A Project Report on

Electviz: Data Visualization and Analysis of Elections

Infosys Springboard Virtual Internship 6.0
Group 1
Batch 11
Team D

Submitted by

1. Neelapu Herambudu
2. Francis Daniel
3. Swathika Murugan
4. Shaik Shabeena Praveen
5. Gopi Krishna
6. Tanmay Parulkar

Internship Duration: 8 weeks (2 months)
2026

1.Project Title:

Electviz: Data Visualization and Analysis of Elections

2.Project Objective:

The objective of ElectViz is to transform raw U.S. election data into a clear, interactive, and trustworthy visual story. The project aims to help users, students, analysts, media teams, and the general public, quickly interpret voting patterns, party performance, and state-wise outcomes using an intuitive Power BI dashboard. By simplifying a complex national dataset, ElectViz enables fast insights, evidence-based conclusions, and more informed democratic understanding.

3.Project Description:

ElectViz is a data visualization project designed to improve how election results are understood by the media and the public. It is an interactive data visualization project designed to analyze, interpret, and present election data in a clear and engaging way.

The problem addressed by this project is that election data is often presented in complex tables and articles, making it difficult for audiences to quickly grasp trends, regional patterns, and key outcomes.

Developed using Power BI, this project transforms large-scale election datasets into intuitive dashboards that help media organizations, analysts, and the public gain meaningful regions and parties insights into voting patterns, party performance, and regional election dynamics. The project demonstrates how data analytics and visualization can enhance political reporting by offering fact-based, data-driven stories. By simplifying complex datasets into visual narratives, ElectViz empowers users to make informed interpretations of election outcomes such as :

- Quick identification of key trends and close races
- Improves audience engagement and trust
- Faster election result analysis
- Easy comparison across

3.1 Project Overview

ElectViz is a multi-page Power BI dashboard developed to analyze historical election data spanning several decades. The dataset contains information related to election year, state, party affiliation, vote counts, vote share, candidates, and

seat allocation.
The dashboard enables:
* Macro-level analysis (national trends)
* Meso-level analysis (state and party comparisons)
* Micro-level analysis (candidate performance)
Through interactive filters and visual elements, ElectViz supports exploratory analysis and informed interpretation of electoral patterns.

3.2 Project Approach

Our team followed a structured and iterative workflow to convert a messy, inconsistent US Election dataset into a fully interactive analytical dashboard.

Step 1 – Understanding the Raw Dataset

- The initial dataset contained several issues
- Missing district names
- Mixed party abbreviations and full names
- Incorrect vote share values (e.g., 0.01 instead of 1%)
- Duplicate and inconsistent state formatting
- Fragmented election type fields (special, runoff, fusion, etc.)

The team first analyzed the structure to identify all cleaning and transformation steps needed.

Step 2 – Data Cleaning & Preparation (Power Query)

Using Power Query, the team:

- Standardized state names and party labels
- Removed duplicates and unused columns
- Handled missing values and inconsistent formatting
- Created a reliable base table ready for modeling
- Ensured numerical fields like candidate votes and total votes were properly typed

This step ensured that the dashboard visuals produced correct insights

Step 3 – Data Modeling & Measure Creation (DAX)

Custom DAX measures were created to support analysis:

- Total Votes

- Total Candidates
- Average Vote Share
- Seats Won by Party
- Year-over-Year Vote Trends
- Party Vote Share %

DAX also fixed logical issues created by the raw dataset (e.g., recalculating correct percentages).

Step 4 – Designing Visual Pages

Three dashboard pages were built with a clear structure:
Page 1 – Election Overview & Trends
Page 2 – Special Elections & Fusion Ticket Analysis
Page 3 – State-wise Candidate Vote Distribution
Page 4 – Party & Candidate Performance
Page 5 – Seat Distribution by Party and State
Page 6 – State Wise Election Participation
Page 7 –Election Results-Final Overview

Step 5 – Usability & Visual Refinements

We refined:

1) Map color contrast
2) Data label visibility
3) Slicer interactivity
4) Clean layout spacing
5) Improved title clarity & page names

3.3 Technology Used

Dataset:

- Microsoft Excel (US_House_Election.csv file)

Dashboard:

- Power BI
- Power Query (M Language)
- DAX (Data Analysis Expressions)

Documentation:

- PDF
- GitHub

3.4 Insights from dashboard

These insights reflect our actual ElectViz visuals:

- **Voting Trends Over Time**

1. Total votes show long-term growth, indicating increasing voter participation.
2. Presidential election years show significant spikes compared to midterms.

- **Party Dominance Patterns**

1. Republican and Democratic parties hold the largest share of votes consistently.
2. Third parties like Libertarian and Independent show niche but stable contributions.
3. Certain states exhibit long-term loyalty to a specific party.

- **State-Level Distribution**

1. Coastal and high-population states contribute significantly larger vote volumes.
2. Several central and mountain-region states show more competitive, shifting trends.

- **Candidate-Level Insights**

1. Top-performing candidates cluster heavily in key battleground states.
2. Candidates with fusion tickets or multiple party endorsements show distinct vote patterns.

