“Electviz: Election Data Visualization for Media Documentation”

A PROJECT SUBMITTED TO -

INFOSYS SPRINGBOARD 6.0

BY

Arjun Srivastava

Mayur Shivaji Raut

Kotha Akshaya

Srilakshmi Sowjanya

TEAM INTRODUCTION

We are a team passionate about data analytics, visualization, and decision intelligence. Through this project, we aimed to understand real-world electoral datasets, clean and model the data, and design dashboards that simplify complex election information.

This project enhanced our skills in teamwork, analytical thinking, BI tool usage, and delivering insights for real-time decision-making.

Project Duration: 8 Weeks

TEAM CONTRIBUTION

Members	Contribution
➢ Arjun Srivastava	➢ Data cleaning, DAX, dashboard insights, final integration
➢ Mayur Shivaji Raut	➢ Data collection, preprocessing, attribute mapping
➢ Kotha Akshaya	➢ DAX measures, KPI setup, visual formatting
➢ Srilakshmi Sowjanya	➢ Documentation, PPT design, testing & improvements

1

ABSTRACT

Elections produce enormous amounts of structured and unstructured data that are challenging to interpret quickly. The **ElectViz Dashboard** project aims to simplify this information for journalists, analysts, and citizens by converting raw electoral data into clear, interactive visuals using **Microsoft Power BI**.

The system highlights party performance, candidate profiles, vote-share patterns, and socio-demographic insights such as gender, education, wealth, and criminal cases.
It enables instant comparisons across states and parties, supporting **data-driven storytelling** during election coverage. The dashboards are optimized for **speed, clarity, and interactivity**, helping users understand *who is winning, where, and by how much*.

OBJECTIVES

- To visualize national and state-level election results in a clean, comprehensible format.
- To identify winning trends by **party, region, gender, education, and assets**.
- To assist media teams in delivering rapid and accurate insights during counting or post-poll analysis.
- To present correlations between candidate wealth, liabilities, and criminal records.
- To establish a scalable dashboard framework reusable for future elections or comparative studies.

PROBLEM STATEMENT

During large-scale elections, thousands of data points are generated regarding candidates, constituencies, and results.
Traditional reporting methods are slow and text-heavy, making it difficult for the public to interpret outcomes or detect patterns.

There is a need for an **interactive visualization system** that can:

- Present detailed insights in real-time.
- Summarize complex metrics such as vote share, turnout, and demographics.
- Allow filtering by multiple parameters (party, state, gender, education).

This project addresses these challenges by developing a **Power BI-based analytical dashboard** that delivers transparency and actionable insights.

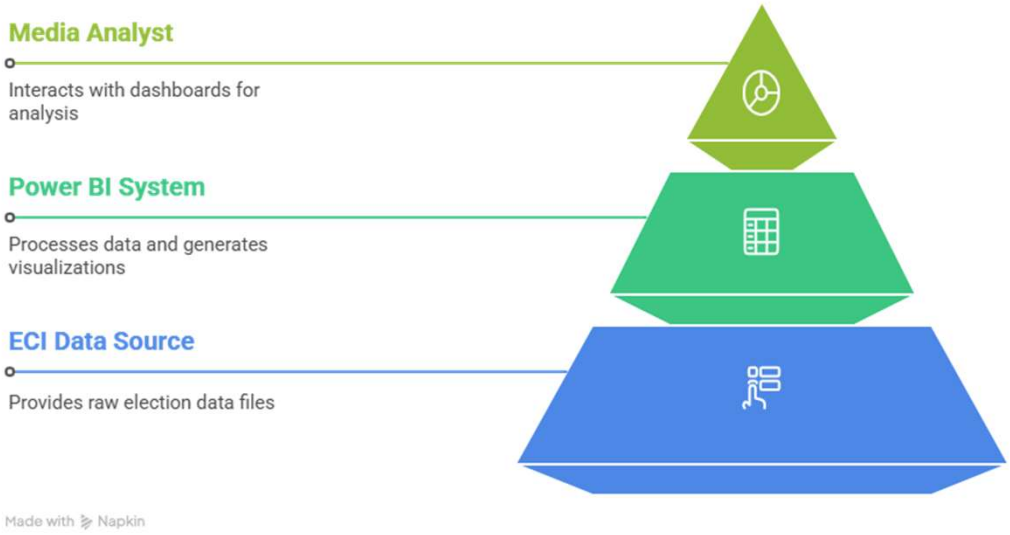
KEY FEATURES

- Interactive KPIs – Displays total candidates, winners, average age, and financial statistics.
- State-wise & Party-wise Analysis – Bar and map visuals for regional and political distribution.
- Demographic Insights – Gender ratio, education level, and caste/category representation.
- Financial & Legal Insights – Correlation between assets, liabilities, and criminal records.
- Voting Pattern Analysis – Comparison of general vs postal votes and postal vote share.
- Performance Indicators – Top candidates by wealth, votes, and criminal cases.
- Dynamic Filtering – State, Party, Gender, Education, and Category filters.

