

Infosys Springboard Virtual Internship 6.0

Internship Completion Report

Team Details

Team Name: Team C

Batch: 11

Group: 1

Team Members:

- Jyoti Pathak
- Jasleen Kaur
- Nikhitha
- Bhavishya Priyadarshini V

Mentor Name : Mrs. Nithyasri S J

Internship Duration : 8 Weeks (2 Months)

Year : 2026

1. Project Title

ElectViz: Election Data Visualization For Media

2. Project Objective

The primary objective of the ElectViz project is to analyze and visualize Indian National-level General Election data using Microsoft Power BI to transform complex and large-scale datasets into clear, interactive, and meaningful insights. The project aims to study voting patterns, voter turnout trends, party performance, regional variations, and gender participation to support data-driven political analysis. By converting raw election data into intuitive dashboards, ElectViz enables media organizations, researchers,

and the public to better understand electoral dynamics and promote informed decision-making through effective data visualization.

The project aims to:

- Transform large-scale raw election data into meaningful visual dashboards.
- Analyze voter turnout, party performance, state-wise and regional voting patterns.
- Study gender participation and representation in elections.
- Enable data-driven storytelling for media, researchers, and the general public.
- Support better understanding of democratic processes through visual analytics.

Overall, the project focuses on making election data **accessible, understandable, and impactful** for informed decision-making and public communication.

3. Project Description

The project can be briefly and simply described under the following subtopics:

3.1 Project Overview

ElectViz is an **interactive data visualization project** developed using Microsoft Power BI to analyze and present Indian General Election data in a clear, structured, and engaging manner. The project focuses on transforming large-scale, complex election datasets into interactive dashboards that provide meaningful insights into voting patterns, voter turnout, political party performance, and regional election dynamics.

The dataset used in this project contains **national-level election data** for the duration of 11 years (1977-2014) from Kaggle, including key attributes such as election year, state, constituency, candidates, political parties, votes polled, and registered electors. Due to the size and complexity of the data, raw information was difficult to interpret directly. ElectViz addresses this

challenge by converting raw election data into visual narratives that are easy to understand and analyze.

The project is designed to support fact-based, data-driven political analysis and is useful for media organizations, analysts, researchers, and the general public. By simplifying complex datasets into visual dashboards, ElectViz enables users to interpret election outcomes more effectively and gain a deeper understanding of democratic participation in India.

3.2 Project Approach :

The **ElectViz project** was executed using a structured, milestone-based methodology to ensure systematic progress from raw data acquisition to advanced analytics and interactive visualization. Each phase was carefully planned to transform unstructured electoral data into meaningful and actionable insights.

3.2.1 Business Understanding:

From a business and analytical perspective, the project was designed to:

- Enhance transparency in electoral data reporting: The project enhances transparency by presenting election data in an accessible, visual, and interactive format.
- Improve reporting efficiency using interactive dashboards: Journalists and analysts can use dashboards to identify voting patterns, turnout changes, and gender participation trends.
- Enable fact-based storytelling for media and research purposes: Party-wise and state-wise performance analysis enables deeper electoral strategy insights.
- Provide insights into gender participation and voter engagement: A key objective was to highlight female participation trends and calculate gender disparity in candidate representation.

The dashboard aims to serve stakeholders such as:

- Media organization
- Research institutions
- Policy analysts

- Public awareness platforms

Key business questions addressed:

- What is the percentage of female candidates contesting elections?
- What is the Gender Gap % between male and female candidates?
- Which states/regions show higher voter participation?
- Which parties perform consistently across years?
- How has voter turnout evolved over time?

By leveraging data analytics and visualization tools, the project demonstrates how electoral data can be transformed into impactful, insight-driven narratives.

3.2.1 Data Collection and Understanding:

The project began with the collection of the **Indian General Election dataset from Kaggle**, which includes national-level historical election records. The dataset comprised multiple attributes such as:

- Election Year
- State and Union Territory
- Constituency Name
- Constituency Reservation Type
- Parliamentary Constituency Number
- Candidate Name
- Candidate Gender
- Political Party Name
- Political Party Abbreviation
- Votes Polled
- Total Electors
- Voter Turnout Ratio (Custom Column)
- Region (Custom Column)

An initial data exploration phase was conducted using Power BI to:

- Understand the dataset structure and relationships
- Examine column data types and distribution
- Identify missing, null, and inconsistent values
- Evaluate the relevance of each attribute for analytical objectives

Column profiling features (valid, error, and empty value checks) were used to assess data quality. This phase ensured a clear understanding of the dataset before transformation and modeling.

3.2.3 Data Cleaning and Preprocessing:

Data cleaning and preprocessing were performed using **Power BI Power Query Editor** to ensure data consistency, accuracy, and readiness for analysis. Since real-world datasets often contain inconsistencies, missing values, and formatting errors, a structured data transformation process was essential before visualization and modeling.

During this stage, the dataset was carefully examined to identify null values, duplicate records, incorrect spellings, and inconsistent formatting across columns such as state names, political party names, and candidate details. Missing or incomplete values were handled appropriately based on relevance, and unnecessary columns were removed to streamline the dataset.

Text fields were standardized by trimming extra spaces, correcting spelling inconsistencies, and ensuring uniform naming conventions. Numerical columns, including total votes and electors, were converted into appropriate numeric data types to enable accurate aggregation and calculations. Date and year fields were reformatted to support time-based trend analysis.

Additionally, new calculated and derived columns were created to enhance analytical capabilities. These transformations included preparing fields required for voter turnout calculations, gender-based analysis, and regional categorization. The applied steps were documented within Power Query to maintain transparency and reproducibility of the data transformation process.

Overall, this preprocessing phase ensured that the dataset was clean, structured, and optimized for efficient modeling, accurate analysis, and meaningful dashboard visualization.

3.2.3.1 Column-wise Data Cleaning and Transformation:

The following modifications were performed:

★ State Column Cleaning

- Replaced inconsistent spellings and duplicate entries.
- Trimmed extra spaces.
- Standardized capitalization.

