Infosys Springboard Virtual Internship 6.0 Completion Report

Team Details: Start date
4-SEP -2025

Names:

Vishnu Vardhan Reddy
Sree Raja Mounika
Ashutosh Roy
Angammal R

Internship Duration: 8 Weeks

1. Project Title

ElectViz Election Data Visualization for Media

2. Project Objective

The primary objective of the "Electviz" project is to strategically transform complex, raw electoral data into interactive, accessible, and comprehensible visual insights using Microsoft Power BI.

This initiative was developed to equip media professionals, journalists, and data analysts with a robust tool to visualize intricate electoral datasets with enhanced ease and accuracy. The project directly addresses the significant operational challenges of traditional election reporting, where processing large datasets in raw numeric formats often leads to critical delays, public confusion, and errors in data interpretation.

The core strategic aims of the project are as follows:

1. **To Automate and Expedite Analysis:** Replace time-consuming, manual data analysis workflows with an automated visualization dashboard that enables real-time insights.

2. **To Enhance Data Comprehension:** Convert complex and unwieldy tabular data into a suite of intuitive and interactive visuals, including charts, tables, and geospatial maps.

3. **To Provide Deeper Insights:** Leverage Power BI's Data Analysis Expressions (DAX) to create dynamic calculated measures, resolving the common lack of comparative insight found in static data.

Ultimately, the "Electviz" dashboard solution empowers media outlets to produce reliable, datadriven reports, rapidly identify emerging trends, and deliver faster, more accurate interpretations of election results, thereby supporting high-paced live news coverage and effective visual storytelling.

3. Project description in detail

- **Overview:**
  - **Electviz** is a Business Intelligence (BI) and data visualization solution designed to transform raw, complex electoral data into interactive, clear, and understandable visual insights. ○ It connects media professionals, journalists, and data analysts with processed, aggregated data on a single platform, enabling rapid, accurate, and data-driven reporting during live election coverage.
- **Technology Stack:**
  - The project is developed entirely within the **Microsoft Power BI Desktop** ecosystem, which serves as the all-in-one platform for data ingestion, transformation, modeling, and visualization.
  - **Data Source (Input):**
    - Utilizes static **CSV/Excel datasets** as the source for raw election data, including Party Details, Candidate Names, Votes Received, and State/Constituency information. ○ **Data Transformation & Modeling (ETL):**
    - **Power Query Editor** is used for the ETL (Extract, Transform, Load) process. This includes cleaning data (removing nulls/errors), transforming tables, and merging related data sources (like party and alliance info).
    - **Power BI Data Model** is used to create and manage relationships (e.g., one-to-many) between the core tables (Party, Candidate, State). ○ **Analysis & Calculation Engine:**
    - **DAX (Data Analysis Expressions)** are used to write formulas that create powerful, dynamic measures for analysis, such as `Total Votes`, `Vote Share %`, and `Margin` (difference between 1st and 2nd place). ○ **Visualization & Dashboard (Frontend):**
    - Built using **Power BI Visuals** to create a clean, interactive, and responsive user interface for analysis. This includes Bar charts, Pie charts, Column charts, Tables, Slicers, and Filters.

- Employs geospatial visuals like **Bing Maps** and **OpenStreetMap** to provide a clear region-wise analysis of vote share and party performance..

- **User-Facing Dashboards & Functionalities:**

Primary User (Media Professional / Journalist / Analyst):

- Access and interact with two primary dashboards for real-time analysis.
- Utilize slicers and filters to drill down into data by state, party, or constituency.
- Cross-filter visuals by clicking on chart elements (e.g., clicking a party to see its performance across all charts).

National Overview Dashboard:

- View high-level KPIs such as Total Constituencies, and seats won by major parties (BJP, INC) and alliances (NDA, I.N.D.I.A). ○ Analyze the "Alliance-wise Vote share" through pie charts.
- Identify "Top Leading Parties" and "Margin-wise Leading Candidates" at a national level.

Security & Access Control Features:

- **Platform-Managed Security:** The project relies on the inherent security and authentication protocols of the **Microsoft Power BI** ecosystem (Desktop and Pro Service).
- **Access & Sharing Control:** (As noted in the "Disadvantages" section of the document) Sharing of dashboards for internal or public use is controlled by the **Power BI Pro licensing model**, which manages who can view and interact with the reports.
- **Data-Level Security (Implicit):** Data is encapsulated within the Power BI data model and file; it is not a live database exposed to the web, which inherently limits direct access to the raw data sources.

Deployment and Hosting:

Dashboard Development & Viewing: The primary platform for development, testing, and viewing the dashboards is the **Microsoft Power BI Desktop** application.

Data Processing & Modeling: All "backend" logic, including ETL (via Power Query Editor) and analytics (via DAX measures), is handled internally within the Power BI file's data model.