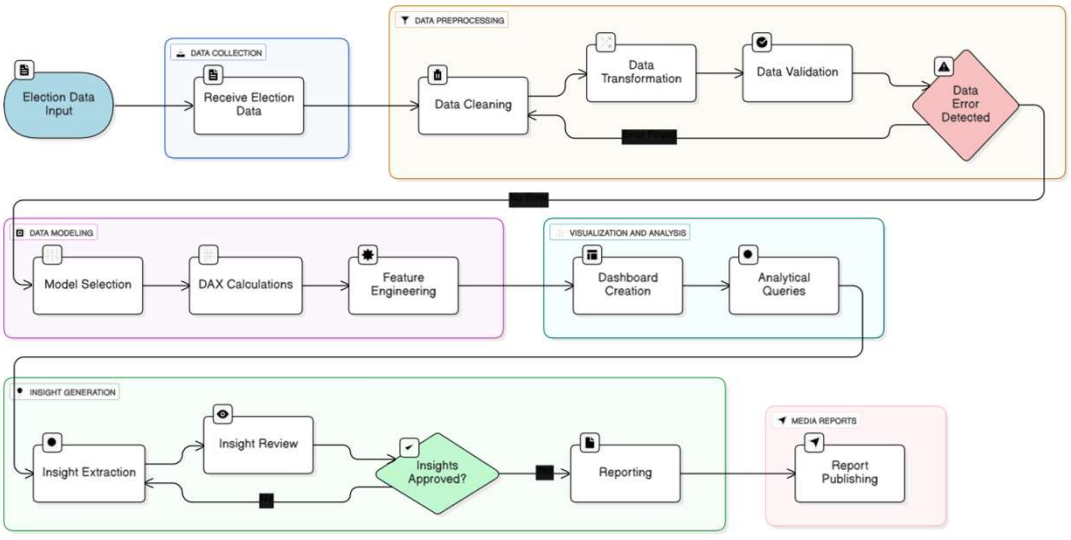
DATA FLOW DIAGRAM:

The **ElectViz System** follows a well-structured flow of data from collection to visualization.
The **Data Flow Diagram (DFD)** explains how election data moves through the system and is transformed into meaningful insights.

Context Diagram



Detailed Data Flow of the ElectViz System



INSIGHT GENERATION

The **ElectViz Dashboard** generates clear and meaningful insights from complex election data using Power BI.
By applying **data modeling** and **DAX calculations**, raw datasets are transformed into KPIs and visual stories that highlight political, demographic, and financial patterns.

Key insights include:

- **Party Performance:** BJP and INC dominate overall results, while regional parties like DMK and AITC lead within their states.
- **Demographics:** Around 86 % of winners are male, and most belong to the 46–65 age group. Post-graduates show a slightly higher success rate.
- **Financial Insights:** Wealthier candidates often secure more seats, and many have declared criminal cases, showing a link between influence and resources.
- **Voting Trends:** States such as Uttar Pradesh and Maharashtra contribute the highest number of winners, with postal votes forming a small yet decisive share.

These insights help media teams and analysts understand **who is winning, where, and why**, making election coverage faster and more data-driven.

FUTURE SCOPE

In the future, the **ElectViz system** can be expanded beyond static visualizations to include **predictive and real-time analytics**.

Planned improvements:

- **Real-time Data Integration:** Connect dashboards to live feeds or APIs from the Election Commission for automatic updates during counting.
- **Predictive Modeling:** Use Python-based models like Random Forest, ARIMA, or Prophet to forecast vote share and winning probabilities.
- **Web & Mobile Accessibility:** Publish dashboards online via Power BI Service for easy access by media, analysts, and the public.
- **Historical Comparison:** Integrate data from previous elections to study long-term voting trends and regional shifts.

These enhancements will make ElectViz a **comprehensive election intelligence tool**, combining visualization, prediction, and live monitoring in one platform.

TECHNOLOGY STACK

The **ElectViz Dashboard** uses a blend of tools for data processing, analysis, and visualization. Each component contributes to transforming raw election data into clear, interactive insights.

1. **Power BI:**
   Main platform for visualization and analysis. Used to design dashboards, create KPIs, maps, and charts, and apply filters for state, party, and category analysis.
2. **Power Query:**
   Handles data cleaning and transformation. Removes duplicates, merges tables, and prepares structured data for reporting.
3. **DAX (Data Analysis Expressions):**
   Used to create calculated measures such as *Average Assets, Vote Share %,* and *Total Winners* for deeper analytical insights.
4. **Microsoft Excel / CSV Files:**
   Stores and organizes raw datasets obtained from the Election Commission of India before importing into Power BI.
5. **Python (Optional):**
   Used for additional analysis or forecasting with libraries like *Pandas* and *Scikit-learn*.

Together, these tools create a **complete analytical ecosystem**, enabling accurate, data-driven election visualization and media-ready reporting.

MODULE DESCRIPTION

The **ElectViz Project** is divided into several modules, each responsible for a specific stage of the data analysis process. These modules ensure a smooth workflow — from collecting raw election data to generating meaningful insights.

1. Data Collection Module

Purpose:
To gather official election data from reliable sources.
Description:
Raw datasets containing information about candidates, parties, votes, and demographics are collected in CSV or Excel format. The data is stored and organized for use in Power BI.