Examples:

- Pondicherry replaced to Puducherry
- Chattisgarh replaced to Chhattisgarh
- Nct Of Delhi replaced to Delhi
- National Capital Territory replaced to Delhi
- Orissa replaced to Odisha
- Uttaranchal replaced to Uttarakhand

This reduced mismatched state entries and ensured proper state-wise aggregation.

★ Party_Name Column Cleaning

Replaced mismatching or different spelling values in *party_name* column.

Originally: **45 distinct party values**
After cleaning: **36 distinct values**

Examples:

- Bharatiya Janta Party, BJP replaced to Bharatiya Janata Party
- Independents replaced to Independent
- Janata Party replaced to Janata Party
- Janta Party (JP) replaced to Janata Party
- BSP replaced to Bahujan Samaj Party
- INC replaced to Indian National Congress
- SP replaced to Samajwadi Party
- AAP replaced to Aam Aadmi Party
- AITC replaced to All India Trinamool Congress
- NCP replaced to Nationalist Congress Party

- IND replaced to Independent
- CPM replaced to Communist Party of India (Marxist)
- AIFB replaced to All India Forward Bloc
- All India Majlis-E-Ittehadul Muslimeen Replaced to All India Majlis-E-Ittehadul Muslimeen

This standardization ensured accurate party-wise comparison and aggregation.

★ Gender Column Standardization

Cleaned inconsistent entries in *gender* column.

- M replaced to Male
- F replaced to Female
- Removed blank/null values
- Trimmed spaces

This enabled correct gender-based calculations.

★ Total Votes Column

- Removed commas from numeric values
- Converted data type from Text to Whole Number
- Removed null or invalid entries

This allowed proper aggregation using SUM function.

5. Total Electors Column

- Converted data type from Text to Whole Number
- Checked for blank values
- Ensured no zero or invalid entries

Used for turnout ratio calculation.

6. Year Column

- Converted from Text to Whole Number
- Standardized year format (YYYY)
- Enabled chronological sorting

Used for year-wise trend analysis.

7. Constituency Column

- Trimmed leading and trailing spaces
- Standardized text formatting
- Removed duplicate entries

Ensured clean constituency-level reporting.

8. Created Custom Columns

Added the following custom columns:

- `voters_turnout_%` → Calculated percentage of votes received over total electors
- `Region` → Classified states into:
 - North
 - South
 - East
 - West

3.2.3.2 Key Data Enhancements and Calculations

- **Gender Gap %** – Created as a DAX measure to calculate the percentage share of female candidates relative to total candidates.
- **Female Candidate** – Calculated using DAX to determine the total number of female candidates contesting elections.
- **Total Votes Received** – Cleaned and converted into numeric format to enable proper aggregation and calculations.
- **Voter Turnout Ratio** – Calculated using Total Votes and Total Electors to measure voter participation.
- **Year** – Standardized into proper date/year format to support time-based trend analysis.
- **State / Region** – Cleaned and standardized by removing duplicates and correcting spelling inconsistencies.

3.2.3.3 Data Standardization:

These preprocessing steps ensured that the dataset was structured, reliable, and ready for advanced analysis.

- **Year** column was standardized into proper date/year format for time-based analysis.
- **Total Votes Received** was cleaned and converted into numeric format to enable aggregation and calculations.
- **State / Region** column was cleaned by:
 - Removing duplicate entries
 - Correcting inconsistent spellings
 - Standardizing naming conventions

3.2.3.4 Handling Missing And Null Values:

- NULL and blank values were replaced appropriately (e.g., “Unknown” where necessary).
- In numerical columns, null values were handled carefully to avoid aggregation errors.

3.2.3.5 Derived Columns & DAX Measures:

Several calculated fields were created to enhance analytical depth:

- **Female Candidate** – Created using DAX query to count female candidates.

```
Female Candidate = CALCULATE(  
    COUNT('ElectViz(INFOSYS)'[Candidate Name])  
    , 'ElectViz(INFOSYS)'[Candidate Gender] = "F")
```

- **Gender Gap %** – Created using DAX to calculate the percentage difference between male and female candidates:

```
Gender Gap % = Divide([Female Candidate], [Total  
Candidates]) * 100
```

- **Total Candidates** – Created using DAX to calculate the total number of candidates fighting elections:

Total Candidates = Count('ElectViz(INFOSYS)')[Candidate Name])

- **Total Votes** - Created using DAX to calculate the total votes received:

Total Votes = SUM('ElectViz(INFOSYS)')[Total Votes Received])

3.2.4 Dashboard Design and Visualization:

After preprocessing, a robust data model was created in Power BI by:

- Establishing relationships between relevant tables
- Ensuring proper cardinality and filtering direction
- Validating cross-filter functionality for slicers

3.2.4.1 Vizualization Strategy:

Interactive dashboards were developed using a combination of:

- KPI Cards (for summary metrics)
- Stacked Bar Charts (for comparisons)
- Line Charts (for trend analysis)
- Donut and Pie Charts (for distribution analysis)
- Maps (for geographic insights)
- Tables and Gauges (for detailed metrics)

3.2.4.2 Analytical Coverage:

The dashboards were designed to provide multiple analytical perspectives, including:

- Election overview and summary statistics
- Gender-based candidate participation
- Regional analysis (North, South, East, West)

- State-wise voting trends
- Candidate and party participation
- Voter turnout analysis
- Top party performance across states and years

Interactive slicers for **Year, State, and Region** were incorporated to allow dynamic filtering and deeper drill-down analysis.

The final dashboard delivers a user-friendly, interactive, and insight-rich visualization platform that transforms raw election data into meaningful analytical outputs.

3.3 Technology Used

- **Microsoft Power BI** – For building interactive dashboards and visualizing election data.
- **Power Query Editor** – For data cleaning, preprocessing, and transformation.
- **DAX (Data Analysis Expressions)** – For creating calculated measures and analytical insights.
- **Microsoft Excel / CSV Files** – For importing and handling structured datasets.
- **Kaggle** – Source of Indian National-level General Election dataset.
- **GitHub** – For version control, project documentation, and collaboration.
- **GitLab** – For repository management and maintaining project versions.