Data Source: The project uses **static CSV/Excel datasets** that are imported locally. It does not connect to a live cloud-hosted database.

Cloud Deployment & Sharing: To deploy the dashboards for public or organizational access, a **Power BI Pro license** is required. This allows the `.pbix` file to be published to the **Microsoft Power BI Service** (cloud), which enables sharing and web-based viewing. ○

Component	Technology Used	Purpose & Details
Data Source	**CSV/Excel Dataset**	**Holds all raw election data (party, candidate, votes, state).**
ETL Tool	**Power Query Editor**	**Used to clean, merge, rename, and transform the raw datasets before analysis.**
Data Modeling	**Power BI Data Model**	**Defines relationships between data tables (e.g., Party, Candidate, State) using keys.**
Analytics Engine	**DAX (Data Analysis Expressions)**	**Creates dynamic, calculated measures for analysis (e.g., Total Votes, Margins, Vote %).**
Visualization / Frontend	**Power BI Visuals (Bar, Pie, Map, Slicers), Charts**	**Builds the interactive dashboards and provides geo-visual representation of data.**
Platform	**Microsoft Power BI Desktop**	**The main development, testing, and viewing platform for the dashboards.**
Hosting/Deployment	**Power BI Service (with Pro License)**	**Required to publish, share, and deploy the dashboards for web or organizational access.**

4. Timeline Overview

Week	Activities Planned	Activities Completed
Week 1	**Project Initiation & Data Sourcing:** Understood project requirements for media reporting. Sourced and collected all raw election data (CSV/Excel files for parties, candidates, votes, states, etc.).	Completed as per Planned
Week 2	**Data Processing (Module 1):** Imported all source files into Power BI. Used Power Query Editor to clean, transform, rename columns, and merge datasets (e.g., adding Party and Alliance info). Handled nulls, duplicates, and errors.	Completed as per Planned
Week 3	**Data Modeling (Module 2):** Moved from Power Query to the Data Model view. Created the schema by defining relationships between the core tables ('Party', 'Candidate', and 'State').	Completed as per Planned
Week 4	**DAX Calculation (Module 3):** Developed the analytical engine. Used DAX to write calculated measures for essential insights, such as Total Votes = SUM(...), Vote Share = DIVIDE(...), and Margin.	Completed as per Planned
Week 5	**Visualization (Module 4):** Began building the individual visual components. Created core charts (Bar chart for Top Parties, Column chart for Margins, Pie chart for Alliance Share, Table for Constituency leaders).	Completed as per Planned
Week 6	**Dashboard 1 (Module 5):** Assembled the "National Overview" dashboard. Combined KPI cards, national-level charts, and summary tables into a single, interactive report page.	Completed as per Planned
Week 7	**Dashboard 2 (Module 5):** Assembled the "State-Level Insights" dashboard. Integrated map visuals (Bing/OpenStreetMap) and added state-specific charts (NOTA analysis, Votes by State). Added all slicers and filters.	Completed as per Planned

Week 8	**Final Review & Documentation:** Conducted end-to-end testing of all dashboards, filters, and slicer interactions. Prepared final documentation and presentation materials.	Completed as per Planned

5a. Key Milestones

Milestone	Description	Date Achieved
Project Kickoff	**The project began by understanding the problem statement (raw election data for media) and defining the objective (interactive visualization). The tool was selected (Power BI), and a clear project plan and module structure were created.**	**04-Sep-2025**
Prototype/First Draft	**Designed the initial data flow (DFD). Imported, cleaned, and transformed all raw CSV/Excel data using Power Query (Module 1). Established the data model and table relationships (Module 2).**	**15-Sep-2025**
Mid-Term Review	**Presented the partially working prototype with a functional data model. Showcased the core DAX calculations (Module 3) and initial individual visuals (bar charts, pie charts) to demonstrate data flow. Received feedback on measures and visual clarity.**	**29-Sep-2025**
Final Submission	**Completed full dashboard integration. Finalized all modules by assembling all visuals into the "National Overview" and "State-Level Insights" dashboards (Module 5). Conducted end-to-end testing of all filters, slicers, and map interactions, and completed the documentation.**	**17-Oct-2025**
Presentation	**Delivered the final project demonstration, showcasing the two fully interactive dashboards. Highlighted how the solution solves the problem for media professionals, explained the architecture, and discussed key learnings and future enhancements.**	

5b. Project execution details

5b. Project execution details (Modules)

1. Data Processing Module

- **Description: This module handles the ingestion, cleaning, and transformation of all raw source data to make it usable for analysis.**
- **Implementation:**

- Imports raw election data from CSV and Excel files directly into Power BI. ○ Uses the Power Query Editor to perform ETL (Extract, Transform, Load) operations.
- Tasks include removing nulls, duplicates, and ad errors. ○ Renames columns for clarity and merges related datasets (e.g., joining Party and Alliance information).