2. Data Cleaning & Transformation Module

Purpose:
To ensure the collected data is accurate, complete, and formatted properly.
Description:
Using **Power Query**, missing values are handled, duplicate entries are removed, and column names are standardized. The final cleaned dataset is ready for modeling and analysis.

3. Data Modeling Module

Purpose:
To establish relationships between different datasets and create calculated measures.
Description:
Relationships between tables such as *State, Party,* and *Candidate* are defined.
Using **DAX formulas**, new fields like *Total Votes, Vote Share %, Average Assets,* and *Criminal Cases per Candidate* are created to support deeper insights.

4. Visualization Module

Purpose:

To present analyzed data in a clear and interactive format.

Description:

Power BI is used to design dashboards with visuals such as **bar charts, donut charts, maps, and KPI cards**. Each visual helps users understand patterns like party dominance, gender ratios, and regional performance at a glance.

5. Reporting & Insight Module

Purpose:

To summarize insights and enable interactive data exploration.

Description:

Users can apply filters (State, Party, Education, Gender) to generate custom views and reports. Dashboards are exported or shared through Power BI for presentation and publication.

6. Forecasting / Future Integration Module (Optional)

Purpose:

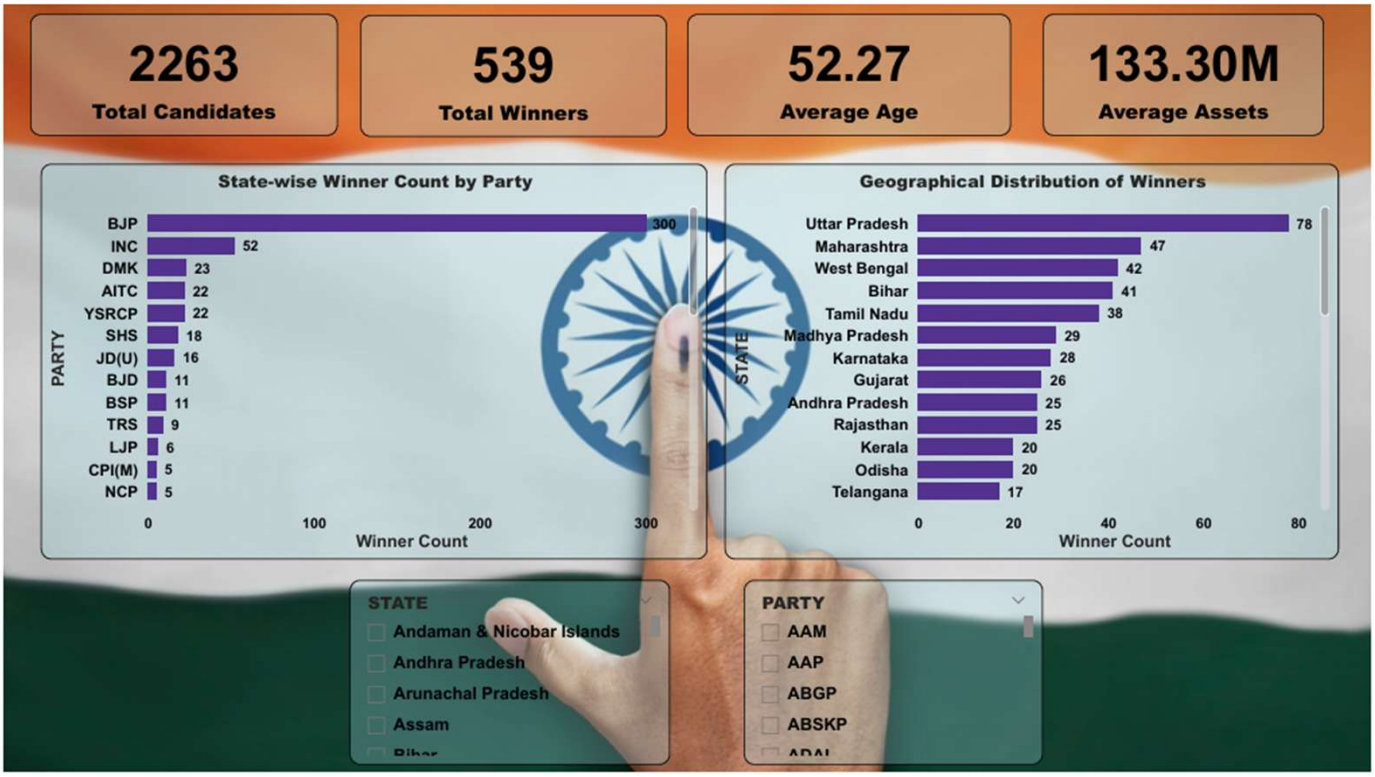
To enhance the system with predictive capabilities.

Description:

Python or Power BI integration can be used to build models like **Random Forest or ARIMA** to forecast vote share or winning probabilities in future elections.