3.4 Insights from the Dashboard

- Revealed clear state-wise and region-wise voting patterns across election years.
- Identified significant variations in voter turnout trends and regional participation levels.
- Highlighted party dominance and vote share distribution across states and regions.
- Analyzed candidate participation and gender representation trends over time.

- Enabled comparative insights into year-wise electoral growth and performance shifts.

3.5 Real-World Impact for Media and Public Communication

ElectViz improves the quality of election reporting by enabling data-driven journalism. Instead of presenting raw numbers or isolated statistics, media professionals can use interactive dashboards to explain trends, compare regions, and highlight key electoral shifts in a visually engaging way. This enhances clarity, reduces misinterpretation, and strengthens the credibility of political coverage.

The project also supports faster decision-making during election periods. Analysts and reporters can quickly filter data by state, year, or political party to generate insights in real time. This makes it easier to identify emerging patterns such as changing voter turnout, regional party dominance, or shifts in candidate participation.

For public communication, ElectViz promotes transparency and accessibility. Citizens can better understand how elections evolve over time, how their state compares with others, and how political participation varies across regions and genders. By simplifying complex datasets into intuitive visuals, the project encourages informed civic engagement and strengthens democratic awareness.

4.0 Time Line Overview

Time Line	Activities Planned	Activities Completed
WEEK-1	Data Collection & Preparation	Collected Indian National-level Election dataset from Kaggle . Performed data cleaning

WEEK-2	Core Dashboard Development	<p>Designed and developed foundational dashboards in Power BI:</p> <ul style="list-style-type: none"> ● Election Overview Dashboard ● Gender Analysis Dashboard ● Regional Analysis Dashboard
WEEK-3	Milestone Documentation(Phase 1)	Prepared documentation and PowerPoint presentation for Milestones 1 & 2. Structured project progress for review
WEEK-4	Analysis Visualization of the dashboards	<p>Developed additional analytical dashboards:</p> <ul style="list-style-type: none"> ● State-Wise Voting Trend Analysis Dashboard ● Candidate Participation Analysis
WEEK-5	Milestone Documentation(Phase 2)	Documented Milestone 3 dashboards and prepared presentation slides for evaluation.
WEEK-6	Data Processing and Dashboard	Created advanced dashboards for deeper

	Development Stages	<p>insights:</p> <ul style="list-style-type: none"> • Voter Turnout Analysis • Top Party Performance Dashboard <p>Finalized dashboards with filters, slicers, and consistent storytelling flow.</p>
WEEK-7	Milestone Documentation(Phase 3)	Completed documentation and presentation preparation for Milestone 4. Summarized advanced analytical insights.
WEEK-8	Final Submission and presentation	Compiled complete project documentation and delivered final presentation covering objectives, implementation, insights, and impact.

5.0 Key Milestones

Milestone	Title	Activities Performed	Activities Completed
Milestone 01	Data Collection & Preprocessing	<ul style="list-style-type: none"> • Collected Indian General Election dataset from Kaggle covering national-level elections of the last 10 	Clean, consistent, and well-structured dataset prepared for accurate analysis and visualization

		<p>years</p> <ul style="list-style-type: none"> Performed data cleaning and preprocessing using Power BI Power Query Editor Handled missing and inconsistent values Standardized state and political party names Converted data types and removed duplicate records Created required calculated fields including Region and Voter Turnout Percentage 	
Milestone 02	Initial Dashboard Development	<ul style="list-style-type: none"> Designed foundational dashboard layout in Power BI Created Election Overview page summarizing votes, candidates, parties, and states Developed Gender Analysis page to analyze male vs female candidate participation Built Regional Analysis page to study election behavior across North, South, East, and West regions 	High-level dashboards providing overview of participation patterns and regional distribution
Milestone 03	Core Visual & Trend Analysis	<ul style="list-style-type: none"> Developed State-wise Voting Trend dashboard to analyze voting patterns across states and election years 	Interactive dashboards highlighting voting trends, party participation, and candidate

		<ul style="list-style-type: none"> Created Candidate Participation Analysis focusing on party-wise and candidate-level insights Implemented interactive visuals, slicers, and filters for comparative analysis 	performance
Milestone 04	Advanced Insights & Performance Analysis	<ul style="list-style-type: none"> Designed Voter Turnout Analysis dashboard to compare turnout across years, states, and regions Built Top Party Performance dashboard to identify dominant political parties based on vote share Enabled year-wise and region-wise performance comparison 	Insight-driven dashboards providing deep analysis of voter engagement and party dominance

5.1 Project Execution Details

The execution of our project was done in a structured and sequential manner in the way as given below:

5.1.1 Data Collection:

- Collected Indian General Election dataset from Kaggle.
- Dataset includes national-level election data covering the last 10 years.
- Data contains information on states, constituencies, candidates, parties, votes, and electors.

5.1.2 Data Cleaning & Preprocessing:

Data cleaning and preprocessing were performed using **Power BI Power Query Editor** to ensure data consistency, accuracy, and readiness for analysis. Since real-world datasets often contain inconsistencies, missing values, and formatting errors, a structured data transformation process was essential before visualization and modeling.

During this stage, the dataset was carefully examined to identify null values, duplicate records, incorrect spellings, and inconsistent formatting across columns such as state names, political party names, and candidate details. Missing or incomplete values were handled appropriately based on relevance, and unnecessary columns were removed to streamline the dataset.

Text fields were standardized by trimming extra spaces, correcting spelling inconsistencies, and ensuring uniform naming conventions. Numerical columns, including total votes and electors, were converted into appropriate numeric data types to enable accurate aggregation and calculations. Date and year fields were reformatted to support time-based trend analysis.

Additionally, new calculated and derived columns were created to enhance analytical capabilities. These transformations included preparing fields required for voter turnout calculations, gender-based analysis, and regional categorization. The applied steps were documented within Power Query to maintain transparency and reproducibility of the data transformation process.