- **Purpose: To convert multiple, raw, and potentially messy data sources into a single, clean, and structured dataset ready for analysis.**

2. Data Modeling Module

- **Description: This module establishes the logical structure of the data by defining the relationships between the different data tables.**
- **Implementation:**
  - **Uses the Power BI Data Model view. ○ Creates relationships (joins) between the core tables: 'Party', 'Candidate', and 'State'. ○ Defines the cardinality of the relationships (e.g., One 'Party' has Many 'Candidates').**
- **Purpose: To build a robust relational model that allows for complex, cross-table analysis and enables interactive filtering across the entire dashboard.**

3. DAX Calculation Module

- **Description: This module serves as the analytical engine, creating new insights and key performance indicators (KPIs) that do not exist in the raw data.**
- **Implementation:**
  - **Uses DAX (Data Analysis Expressions) to write formulas for calculated measures.**
  - **Key measures created:**
    - **Total Votes = SUM('Election'[Votes])**
    - **Vote Share = DIVIDE([Total Votes], CALCULATE(SUM('Election'[Votes]), ALL('Election')))**
    - **Margin (Difference between 1st and 2nd candidates' votes).**
- **Purpose: To generate the dynamic, analytical insights (like percentages and margins) required by media professionals for their reporting.**

4. Visualization Module

- **Description: This module focuses on creating all the individual visual elements (charts, tables, and maps) that will be used to display the data.**
- **Implementation:**
  - **Uses Power BI's native visualization pane. ○ Builds specific visuals like:**
    - **Bar chart: Top Parties by Votes**
    - **Column chart: Margins by Candidate**
    - **Pie chart: Alliance-wise Vote Share**
    - **Table: Leading Candidate and Party by Constituency**
    - **Maps (Bing/OpenStreetMap): To show state-wise performance.**

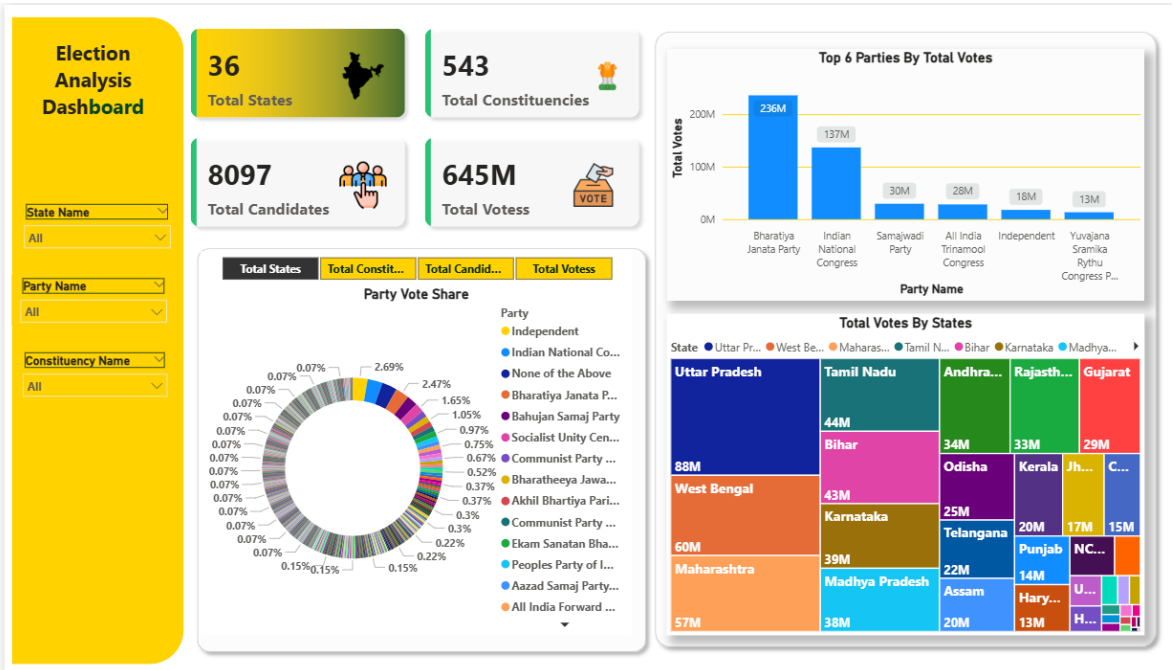
- **Purpose: To translate complex data and numbers into clear, intuitive, and easy-tounderstand visual formats for quick interpretation.**