PROJECT DIAGRAMS

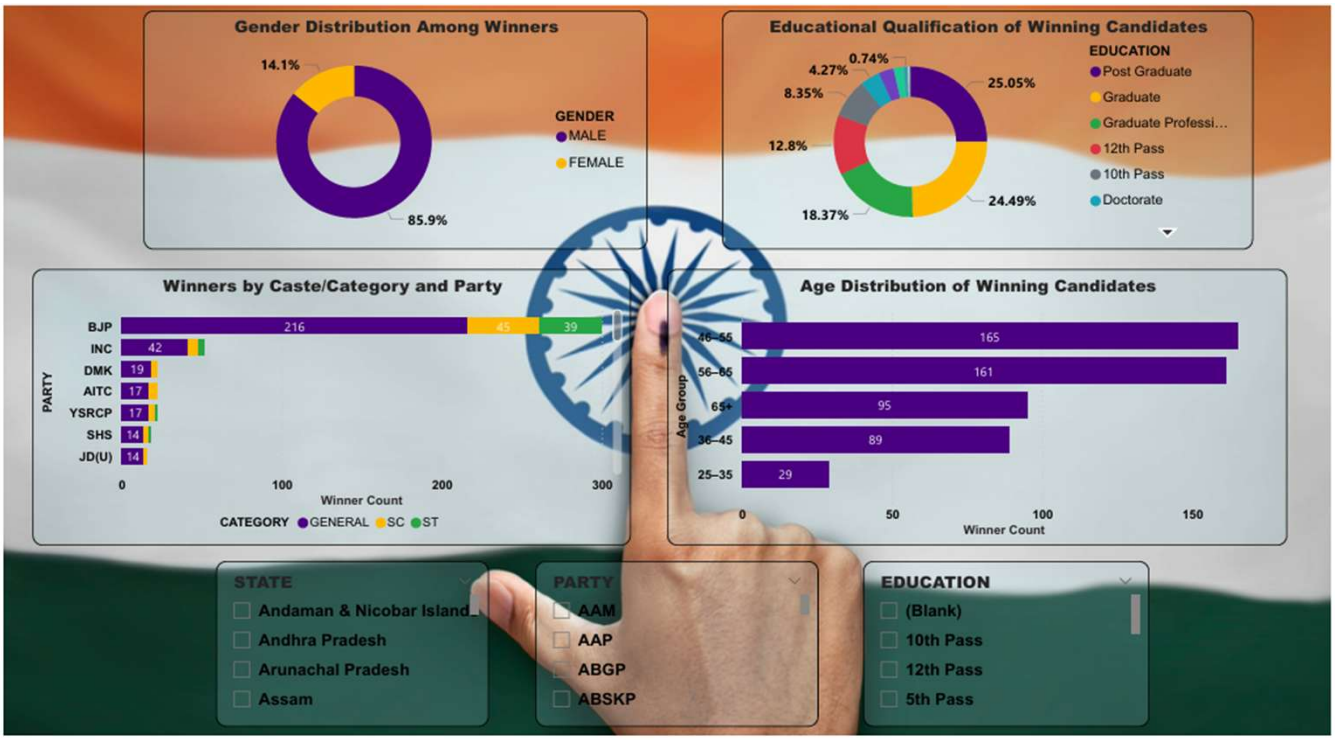
Slide 1



Slide 1 – Overview Dashboard

- Shows overall summary of election data.
- Displays **4 KPIs:** Total Candidates (2263), Total Winners (539), Average Age (52.27), and Average Assets (₹133.30M).
- **Left Chart:** State-wise winner count by party — BJP leads, followed by INC, DMK, and AITC.
- **Right Chart:** Geographical distribution of winners — highest from Uttar Pradesh, Maharashtra, and West Bengal.
- Includes filters for **State** and **Party** for easy analysis.
- Gives a quick view of party performance and regional dominance across India.