- **Election Type Behavior**

1. Special elections and runoffs have substantially lower participation.
2. Vote share and margins fluctuate more heavily in these non-standard elections.

3.5 REAL-WORLD IMPACT (MEDIA & PUBLIC COMMUNICATION)

For Newsrooms & Media

1. Journalists can rapidly filter by year, state, and party for instant storytelling.
2. The visualization layout replicates real-time election-night dashboards.

3. Helps decode complex patterns like vote swings, seat changes, and party shifts.
4. Converts thousands of messy records into clean, broadcast-ready graphics.

For the General Public

1. Simplifies election results for non-experts.
2. Builds transparency in democratic processes.
3. Enables citizens to explore historical patterns, turnout, and vote share.
4. Helps students and analysts understand long-term political shifts.

For Researchers & Policy Analysts

1. Generates insights on party performance history.
2. Useful for studying turnout patterns and electoral behaviors.
3. Can be extended to demographic analysis for deeper political research.

4. Timeline Overview:

Weekly Activit*ies*	**Activities Achieved**	**Activities Summary**
Week 1	Dataset Collection and Understanding	Collected the U.S. election dataset and understood its structure, variables, and scope for analysis.
Week 2	Data Cleaning and Transformation	Cleaned missing values, standardised formats, and transformed data to ensure consistency and accuracy.
Week 3	Data Modelling and DAX measure creation	Created data relationships, built the data model, and developed required DAX measures for analysis.
Week 4	Dashboard Development and Optimisation	Designed interactive dashboard pages and optimised visuals for performance and clarity.
Week 5	Insight validation and Analysing	Validated insights, cross-checked calculations, and analysed trends for

		meaningful interpretation.
Week 6	Dashboard Completion and Documentation	Finalized dashboard pages and prepared detailed project documentation.
Week 7	Optimization and documentation	Improved dashboard performance, refined visuals, and enhanced documentation quality.
Week 8	Documentation and Final Presentation	Completed final documentation and prepared the project for presentation and submission.

5a. Key Milestones:

Name	Description	Date Achieved
Milestone 1	Focused on understanding project objectives and preparing the election dataset for analysis. Performed data cleaning and preprocessing to ensure accuracy, consistency, and reliability.	09/01/2026
Milestone 2	Developed initial dashboard pages including Election Overview, Special Elections, and State-wise Vote Distribution. Provided high-level trends and regional insights into voting patterns.	16/01/2026

Milestone 3	Built advanced analytical pages to evaluate party dominance, candidate performance, and seat distribution. Enabled deeper insights into political control and power shifts over time.	23/01/2026
Milestone 4	Completed the dashboard with voter participation analysis and consolidated results overview. Focused on summarizing insights and improving storytelling for clear presentation.	30/01/2026

Milestone 1: Project Initialisation, Dataset Setup, Data cleaning and Preprocessing

DATASET DESCRIPTION:

The dataset used in this project is a **United States House of Representatives election dataset** extracted from the file *house.csv*. It contains 32,452 records and 20 columns. Each row represents one candidate's performance in a specific congressional election. The dataset includes information about the election year, state, congressional district, election stage, political party, and vote counts. This dataset is suitable for analyzing election trends, party performance, candidate success, and voting behavior across time and geography.

- **Source: U.S. House Election Dataset (house.csv)**
- **Time Period: 1976 – 2018**
- **Granularity: Candidate-level election data**

COLUMN INFORMATION:

The dataset contains the following key columns:

- **year**-Represents the election year in which the congressional race took place.
- **state**-The full name of the U.S. state where the election occurred.
- **state_po**-The two-letter postal abbreviation of the state (e.g., CA, TX, NY).

- **district**-The congressional district number for the election.
- **office**-Indicates the political office being contested (House of Representatives).
- **stage**-Specifies the election stage, such as General, Primary, or Runoff.
- **runoff**-Indicates whether the election was a runoff election. Some values were missing in the original data.
- **special**-A Boolean column indicating whether the election was a special election.
- **candidate**-The name of the candidate who contested the election.
- **party**-The political party affiliation of the candidate.
- **candidate votes** -The number of votes received by the candidate.
- **total votes**-The total number of votes cast in that election race.
- **fusion ticket**-Indicates whether the candidate ran under a fusion ticket.
- **state_fips, state_cen, state_ic, version**-These are technical or geographic identifier fields that are not required for analytical purposes.

Milestone2: Dashboard development

Page 1: Election Overview and Trends

Purpose of the Page

This page provides a high-level summary of election data across years, enabling users to understand overall voting patterns and participation trends.

Visuals Used

1. **KPI Cards**
    - Total Votes Cast
    - Total Candidates Participated
    - Average Vote Share
    - Purpose: Quickly summarize key election statistics**.**
2. **Line Chart / Area Chart (Voting Trends Over Time)**
    - Displays vote counts across different election years.
    - Helps track participation growth over time.
3. **Map Visualization**
    - Shows geographic vote distribution across states.
    - Highlights regions with higher voting activity.