5. Dashboard Module

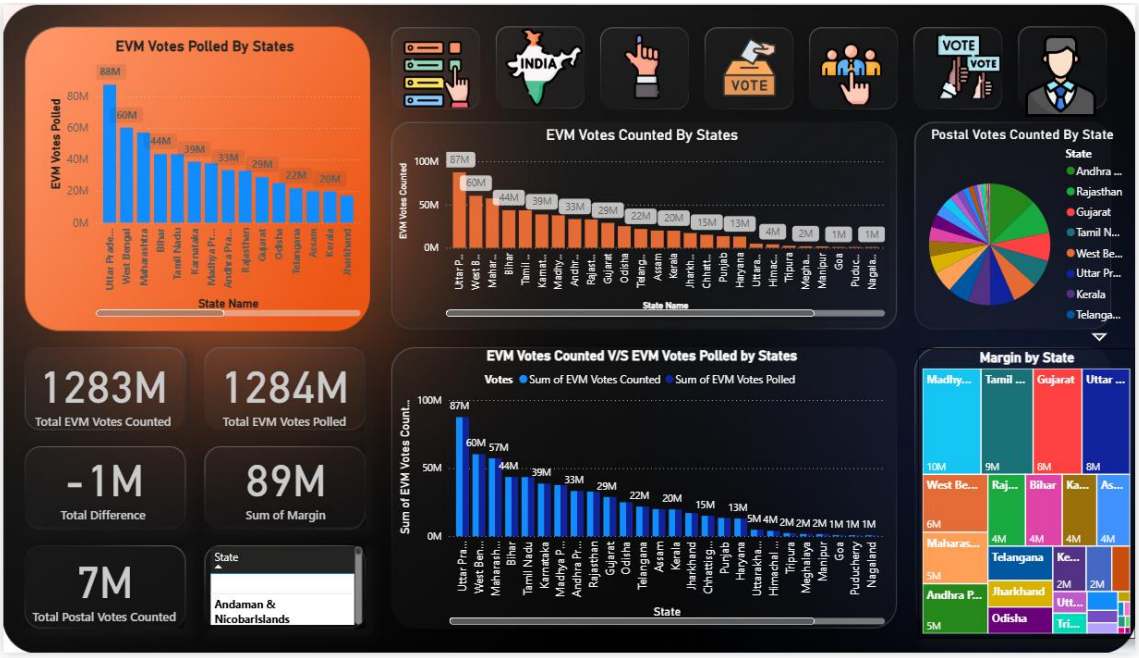
- **Description: This module is the final presentation layer, which combines all the individual visuals into two cohesive and interactive dashboards for the end-user.**
- **Implementation:**
  - **Dashboard 1: National Overview: Assembles all visuals related to the complete national results (party/alliance seats, vote shares, top candidates).**
  - **Dashboard 2: State-Level Insights: Assembles all visuals for regional analysis, including state-specific data, NOTA analysis, and the interactive map.**
  - **Adds slicers and filters to allow users to drill down by state, party, or constituency.**
- **Purpose: To deliver the final, polished product to media professionals, enabling them to explore the data, identify trends, and use the visuals for live reporting.**

6 .Snapshots / Screenshots

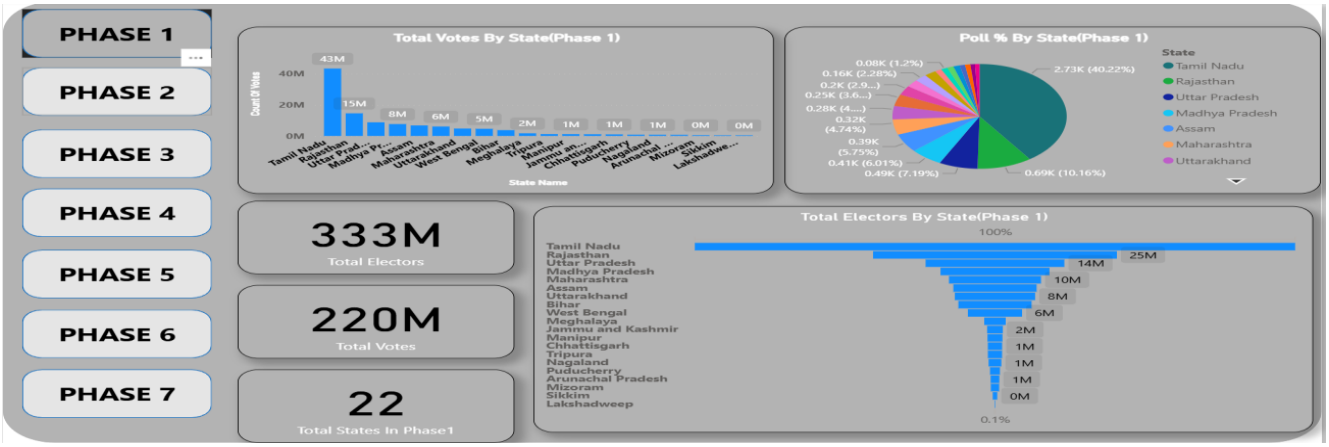
Main Dashboard :



