa_jjaEI>

[6] <https://www.youtube.com/watch?v=Beozx9pUajA>

[7] <https://www.youtube.com/watch?v=RGOj5yH7evk>

[8] <https://www.youtube.com/playlist?list=PLc20sA5NNOvrsn3a78ewy2VTCXVV47NB4>

[9] <https://www.youtube.com/watch?v=XVv6mJpFOb0>

[10] <https://www.youtube.com/watch?v=8dTpNajxaH0>

[11] <https://www.youtube.com/watch?v=2hPCX-p_X8Q>

[12] <https://www.youtube.com/watch?v=mBoX_JCKZTE>

[13] <https://www.youtube.com/watch?v=UOsRrxMKJYk>

* 1. **Website and GitHub Link**

After completing my data analysis tasks, all links of google colab were uploaded on the GitHub.

Also, created a website and hosted it on GitHub, where insights from project were shared. Used Html, CSS, JavaScript to create website. Power Bi dashboard for the project was embedded on the website.

* **Link To Our Website:**

<https://20ayush04.github.io/Lok_Sabha_sys/>

* **GitHub Link for All Tasks:**

<https://github.com/20ayush04/Lok_Sabha_sys>