4. **Slicers**
    - Year filter
    - Party filter
    - State filter
    - Allow users to interactively explore trends**.**

Insights from Page 1

- Voter participation increases steadily over time.
- Presidential election years show higher vote counts compared to midterm elections.
- Large population states consistently contribute the highest votes.
- Certain states demonstrate consistent party dominance over decades.
- Users can easily detect participation fluctuations across election cycles.

Page 2: Special Elections and Fusion Ticket Analysis

Purpose of the Page

This page focuses on special election types and fusion ticket participation, helping analyze less common election scenarios.

Visuals Used

1. **Bar/Column Chart – Election Type Distribution**
    - Displays counts of election types such as general, special, runoff, etc.
    - Shows frequency of each election type.
2. **Stacked Bar Chart – Party Performance in Special Elections**
    - Compares party performance in special or runoff elections.
3. **Fusion Ticket Participation Visual**
    - Displays candidates supported by multiple parties.
    - Shows vote share distribution in fusion cases.
4. **Interactive Filters**
    - Election type slicer

- Year and party filters for deeper analysis.

Insights from Page 2

- Special elections occur less frequently compared to general elections.
- Voter turnout is usually lower in special elections.
- Vote margins fluctuate more in runoff or special contests.
- Fusion tickets occasionally impact vote share distribution**.**
- Certain parties benefit more from coalition or fusion support**.**

Page 3: State-wise Candidate Vote Distribution

Purpose of the Page:

This page analyzes how candidate votes are distributed across states and parties.

Visuals Used:

1. **Stacked Bar Chart – Vote Distribution by State**
   - Shows vote counts of candidates across states.
   - Allows comparison between party performances.
2. **Table Visualization**
   - Displays candidate names, party affiliation, votes received, and vote share.
   - Helps identify top candidates.
3. **State Map Visualization**
   - Shows geographic voting distribution.
   - Highlights competitive and dominant regions**.**
4. **Slicers**
   - State filter
   - Party filter
   - Candidate filter

Insights from Page 3

- Vote distribution varies significantly across states.

- Battleground states show close competition between parties.
- Some states consistently support one dominant party.
- High-performing candidates are concentrated in politically competitive states.
- Smaller parties perform better in specific regions rather than nationally.

Page 4: Party & Candidate Performance

Purpose of the Page

This page evaluates how political parties and individual candidates perform across elections, helping users identify leading candidates and dominant parties.

Visuals Used

1. **Bar Chart – Top Performing Candidates**
    - Displays candidates with the highest votes received.
    - Helps identify influential candidates and strong constituencies.
2. **Party-wise Vote Share Chart**
    - Compares vote share percentages among parties.
    - Shows overall party strength.
3. **Candidate Performance Table**
    - Lists candidates, parties, votes received, and vote share.
    - Enables detailed candidate-level comparison.
4. **Slicers**
    - Year filter
    - Party filter
    - State filter
    - Enables focused performance analysis.

Insights from Page 4

- Major parties consistently secure higher vote shares.
- Certain candidates repeatedly perform strongly across elections.

- Candidate success often correlates with party strength in specific states.
- Smaller parties succeed in limited regions but struggle nationally.
- Competitive states produce closely matched candidate performances.

Page 5: Seat Distribution by Party and State

Purpose of the Page

This page analyzes how seats are distributed among parties across states, helping users understand political dominance and representation patterns.

Visuals Used

1. **Map Visualization – Seat Distribution**
    - Shows which party dominates in each state based on seat wins.
2. **Stacked Bar Chart – Seats Won by Party**
    - Displays seat counts secured by different parties.
3. **State-wise Seat Distribution Table**
    - Provides detailed seat allocation by party and state.
4. **Filters**
    - Year and party slicers allow comparison across elections.

Insights from Page 5

- Two major parties dominate most states.
- Some states consistently support one party over multiple elections.
- Battleground states frequently change party dominance.
- Seat distribution sometimes differs from vote share due to electoral structures.
- Regional political loyalty strongly impacts seat allocation.

Page 6: State-wise Election Participation

Purpose of the Page

This page compares voter participation levels and competitiveness across states.

Visuals Used

1. **Bar Chart – Total Votes by State**
    - Displays voter participation levels across states.
2. **Stacked Chart – Seats vs Votes Comparison**
    - Compares total votes cast with seats won.
3. **State Comparison Table**
    - Shows vote totals, participation levels, and party performance.
4. **Interactive Slicers**
    - Year and party filters for dynamic analysis.

Insights from Page 6

- Larger states consistently show higher participation.
- Certain states show growing participation over time.
- Competitive states show narrow vote margins.
- Smaller states may have fewer votes but strong party dominance.
- Participation trends reflect demographic and political shifts.

Page 7: Election Results – Final Overview

Purpose of the Page

This final page summarizes overall election outcomes, providing a complete overview of results and performance indicators.