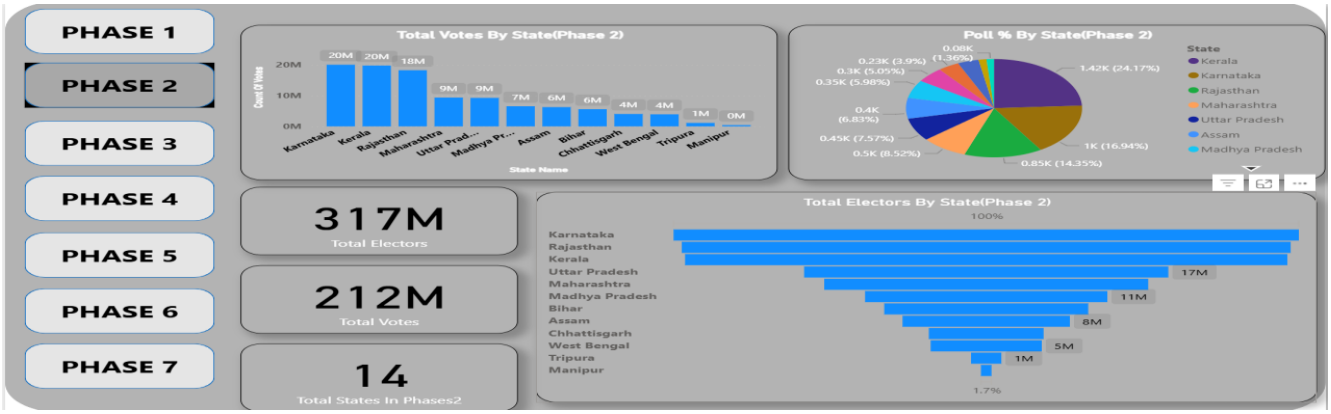
Polled V/S Counted Votes :



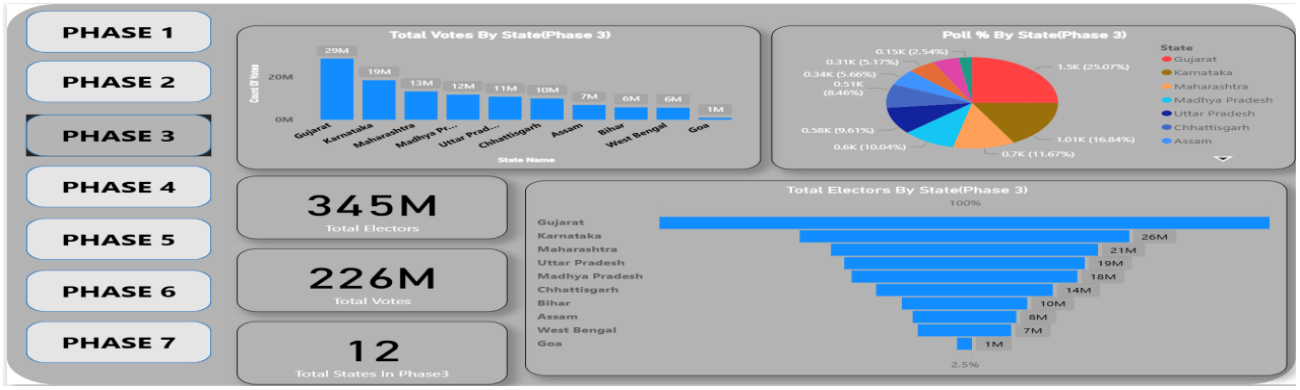
Phase 1 Dashboard :



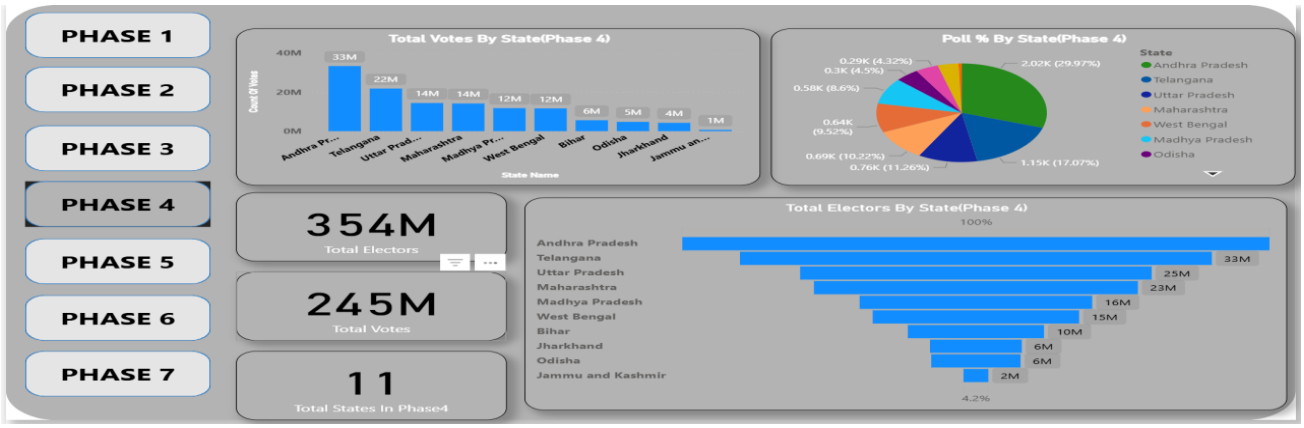
Phase 2 Dashboard :