Overall, this preprocessing phase ensured that the dataset was clean, structured, and optimized for efficient modeling, accurate analysis, and meaningful dashboard visualization.

5.1.2.1 Column-wise Data Cleaning and Transformation:

The following modifications were performed:

★ State Column Cleaning

- Replaced inconsistent spellings and duplicate entries.

- Trimmed extra spaces.
- Standardized capitalization.

Examples:

- Pondicherry replaced to Puducherry
- Chattisgarh replaced to Chhattisgarh
- Nct Of Delhi replaced to Delhi
- National Capital Territory replaced to Delhi
- Orissa replaced to Odisha
- Uttaranchal replaced to Uttarakhand

This reduced mismatched state entries and ensured proper state-wise aggregation.

★ Party_Name Column Cleaning

Replaced mismatching or different spelling values in *party_name* column.

Originally: **45 distinct party values**
 After cleaning: **36 distinct values**

Examples:

- Bharatiya Janta Party, BJP replaced to Bharatiya Janata Party
- Independents replaced to Independent
- Janata Party replaced to Janata Party
- Janta Party (JP) replaced to Janata Party
- BSP replaced to Bahujan Samaj Party
- INC replaced to Indian National Congress
- SP replaced to Samajwadi Party
- AAP replaced to Aam Aadmi Party
- AITC replaced to All India Trinamool Congress
- NCP replaced to Nationalist Congress Party
- IND replaced to Independent
- CPM replaced to Communist Party of India (Marxist)
- AIFB replaced to All India Forward Bloc
- All India Majlis-E-Ittehadul Muslimeen Replaced to All India Majlis-E-Ittehadul Muslimeen

This standardization ensured accurate party-wise comparison and aggregation.

★ Gender Column Standardization

Cleaned inconsistent entries in *gender* column.

- M replaced to Male
- F replaced to Female
- Removed blank/null values
- Trimmed spaces

This enabled correct gender-based calculations.

★ Total Votes Column

- Removed commas from numeric values
- Converted data type from Text to Whole Number
- Removed null or invalid entries

This allowed proper aggregation using SUM function.

5. Total Electors Column

- Converted data type from Text to Whole Number
- Checked for blank values
- Ensured no zero or invalid entries

Used for turnout ratio calculation.

6. Year Column

- Converted from Text to Whole Number
- Standardized year format (YYYY)
- Enabled chronological sorting

Used for year-wise trend analysis.

7. Constituency Column

- Trimmed leading and trailing spaces
- Standardized text formatting
- Removed duplicate entries

Ensured clean constituency-level reporting.

8. Created Custom Columns

Added the following custom columns:

- `voters_turnout_%` → Calculated percentage of votes received over total electors
- Region → Classified states into:
 - North
 - South
 - East
 - West

5.1.2.2 Key Data Enhancements and Calculations

- **Gender Gap %** – Created as a DAX measure to calculate the percentage share of female candidates relative to total candidates.
- **Female Candidate** – Calculated using DAX to determine the total number of female candidates contesting elections.
- **Total Votes Received** – Cleaned and converted into numeric format to enable proper aggregation and calculations.
- **Voter Turnout Ratio** – Calculated using Total Votes and Total Electors to measure voter participation.
- **Year** – Standardized into proper date/year format to support time-based trend analysis.
- **State / Region** – Cleaned and standardized by removing duplicates and correcting spelling inconsistencies.

5.1.2.3 Data Standardization:

These preprocessing steps ensured that the dataset was structured, reliable, and ready for advanced analysis.

- **Year** column was standardized into proper date/year format for time-based analysis.

- **Total Votes Received** was cleaned and converted into numeric format to enable aggregation and calculations.
- **State / Region** column was cleaned by:
 - Removing duplicate entries
 - Correcting inconsistent spellings
 - Standardizing naming conventions

5.1.2.4 Handling Missing And Null Values:

- NULL and blank values were replaced appropriately (e.g., “Unknown” where necessary).
- In numerical columns, null values were handled carefully to avoid aggregation errors.

5.1.2.5 Derived Columns & DAX Measures:

Several calculated fields were created to enhance analytical depth:

- **Female Candidate** – Created using DAX query to count female candidates.

```
Female Candidate = CALCULATE(
  COUNT('ElectViz(INFOSYS)'[Candidate Name]),
  'ElectViz(INFOSYS)'[Candidate Gender] = "F")
```

- **Gender Gap %** – Created using DAX to calculate the percentage difference between male and female candidates:

```
Gender Gap % = Divide([Female Candidate], [Total Candidates]) * 100
```

- **Total Candidates** – Created using DAX to calculate the total number of candidates fighting electionss:

```
Total Candidates = Count('ElectViz(INFOSYS)'[Candidate Name])
```

- **Total Votes** - Created using DAX to calculate the total votes received:

Total Votes = SUM('ElectViz(INFOSYS)'[Total Votes Received])

5.1.3 Feature Engineering & Calculations:

- Created a **Region** column to classify states into North, South, East, and West.
- Calculated **Voter Turnout Percentage** using total votes polled and registered electors.
- Created additional calculated measures to support aggregation and trend analysis.
- Developed year-wise and state-wise calculated metrics to enable comparative trend analysis across multiple election cycles.
- Created dynamic DAX measures for party vote share and candidate participation to support interactive filtering and detailed insights.

5.1.4 Data Modeling:

- Established relationships between key fields such as state, year, party, and candidate.
- Optimized data model to enable cross-filtering and interactive visuals.
- Ensured accurate summarization across multiple dashboard pages.
- Designed a structured data schema to minimize redundancy and improve query performance within Power BI.
- Created calculated columns and measures using DAX to support dynamic aggregations and comparative analysis.
- Validated relationships and filter directions to ensure consistent and accurate data representation across all dashboard visuals.

5.1.5 Dashboard Development:

- Designed interactive dashboards using **Power BI**.