Visuals Used

1. **KPI Cards**
    - Total Votes
    - Seats Won
    - Winning Party
    - Candidate Count
2. **Stacked Bar Chart – Party Results Comparison**
    - Displays total performance of parties.

3. **Table View – Final Results Summary**
    - Provides state-wise or party-wise result summaries.
4. **Slicers**
    - Year and party filters for quick comparisons.

Insights from Page 7

- Provides a quick understanding of election winners and overall outcomes.
- Highlights which party dominated overall results.
- Shows how seats and votes translate into final political control.
- Allows rapid comparison between election years.
- Serves as the dashboard's executive summary.

5b. Project Execution Details

The project was executed using an iterative development methodology, where each phase of the dashboard was built, reviewed, and refined progressively. After completing every set of visuals, feedback was incorporated to improve clarity, accuracy, and analytical depth.

- Step 1 – Understanding the Raw Dataset: The team analyzed the structure of the messy US Election dataset to identify necessary cleaning and transformation steps for issues like missing district names and inconsistent party labels.
- Step 2 – Data Cleaning & Preparation (Power Query): Using Power Query, the team standardized state and party names, removed duplicates, and ensured numerical fields were properly typed to create a reliable base table.
- Step 3 – Data Modeling & Measure Creation (DAX): Custom DAX measures, such as Total Votes and Seats Won by Party, were created to support complex analysis and fix logical issues in the raw data.
- Step 4 – Designing Visual Pages: A multi-page dashboard was built to cover various analytical depths, including election trends, special elections, and state-wise candidate distributions.
- Step 5 – Usability & Visual Refinements: The final dashboard was polished by improving map color contrast, data label visibility, and slicer interactivity to enhance clarity and user experience.

Technical Skills:

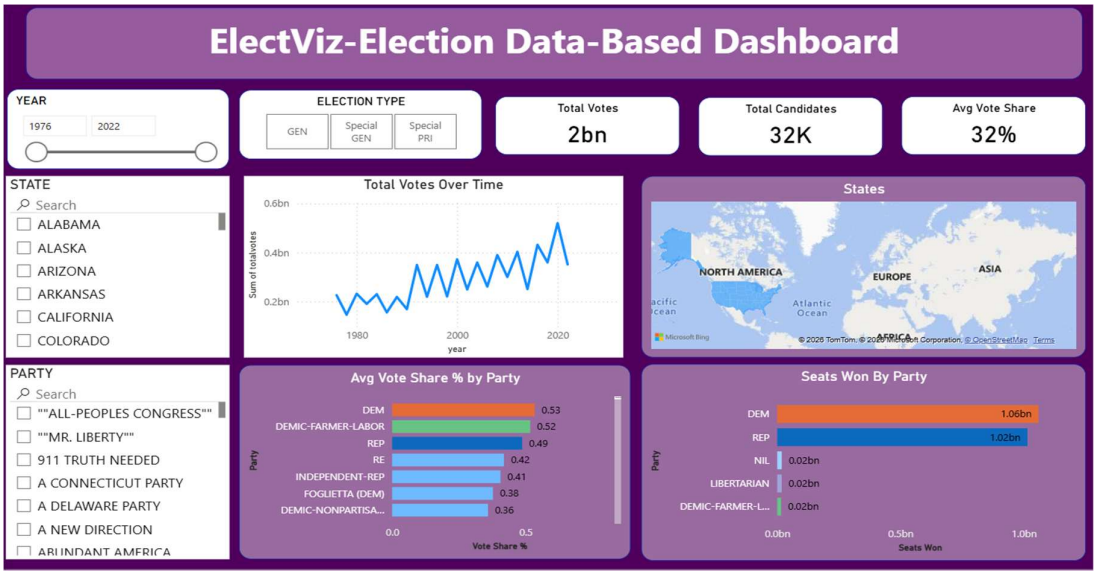
- Data cleaning and preprocessing techniques in Power BI.
- Data transformation using Power Query Editor.
- Data modeling and creating relationships between tables.
- Writing and optimizing DAX measures and calculated columns.
- Creating interactive dashboards and reports.
- Using different visualizations (Bar, Line, Map, KPI, Donut, etc.).
- Applying filters, slicers, drill-through, and tooltips.
- Performance optimization of dashboards.
- Insight validation and analytical interpretation of election data.
- Report documentation and presentation preparation.

Soft Skills:

- Improved communication skills while discussing insights and reviews.
- Enhanced analytical thinking and problem-solving ability.
- Better time management through weekly milestone completion.
- Team collaboration and coordination skills.
- Increased confidence in presenting technical work.
- Ability to take initiative and work independently.
- Attention to detail in data accuracy and validation.
- Improved storytelling and data interpretation skills.