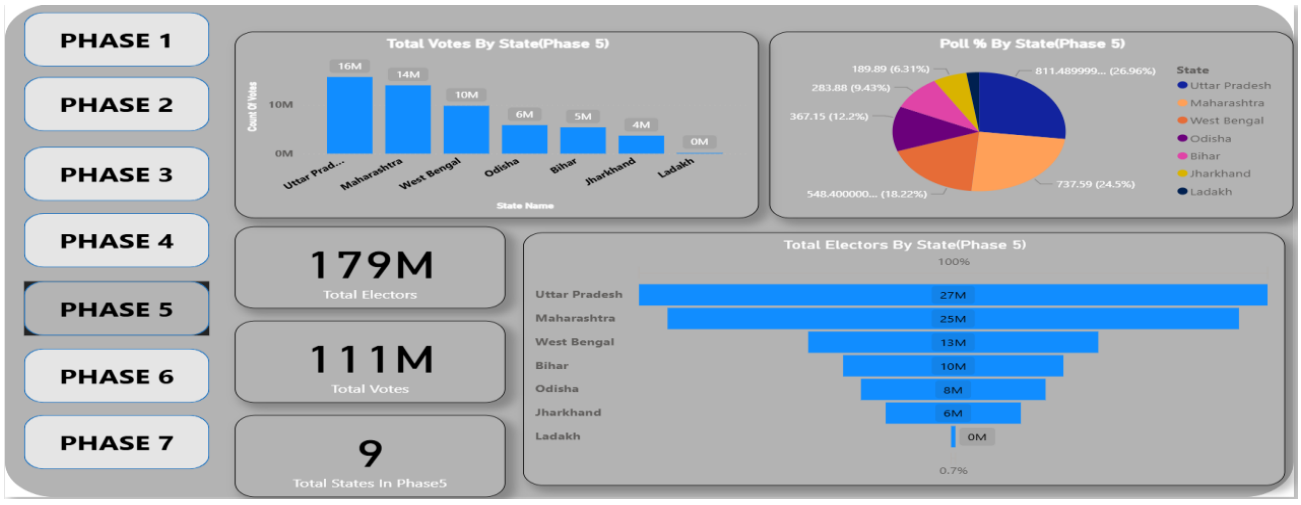
Phase 3 Dashboard :