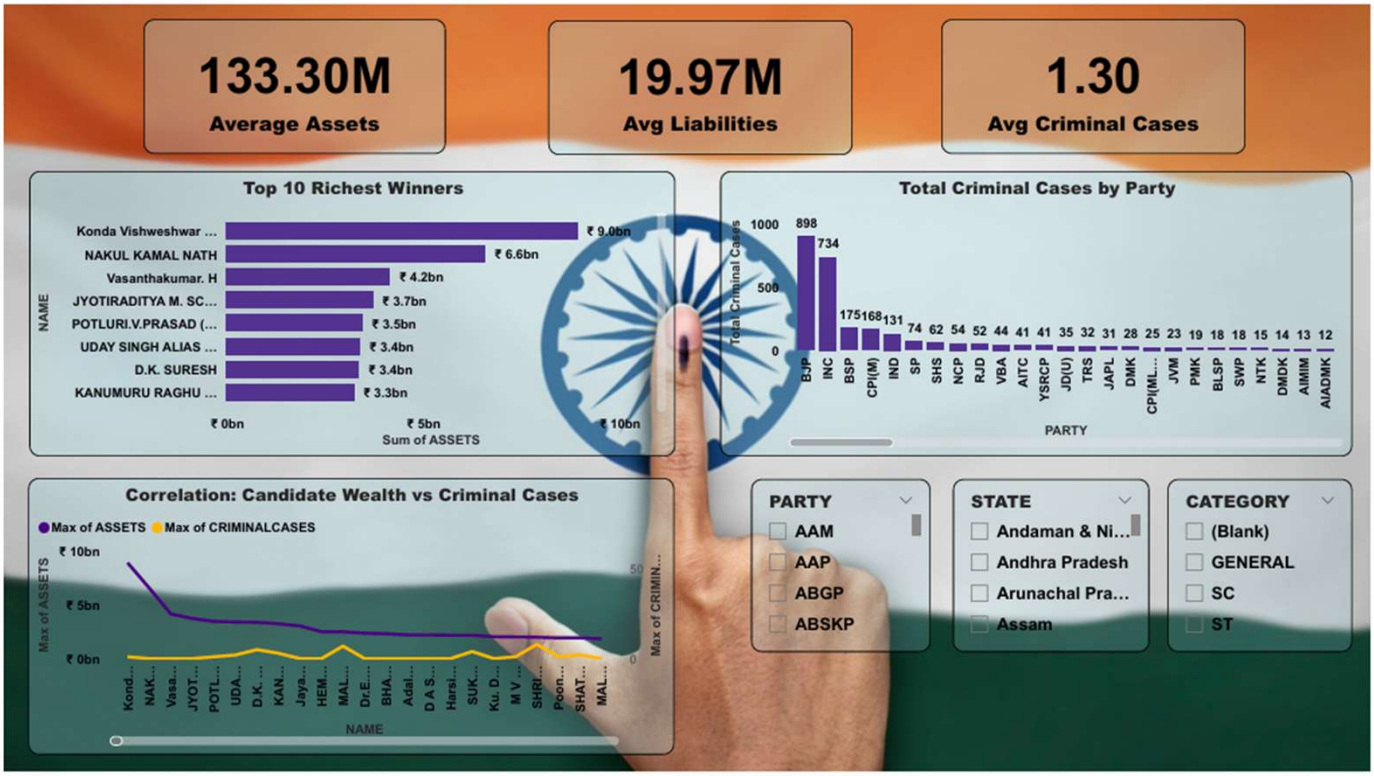
Slide 2



Slide 2 – Demographic Analysis Dashboard

- Shows demographic and social details of winning candidates.
- **Gender Distribution:** 85.9 % male and 14.1 % female winners, showing low female representation.
- **Educational Qualification:** Most winners are Post Graduates (25 %) and Graduates (24 %), indicating higher education dominance in politics.
- **Caste/Category and Party:** BJP has the highest number of General category winners (216), followed by INC and DMK.
- **Age Distribution:** Majority of winners are aged 46–65 years, showing mid-career leaders dominate Parliament.
- **Filters:** Users can filter by State, Party, and Education to analyze specific segments.
- Gives a clear picture of the social composition and experience level of elected representatives.