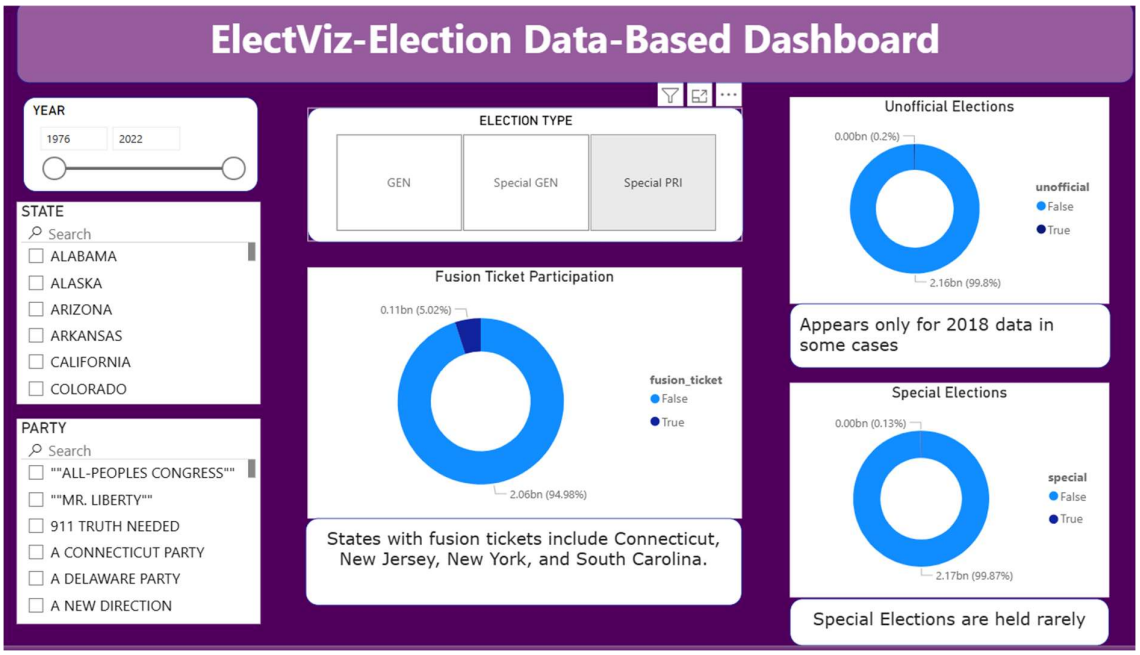
6. Snapshots / Screenshots – Page-wise Explanation

6.1 Election Overview and Trends:



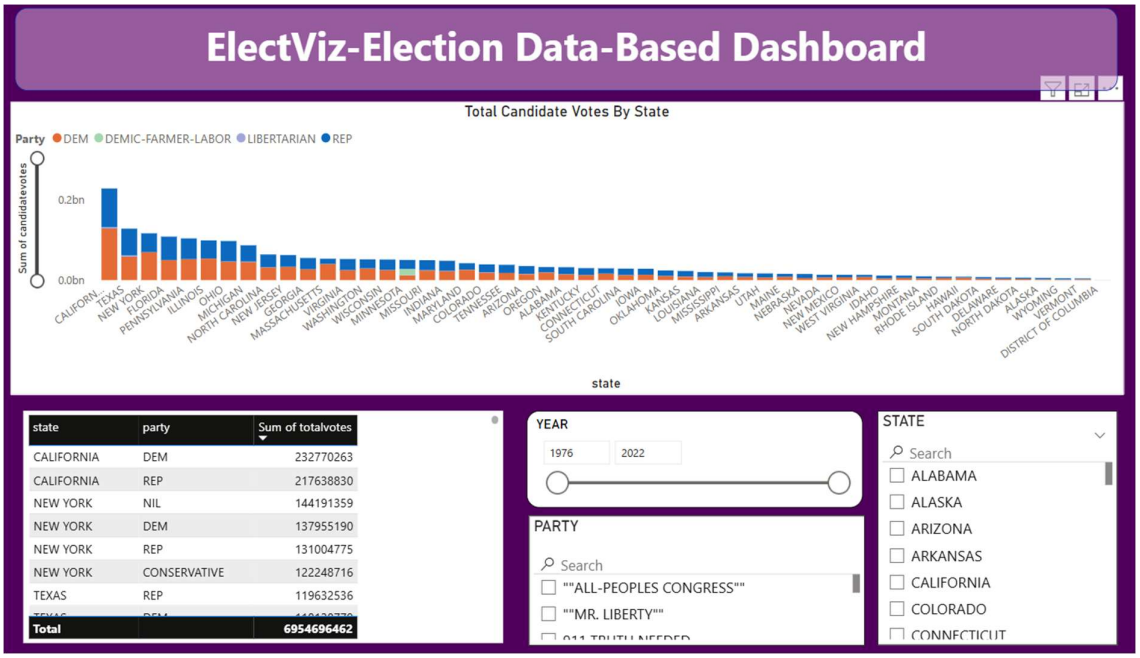
Summarizes key metrics such as total votes, candidate count, and average vote share. Provides temporal and geographic insights into voter participation**.**