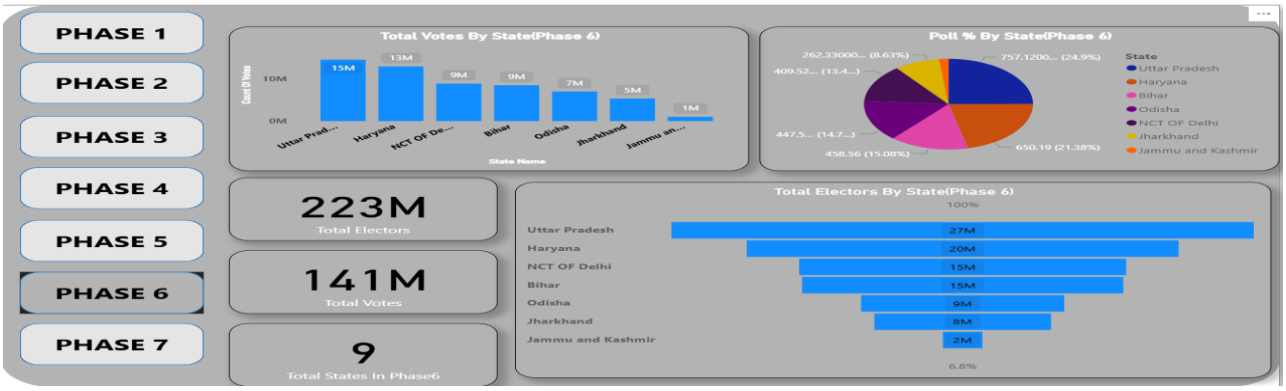
Phase 4 Dashboard



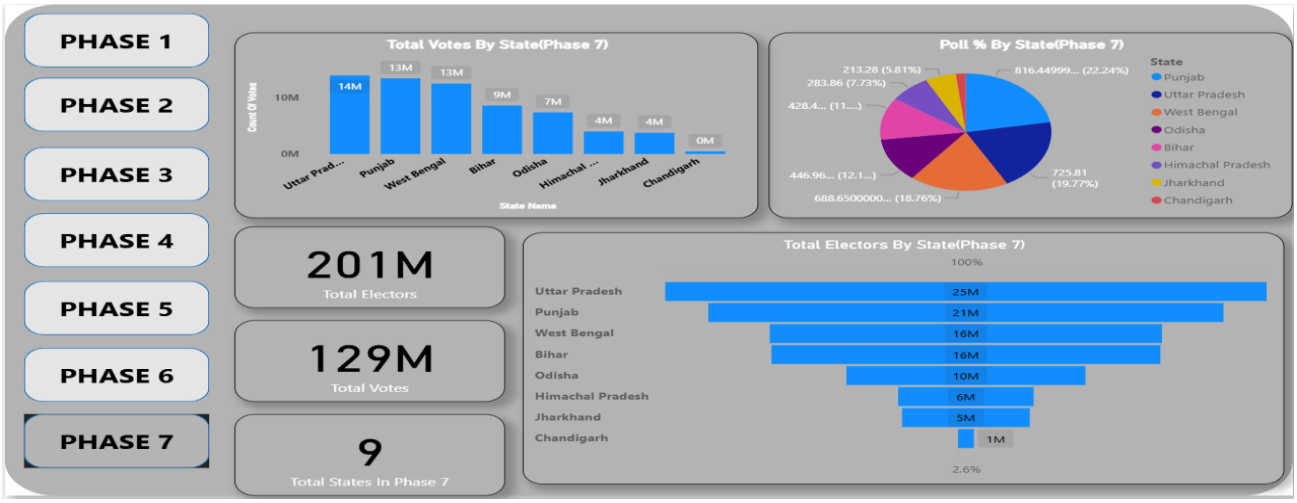
Phase 5 Dashboard :



Phase 6 Dashboard :



Phase 7 Dashboard :



7. Challenges Faced

During the development of the Electviz Election Data Visualization for Media, several challenges were encountered across technical, data, and operational domains:

a. Technical Challenges

- **DAX Calculation Complexity: Writing accurate and efficient DAX formulas for complex measures like 'Vote Share %' (which required ignoring existing filters) and 'Margins' was challenging. Resolution: Used the CALCULATE and ALL functions to correctly determine**

percentages against the grand total, not just the filtered subset. Each DAX measure was tested individually in a table before being added to a visual.

- **Data Model Relationships: Ensuring the multiple source files (Party, Candidate, State, Alliance) were correctly linked in the data model was difficult. An incorrect relationship (e.g., wrong cardinality) would break all filters. Resolution: Utilized the Power BI Data Model view to establish clear, single-direction relationships between the tables (e.g., 'Party' table filters 'Candidate' table), ensuring a robust and logical schema.**

- **Map Visualization Errors: The map visuals (Bing and OpenStreetMap) failed to render data or showed incorrect locations during initial development. Resolution: Ensured all location data (State names) was clean and unambiguous in Power Query. Confirmed an active internet connection was maintained, as these visuals rely on external services to render.**

b. Data & Performance Challenges

- **Static Data Source: The dashboard was designed using static CSV/Excel files from the 2024 election, meaning it could not update automatically during a live event. Resolution: For this project's scope, the data was documented as a static snapshot. The "Future Enhancements" section notes that this could be resolved by connecting the Power BI dashboard to a live database or API.**

- **Performance with Large Data: As the dataset grew with constituency-level detail, the dashboard visuals (especially complex DAX measures) began to load slowly. Resolution: Optimized the data model by removing unused columns in Power Query and ensuring data types were set correctly (e.g., using whole numbers instead of decimals where possible) to reduce the model's memory footprint and improve refresh speed.**

c. Deployment & Operational Challenges

- **Limited Sharing and Accessibility: Sharing the final interactive dashboard with media professionals outside the organization was not possible with the free Power BI Desktop version. Resolution: The project was fully developed and tested in Power BI Desktop. It was noted in the documentation that distribution requires a Power BI Pro license to publish the report to the Power BI Service for web and mobile viewing. 8. Learnings & Skills Acquired**

a. Technical Learnings

- **Developed an end-to-end data visualization solution using Microsoft Power BI.**

- **Mastered ETL (Extract, Transform, Load) processes using the Power Query Editor to clean, merge, and transform multiple raw data sources.**

- **Gained proficiency in Data Modeling, creating logical relationships between different data tables to enable complex filtering and analysis.**