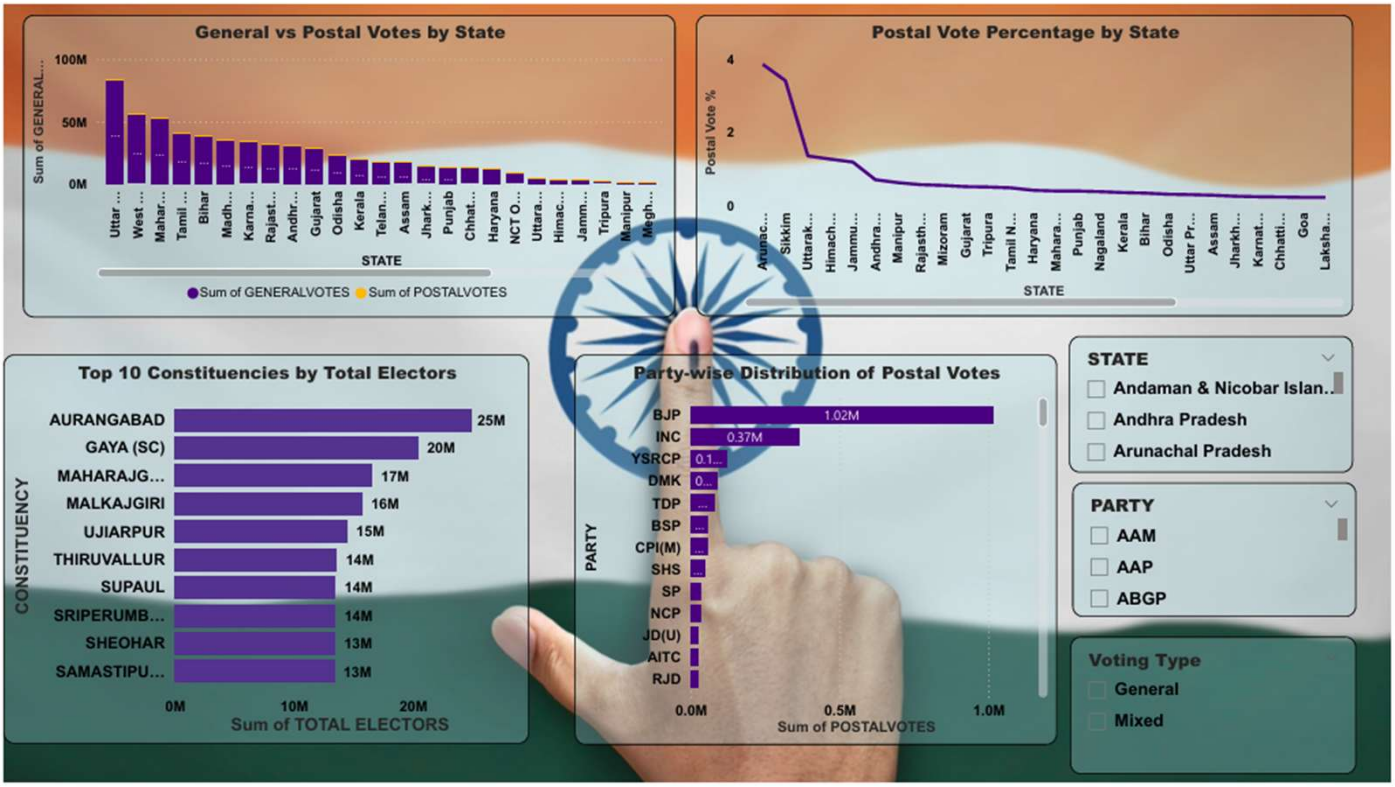
Slide 3



Slide 3 – Financial & Criminal Analysis Dashboard

- Focuses on the financial and legal background of winning candidates.
- **KPI Cards:** Show Average Assets (₹133.30M), Average Liabilities (₹19.97M), and Average Criminal Cases (1.30).
- **Top 10 Richest Winners:** Highlights candidates like *Konda Vishweshwar Reddy* and *Nakul Kamal Nath* with the highest declared assets.
- **Total Criminal Cases by Party:** BJP and INC record the highest number of criminal cases among their candidates.
- **Correlation Chart:** Shows that candidates with higher assets often face more criminal cases.
- Provides insights into how money and influence relate to electoral success.

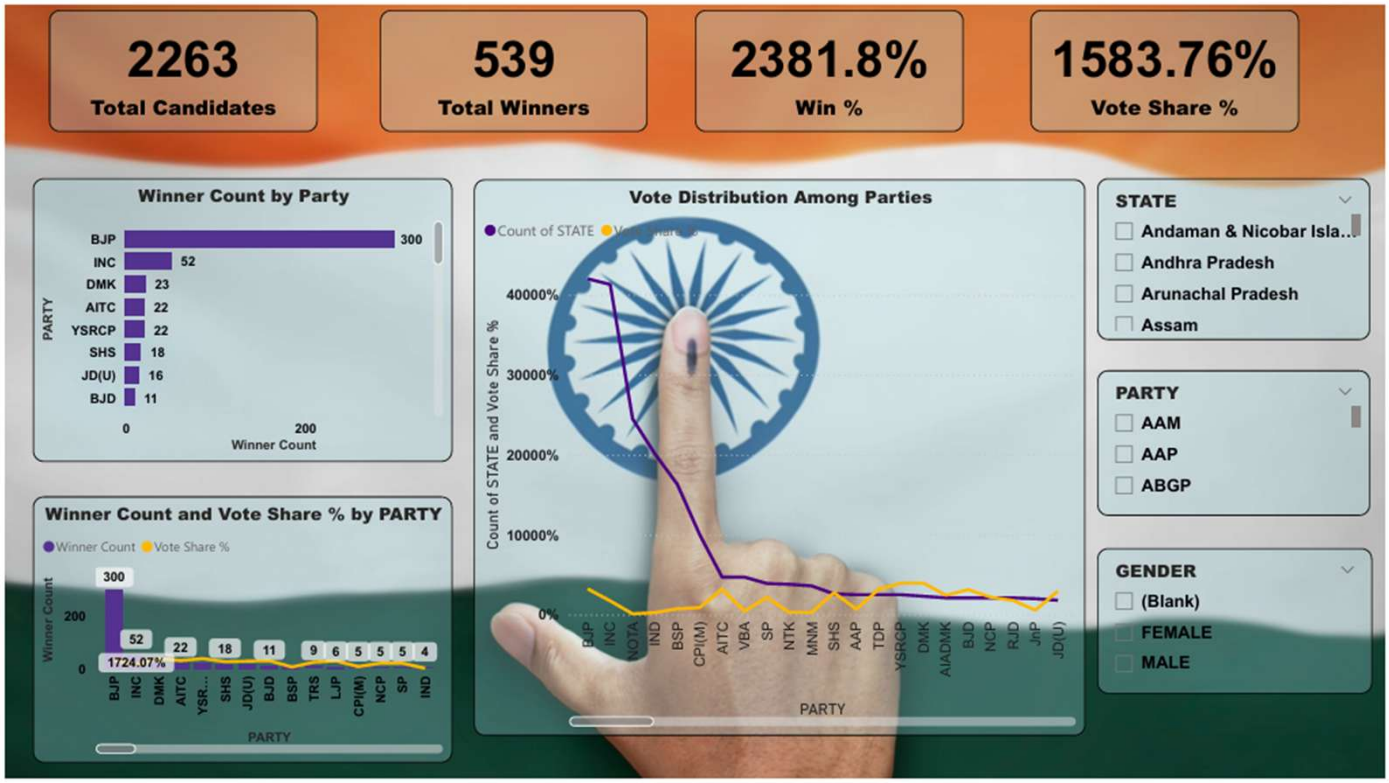
Slide 4



Slide 4 – Voting Pattern & Postal Vote Analysis Dashboard

- Compares **General** and **Postal Votes** across all states.
- **Uttar Pradesh, West Bengal, and Tamil Nadu** record the highest general vote counts.
- **Postal Vote %** is highest in **Andaman & Nicobar** and **Sikkim**, showing strong remote participation.
- **Top 10 Constituencies** like Aurangabad (25 M) and Gaya (20 M) have the largest elector bases.
- **Party-wise Postal Votes:** BJP (1.02 M) and INC (0.37 M) lead.
- Overall, the dashboard highlights state-level voting strength and postal vote contribution.