6.2 Special Elections and Fusion Ticket Analysis



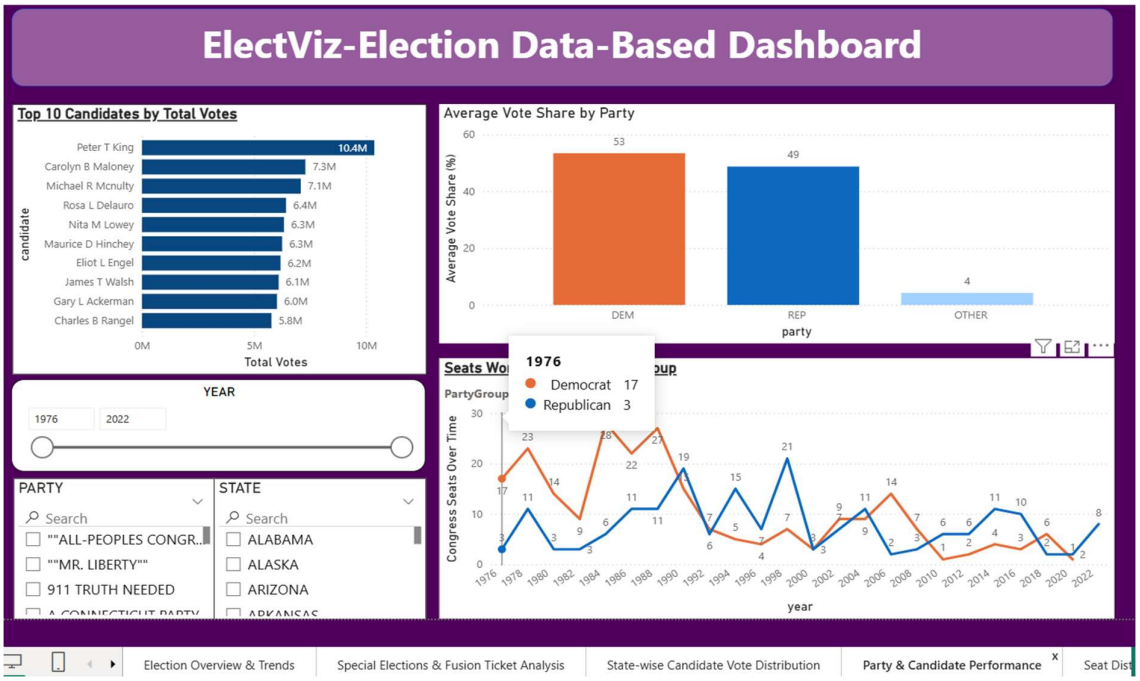
Analyzes the frequency and distribution of special elections and fusion tickets, highlighting their limited occurrence.

6.3 State-wise Candidate Vote Distribution



Examines vote distribution across states and parties, enabling comparative analysis of regional political behavior.