- **Learned to write advanced DAX (Data Analysis Expressions) formulas to create custom calculated measures (Total Votes, Vote Share %, Margins) essential for deep analysis.**

- **Acquired hands-on experience in building interactive dashboards with a variety of visuals, including bar charts, pie charts, tables, and geo-spatial maps (Bing/OpenStreetMap).**

- **Understood the development lifecycle of a BI project, from data ingestion to final dashboard deployment.**

b. Soft Skills & Professional Growth

- **Improved analytical and problem-solving skills while debugging DAX formulas and data model errors.**

- **Strengthened data storytelling abilities by translating millions of raw data points into clear, concise, and compelling visual insights for a non-technical audience.**

- **Enhanced communication and documentation abilities by preparing a structured project report with clear modules, objectives, and conclusions.**

- **Learned to adapt to a data-centric project, maintaining a high standard of accuracy and data integrity, which is critical for media reporting.**

c. Overall Takeaway

This project provided real-world experience in data analytics, bridging the gap between raw, complex datasets and meaningful, interactive information. Beyond just learning a software tool, it fostered a mindset toward data-driven journalism, demonstrating how visual analytics can empower media professionals to report on election results more accurately, transparently, and compellingly.

9. Testimonials from team

Testimonial – Vishnu Vardhan Reddy

My experience developing the Electviz Election Data Visualization for Media project has been a truly enriching and insightful part of my B.Tech program. It provided a valuable opportunity to apply data science concepts to a real-world scenario. Throughout the project, I gained significant hands-on experience in the end-to-end business intelligence lifecycle, enhanced my technical and problemsolving skills, and learned to approach complex data challenges more effectively. Mastering tools like Power BI, Power Query, and DAX to transform raw data into a compelling story has been a wonderful opportunity to grow both technically and analytically.

Testimonial – Sree Raja Mounika

Working on the Electviz project was an incredible learning experience. Seeing raw, messy CSV files of election data be transformed using Power Query into a clean, interactive National Dashboard was amazing. This project truly strengthened my data modeling and visualization skills and showed me the power of BI in making data understandable for everyone.

Testimonial - Ashutosh Roy

My experience with the Electviz project was challenging but highly rewarding. Writing complex DAX measures to calculate vote margins and percentages taught me more than any textbook could. It was fascinating to build a tool with a direct application for media and journalism, turning complex numbers into a clear visual story. I'm grateful for the opportunity to have worked on a project with such practical, real-world relevance.

Testimonial – Angammal R

This project was a fantastic opportunity. Developing the Electviz Election Data Visualization for Media from start to finish gave me hands-on experience in the complete data analytics lifecycle. From the initial data cleaning to building the final State-Level Dashboard with map visuals, I enhanced my technical and analytical abilities. This project has been a great learning experience and has made me much more confident in my data science skills.

10. Conclusion

The project Electviz Election Data Visualization for Media demonstrates how data analytics and visualization can transform complex electoral data into clear, insightful, interactive information. Using Microsoft Power BI, the project successfully converts large election datasets into dashboards

that are easy to explore and understand, displaying a comprehensive view of the 2024 Indian General Election by party, alliance, candidate, and region in a visually engaging format.

This project allowed us to strengthen our technical skills in Power BI, particularly in data modeling, Power Query (ETL), and writing complex DAX measures. Working on Electviz helped us integrate data engineering principles with data storytelling, proving how modern analytics tools can make information more accessible, accurate, and transparent for media professionals and the public.

We are deeply grateful for the academic framework that allowed us to pursue this project. The challenge of translating millions of rows of data into a clean, interactive dashboard not only improved the project's quality but also significantly enhanced our problem-solving and analytical abilities.

This project experience has greatly impacted my academic and career growth, inspiring us to pursue a career that combines data analytics with journalism and business intelligence. The skills, exposure, and confidence gained through this program will undoubtedly serve as a strong foundation for our future endeavors in the field of data science.

11. Acknowledgements

We would like to express our sincere gratitude to the faculty of the B.Tech (CSE - Data Science and Analytics) program for providing me with the invaluable opportunity to develop this project. This documentation has been an enriching experience, enabling me to gain practical exposure, enhance my technical skills, and apply theoretical knowledge to real-world problem-solving.

We extend my heartfelt thanks to my project guide for their constant support, guidance, and encouragement throughout the course of this project. Their valuable insights, constructive feedback, and mentorship played a crucial role in helping me understand the concepts thoroughly and complete the project successfully.

We would also like to thank our peers and colleagues for the insightful discussions and for helping create a structured and collaborative learning environment.

Finally, We are deeply grateful to the Election Commission of India, whose publicly available datasets formed the foundation of this project. This project has been a milestone in my learning journey, and We will always carry forward the knowledge and experience gained here into our future endeavors.