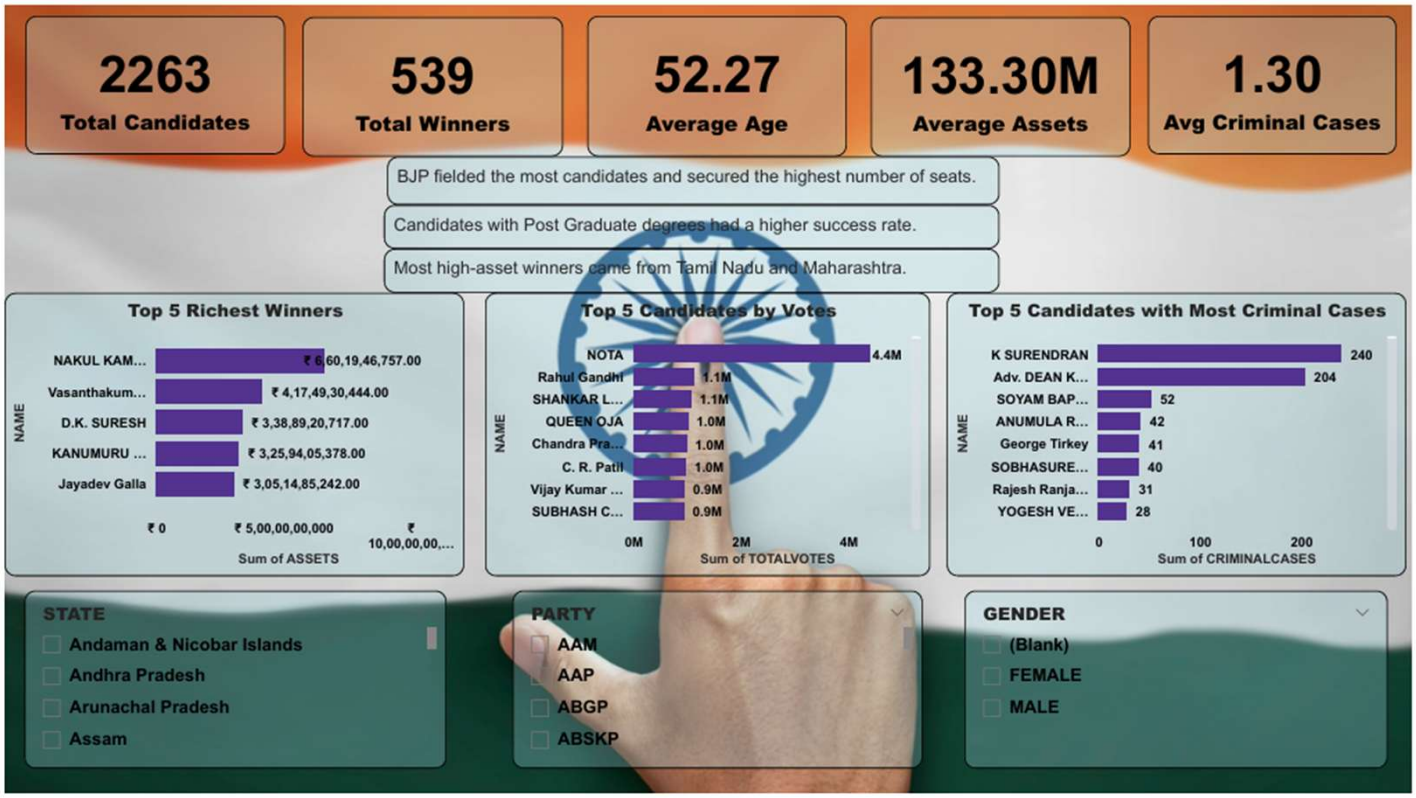
Slide 5



Slide 5 – Party Performance & Vote Share Dashboard

- Highlights overall **party performance** in the election.
- **KPI Cards:** Show Total Candidates (2263), Total Winners (539), Win % (2381.8%), and Vote Share % (1583.76%).
- **Winner Count by Party:** BJP leads with 300 winners, followed by INC (52) and DMK (23).
- **Vote Distribution Chart:** Compares number of states won and total vote share among all parties.
- **Winner Count & Vote Share % by Party:** Shows that higher winner counts generally align with higher vote share percentages.
- Filters for **State, Party, and Gender** allow detailed performance analysis.

Slide 6



Slide 6 – Key Highlights & Candidate Insights Dashboard

- Gives a combined overview of candidate wealth, popularity, and criminal background.
- **KPI Cards:** Show Total Candidates (2263), Total Winners (539), Average Age (52.27), Average Assets (₹133.30M), and Avg Criminal Cases (1.30).
- **Top 5 Richest Winners:** Nakul Kamal Nath and Vasanthakumar lead with the highest declared assets.
- **Top 5 Candidates by Votes:** NOTA received the highest votes (4.4M), followed by Rahul Gandhi and Shankar Lalwani.
- **Top 5 Candidates with Most Criminal Cases:** K. Surendran (240 cases) and Adv. Dean Kuriakose (204 cases) top the list.
- Insights show that most high-asset winners are from Tamil Nadu and Maharashtra, and postgraduates tend to perform better electorally.