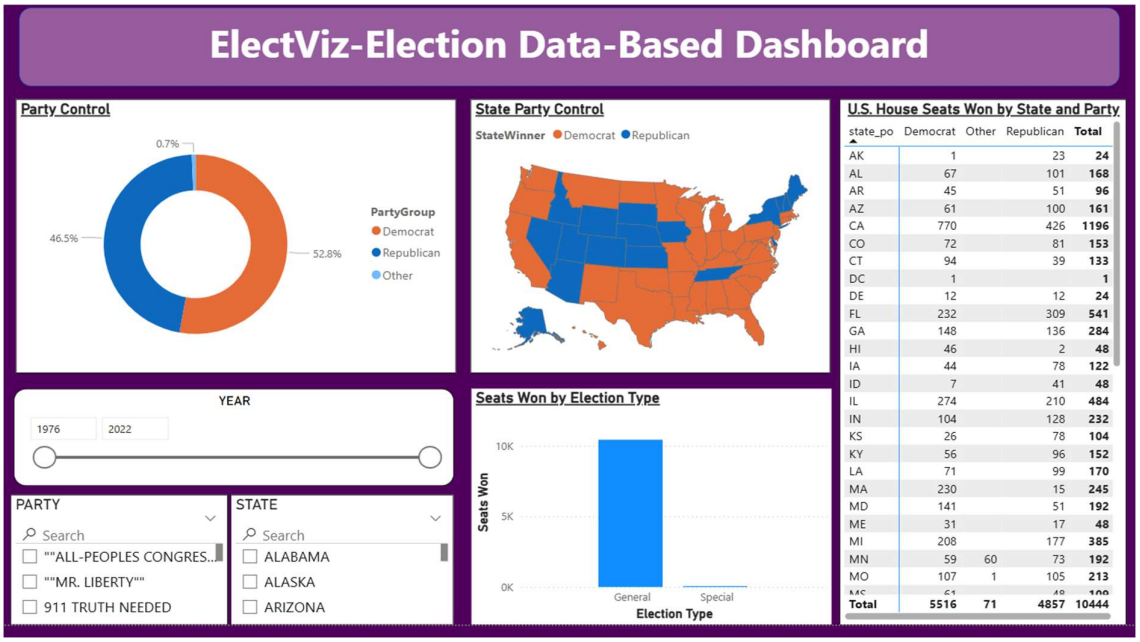
6.4 Party and Candidate Performance



Identifies top-performing candidates and evaluates average vote share across

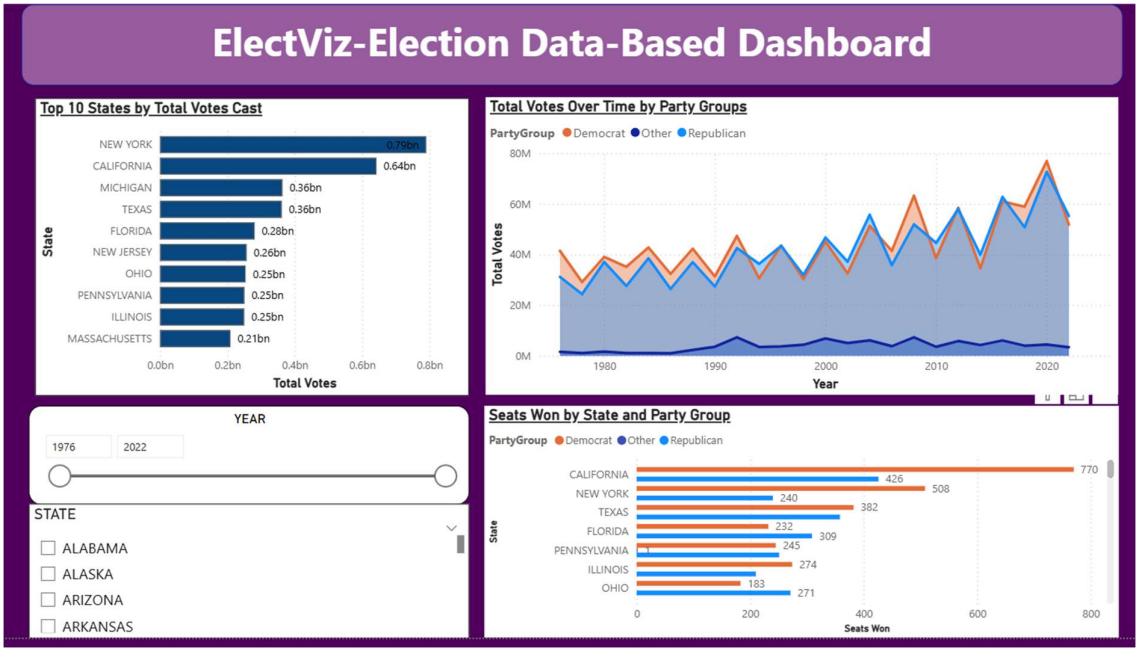
parties.

6.5Seat Distribution by Party and State



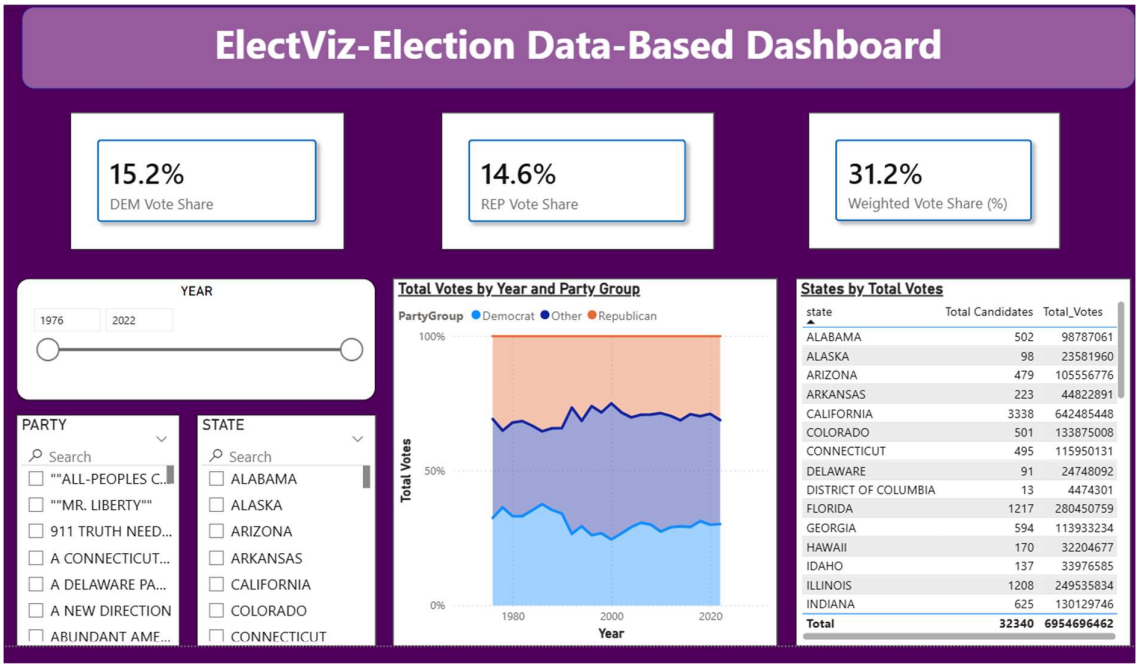
Illustrates party dominance at national and state levels through maps and tables.

6.6State-wise Election Participation and Comparison



Compares total votes and seat wins across states and party groups.

6.7Vote Share and Party Group Analysis



Analyzes weighted vote share and long-term party performance trends.

7. Challenges Faced

7.1 Managing Heterogeneous Election Data

The election dataset contained records from multiple years, states, and parties with varying formats and structures, making it challenging to ensure uniformity and analytical consistency across the dashboard.

7.2 Designing Meaningful Data Relationships

Establishing correct relationships between fact tables and multiple dimension tables required careful validation to avoid incorrect aggregations and misleading insights.

7.3 Handling High Cardinality Dimensions

Large numbers of parties and candidates increased visual complexity and

affected dashboard readability, requiring thoughtful aggregation and filtering strategies.

7.4 Balancing Visual Detail with Interpretability

Including comprehensive information while maintaining visual clarity was challenging, as excessive detail could overwhelm users and dilute key insights.

7.5 Ensuring Analytical Accuracy Across Filters

Maintaining consistent and accurate results when multiple slicers (year, state, party, election type) were applied simultaneously required extensive testing and refinement.

8. Learnings and Skills Acquired

8.1 Advanced Data Structuring Techniques

Developed a deeper understanding of how to structure and normalize complex datasets to support scalable and efficient analytical models.

8.2 Contextual Data Interpretation

Learned to interpret numerical outputs within the broader political and historical context, improving the quality and relevance of insights generated.

8.3 Visualization Strategy and Design Thinking

Gained experience in selecting appropriate visualizations based on the nature of data and the analytical question being addressed.

8.4 Performance-Aware Dashboard Development

Acquired knowledge on optimizing calculations, reducing unnecessary visuals, and improving overall dashboard responsiveness.

8.5 Analytical Storytelling for Decision Support

Learned how to connect individual visuals into a coherent narrative that guides users toward data-driven understanding and informed conclusions.

9. Testimonials from Team

Neelapu Herambudu

“Working on ElectViz strengthened my understanding of data modeling and visualization while giving me real-world experience in transforming complex election data into meaningful insights.”

Francis Daniel

“This project helped me improve my analytical thinking and Power BI skills. Collaborating with the team made the development process efficient and enjoyable.”

Swathika Murugan

“ElectViz allowed me to gain practical exposure to dashboard design and data storytelling, helping me understand how visuals can simplify complex datasets.”

Shaik Shabeena Praveen

“Through this project, I enhanced my skills in data cleaning and visualization while learning the importance of teamwork in delivering impactful analytical solutions.”

Gopi Krishna

“The project provided hands-on experience in solving real data challenges and improved my ability to interpret and present analytical findings clearly.”

Tanmay Parulkar

“Working with the team on ElectViz improved both my technical and collaboration skills while giving me confidence in building professional analytical dashboards.”

These testimonials summarize the collaborative learning and professional growth achieved by the team during the ElectViz project development.

10. Conclusion

The ElectViz project successfully demonstrates how data visualization and analytics can simplify complex election datasets and transform them into meaningful insights for analysts, media professionals, and the general public. By developing an interactive Power BI dashboard, the project enables users to explore voting trends, party performance, state-wise participation, and candidate success through intuitive visuals and dynamic filters. The structured approach to data cleaning, modelling, and visualization ensured that the insights generated are accurate, accessible, and useful for understanding electoral behaviour over time.

Overall, the project highlights the importance of data-driven storytelling in modern political analysis and public communication. Through collaborative teamwork and iterative dashboard development, the team enhanced both technical and analytical skills while delivering a professional analytical solution. ElectViz not only improves election data interpretation but also serves as a scalable model for future extensions involving demographic analysis and advanced political research.

11. Acknowledgements & References

We would like to express our heartfelt gratitude to the Infosys Springboard Team for organizing this valuable internship opportunity, which provided us with a platform to apply our analytical and visualization skills in a real-world scenario. We also sincerely thank **Mrs. Nithyasri S J** for her continuous support, mentorship, and constructive feedback throughout the internship, which greatly helped in refining our project and enhancing our understanding of data analytics and visualization practices.

We would also like to acknowledge the collaborative efforts of our team members:

Neelapu Herambudu,
Francis Daniel,
Swathika Murugan,
Shaik Shabeena Praveen,
Gopi Krishna,
Tanmay Parulkar,

whose cooperation, dedication, and mutual support played a crucial role in successfully completing the ElectViz project. The teamwork, shared ideas, and collective problem-solving efforts significantly contributed to the development of an impactful analytical dashboard.

This internship has been a significant milestone in our professional and technical journey, allowing us to gain practical exposure to data analytics and dashboard development. The project development was supported by various learning resources, including the US Election Dataset sourced from Kaggle, Microsoft Power BI Documentation, and the Infosys Springboard Virtual Internship course materials, which provided essential guidance and technical knowledge throughout the project.