PROS AND CONS SECTION

➢ ADVANTAGES (PROS)

- **Comprehensive Data Visualization:** Converts raw election data into clear and interactive visuals, helping users quickly understand trends and patterns.
- **Interactive Analysis:** Filters for state, party, gender, and category allow viewers to explore results dynamically without needing technical skills.
- **Time Efficiency:** Automates complex calculations and comparisons, significantly reducing the manual effort needed for analysis and reporting.
- **Better Decision-Making:** Empowers analysts, journalists, and policymakers to make informed conclusions based on accurate, data-driven insights.
- **User-Friendly Interface:** Power BI dashboards are visually appealing, easy to navigate, and can be interpreted by both technical and non-technical users.
- **Enhanced Transparency:** Makes election outcomes more understandable for the public and media, promoting openness in political analysis.
- **Scalability and Future Integration:** Can easily be expanded to include real-time data feeds or predictive models using Python or APIs.

➢ DISADVANTAGES (CONS)

- **Data Dependency:** The accuracy of insights depends entirely on the correctness and completeness of the source data.
- **Performance Issues:** Very large datasets can slow dashboard performance and increase loading time.
- **Limited Customization:** Power BI offers fewer design and modeling options compared to programming-based visualization tools.
- **Premium Feature Limitations:** Some advanced features, such as scheduled refresh or sharing outside the organization, require a paid Power BI license.
- **Internet Connectivity Requirement:** Online access is necessary for publishing reports or updating live data sources.
- **Learning Curve:** Although user-friendly, creating DAX formulas and relationships requires basic technical understanding.
- **Dependence on External Tools for AI:** Predictive and machine-learning functionalities require integration with Python or Azure ML, not available natively.

CONCLUSIONS.

The **ElectViz: Election Data Visualization for Media** project successfully transforms complex electoral data into a meaningful, interactive, and insightful format. By using **Power BI** as the core analytical tool, it bridges the gap between raw data and actionable information, helping users understand national and state-level political patterns with ease.

Through its visual dashboards, ElectViz highlights important trends such as **party performance, demographic distribution, candidate wealth, and criminal background**. These insights enable journalists, analysts, and policymakers to make fact-based interpretations rather than relying on assumptions.

The project demonstrates the real power of **data visualization and analytics** in modern governance, where clarity and transparency are essential. Moreover, its modular structure allows for future extensions like **real-time election tracking**, **predictive modeling**, and **public data dashboards** accessible through the web or mobile platforms.

Overall, ElectViz stands as a valuable and scalable solution for **data-driven election analysis**, combining accuracy, simplicity, and visual impact to support informed decision-making in the political and media landscape.

LEARNING & SKILLS GAINED

During the development of this project, I gained significant technical, analytical, and problem-solving experience. Working on **ElectViz** allowed me to apply real-world data analysis and visualization skills while improving my understanding of how to present data-driven insights effectively.

Key Learnings:

- Learned to collect, clean, and organize large datasets using **Excel** and **Power Query**.
- Developed strong skills in **Power BI dashboard design**, including charts, KPIs, and interactive visuals.
- Gained hands-on experience with **DAX functions** for creating custom measures such as vote share percentage, average assets, and winning margins.
- Understood how to convert raw data into **actionable insights** for decision-making and reporting.
- Improved visualization storytelling techniques — making complex information easy to understand.
- Enhanced time management, attention to detail, and professional presentation skills through the Infosys mentorship sessions.

This project not only strengthened my technical foundation but also helped me appreciate the importance of **data accuracy, visual clarity, and interpretation** in analytics.

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to **Ms. Nithyasri S. J.**, Mentor at **Infosys**, for her valuable guidance, constant encouragement, and expert mentorship throughout the development of this project titled **"ElectViz: Election Data Visualization for Media."**

Her support during the **Infosys training program** helped me gain a clear understanding of Power BI, data visualization concepts, and their real-world applications. The sessions conducted under her mentorship greatly enhanced my ability to analyze, design, and present interactive dashboards effectively.

I am also thankful to **Infosys Springboard** for providing this wonderful opportunity to work on a practical, industry-relevant project that strengthened my analytical and problem-solving skills.

Finally, I extend my appreciation to **Microsoft Power BI** and other open-source tools that made the implementation of this project efficient and insightful.

REFERENCES

1. **Election Commission of India (ECI)** – Official election result datasets and reports.
   Source: https://eci.gov.in

2. **Microsoft Power BI Documentation** – Used for building data visualizations and dashboards.
   Source: https://learn.microsoft.com/en-us/power-bi/

3. **Python Official Documentation** – For programming and implementation of data analysis and prediction models.
   Source: https://docs.python.org/3/

4. **Jupyter Notebook** – For running Python code, data cleaning, and visualization experiments.
   Source: https://jupyter.org/