- Used KPIs, charts, tables, maps, and slicers for effective visualization.
- Developed multiple analytical pages including:
 1. Election Overview
 2. Gender Analysis
 3. Regional Analysis
 4. State-wise Voting Trends
 5. Voter Turnout Analysis
 6. Party Performance Analysis

5.1.6 Validation & Refinement:

- Verified accuracy of calculations and visual outputs.
- Improved layout, color consistency, and readability.
- Ensured dashboards aligned with project objectives and user requirements.
- Collected feedback from team members and mentors to refine visual elements and enhance overall user experience and dashboard performance.

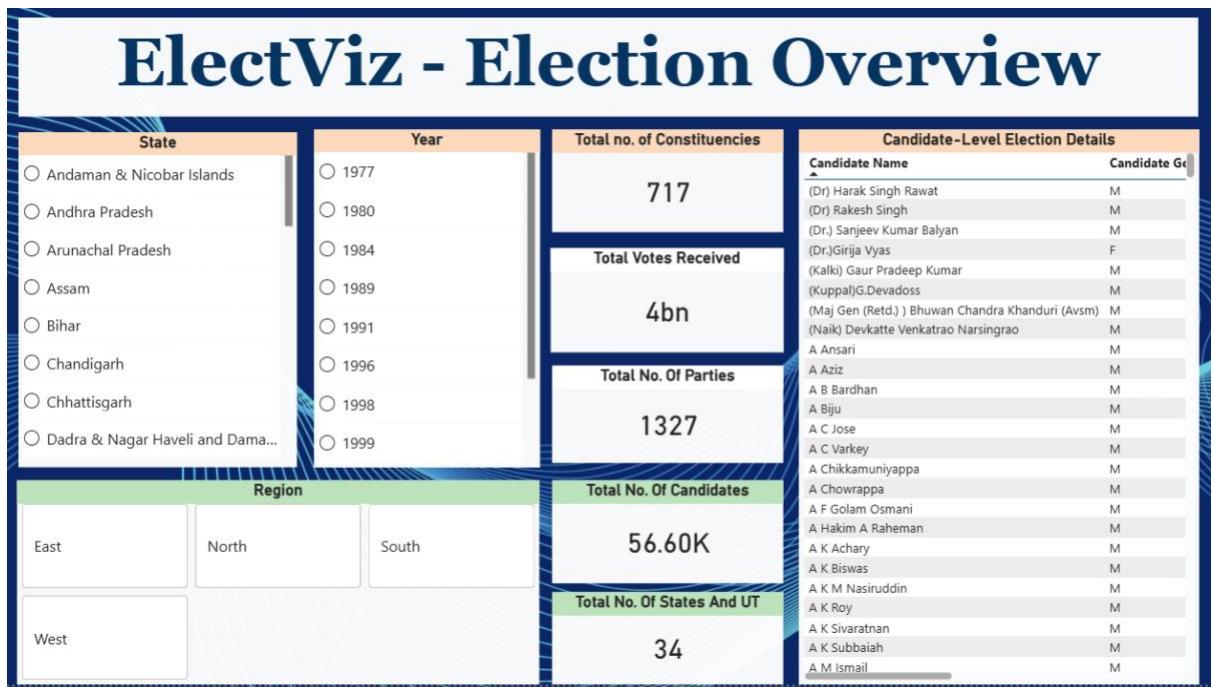
5.1.7 Final Outcome:

- Delivered a clean, interactive, and insight-driven election analytics dashboard.
- Enabled clear understanding of voting patterns, participation trends, and party dominance.
- Demonstrated practical application of data analytics and visualization skills.
- Provided a structured and scalable framework that can be extended for real-time election tracking and advanced predictive analysis.
- Enhanced transparency and supported data-driven storytelling for media professionals, researchers, and public communication.

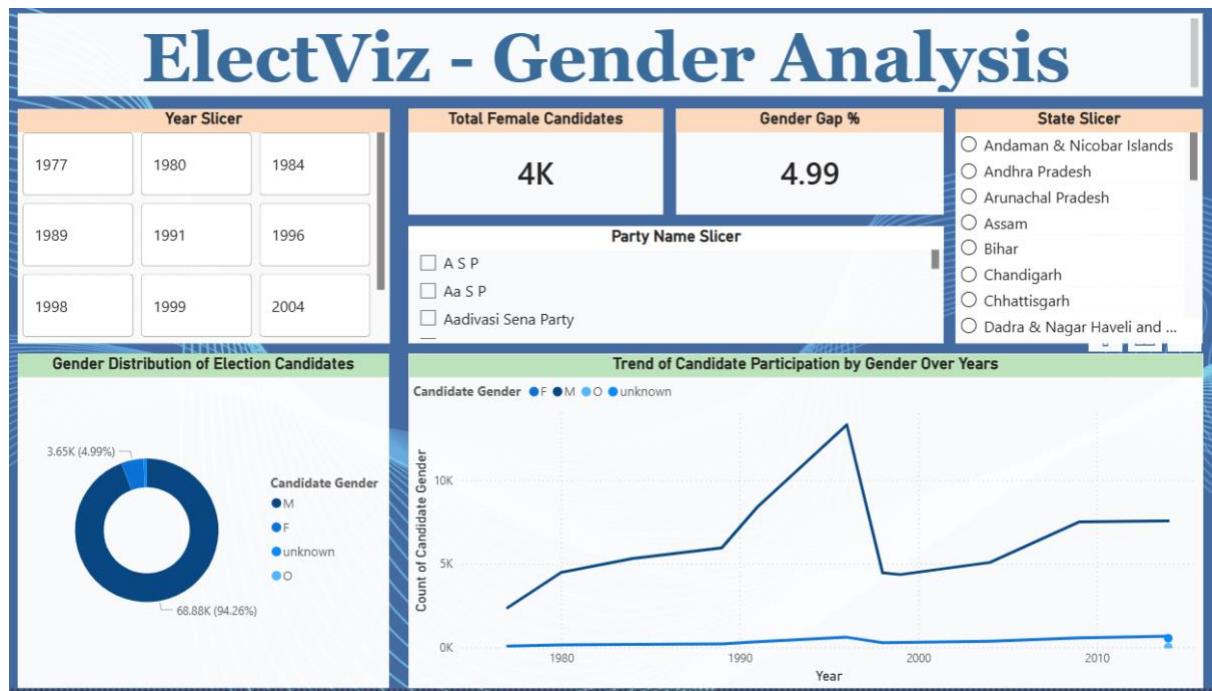
6.0 Snapshots Of The Dashboard

The Dashboard consists of 7 interactive pages as given below:

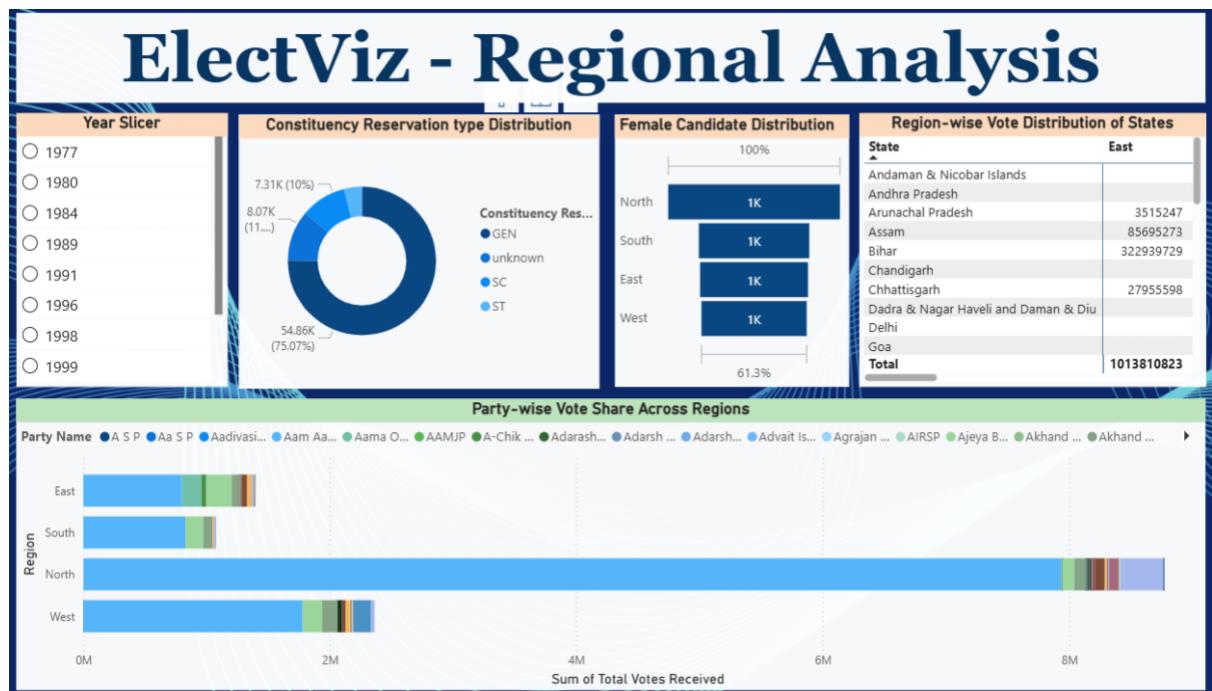
6.1 Election Overview:



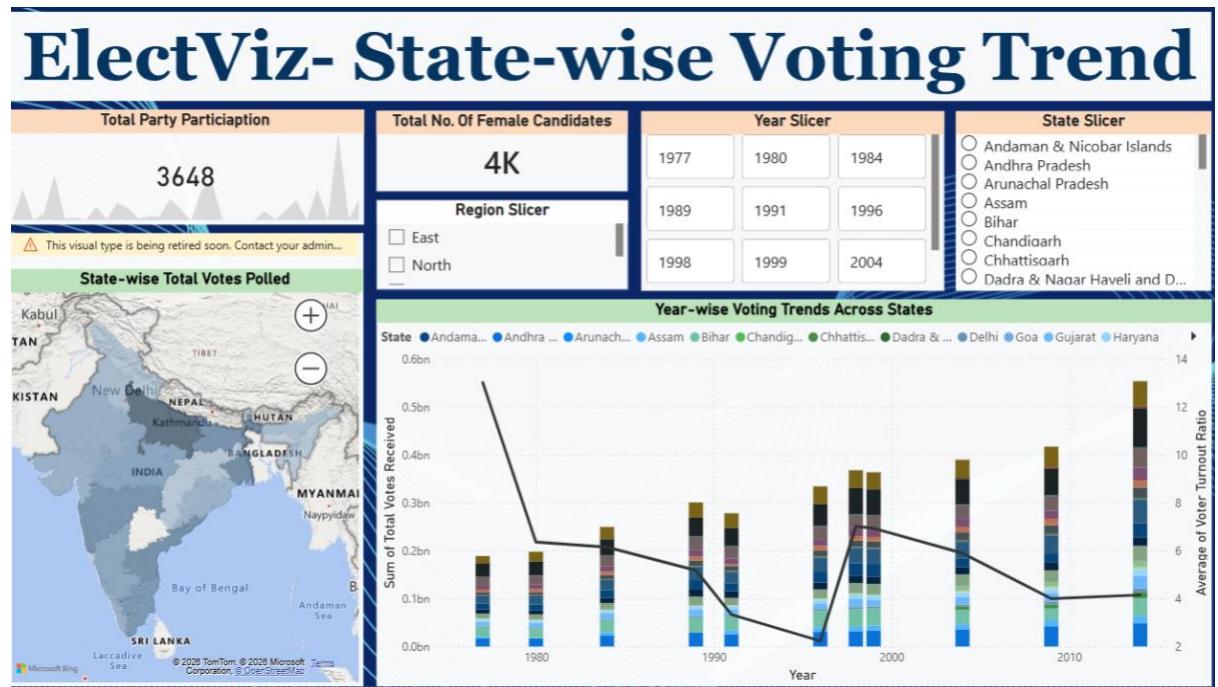
6.2 Gender Analysis:



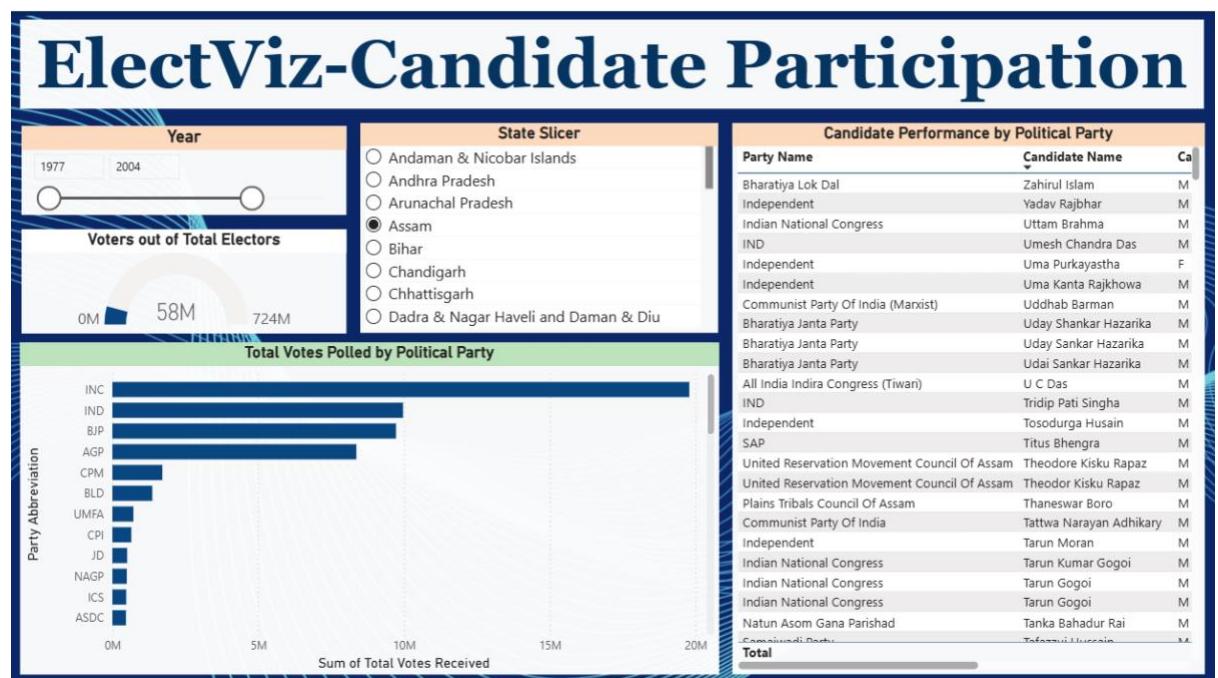
6.3 Regional Analysis:



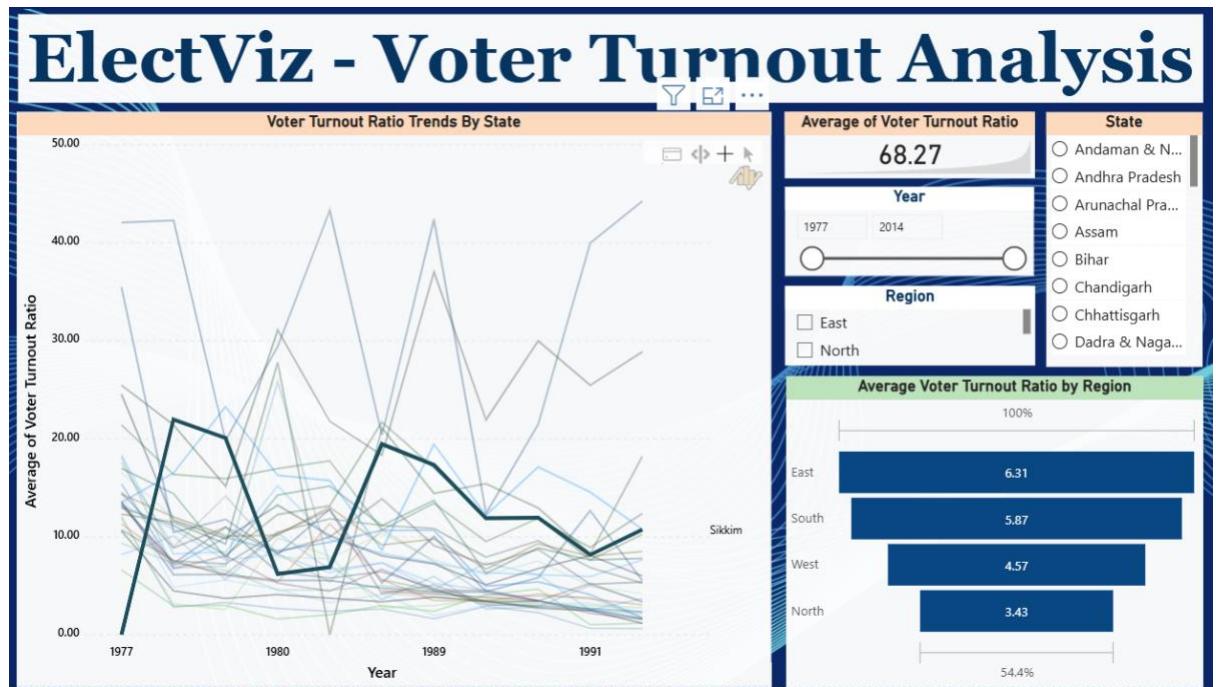
6.4 State-Wise Voting Trend Analysis:



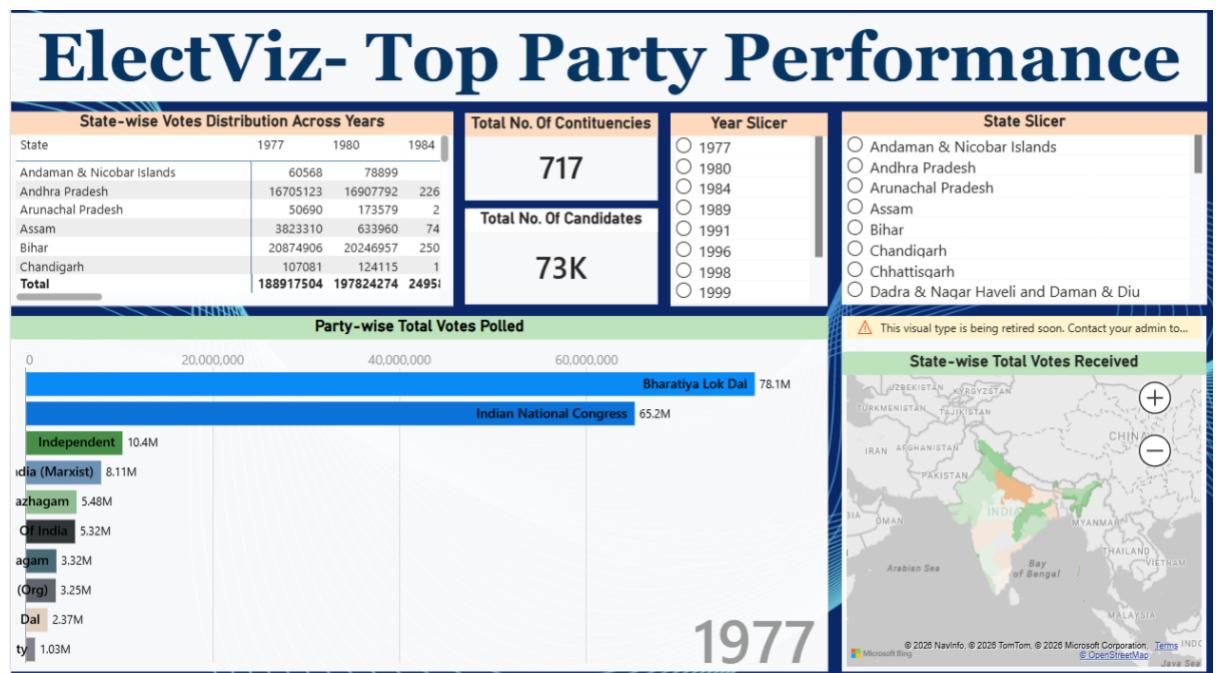
6.5 Candidate Participation Analysis:



6.6 Voter Turnout Analysis:



6.7 Top Party Performance:



7.0 Challenges Faced

During the development of the ElectViz project, the team encountered several challenges related to data handling, technical implementation, collaboration, and visualization design. These challenges were addressed systematically to ensure successful project completion.

7.1 Data Integration and Cleaning

One of the major challenges was working with a large and complex election dataset collected from an open public source. The dataset contained missing values, inconsistent state and party names, duplicate records, and incorrect data formats, which made direct analysis difficult. Integrating different election attributes such as votes, candidates, and elector details required careful validation to avoid inaccurate insights. Significant effort was required to clean, standardize, and structure the data before it could be used for visualization.

7.2 Technical Alignment in Power BI

Ensuring proper data modeling and technical alignment in Power BI was another challenge. Establishing correct relationships between tables, maintaining filter flow across visuals, and ensuring slicers worked consistently across all dashboard pages required multiple iterations. Managing calculated fields, measures, and aggregations while avoiding errors and incorrect results was also a critical technical challenge during dashboard development.

7.3 Communication and Time Management

As this was a group project, effective communication and time management were essential but challenging. Coordinating tasks among team members, aligning individual contributions, and meeting milestone deadlines required continuous discussion and planning. Differences in understanding, availability, and workload distribution occasionally caused delays, which were resolved through regular communication and collaborative problem-solving. To overcome these challenges, the team established clear task allocation and set weekly progress checkpoints to monitor deliverables. Regular meetings were conducted to review work status, clarify doubts, and ensure alignment with project objectives. By prioritizing responsibilities and

maintaining open communication channels, the team improved coordination, minimized misunderstandings, and successfully met project deadlines in a structured and organized manner.

7.4 Dashboard Optimization and Performance Issues

Due to the large volume of data and multiple visuals, the Power BI dashboards initially faced performance issues, including slow loading times and delayed visual responses. Optimizing dashboard performance required reducing unnecessary visuals, simplifying calculations, optimizing data models, and ensuring efficient use of filters and slicers. Achieving a balance between detailed analysis and smooth performance was a key challenge.

7.5 Design Consistency and Storytelling Alignment

Maintaining visual consistency and a clear storytelling flow across multiple dashboard pages was another important challenge. Ensuring consistent color schemes, chart types, labels, and layouts while aligning visuals with meaningful insights required careful design planning. The team focused on structuring dashboards in a logical sequence so that users could easily understand the story behind the data without confusion. Continuous refinement and feedback-based improvements were implemented to ensure clarity and uniformity. This approach helped create a seamless user experience where insights flowed naturally from one dashboard to the next, enhancing overall readability and impact.

8.0 Learnings & Skills Acquired

During the period of internship we acquired many important skills which can be summarized as below:

8.1 Technical Skills

- Gained hands-on experience in **Power BI** for building interactive and dynamic dashboards.
- Learned **data cleaning and preprocessing techniques**, including handling missing values, data normalization, and standardization.

- Worked with **large-scale election datasets**, improving efficiency in managing and transforming structured data.
- Implemented advanced **data visualizations** such as maps, stacked charts, line charts, ribbon charts, and KPI cards.
- Used **DAX calculations** to derive insights such as voter turnout ratios, total votes, averages, and party-wise aggregations.
- Applied consistent **dashboard design principles** including layout, color themes, and user-friendly slicers.

8.2 Analytical and Problem-Solving Skills

- Developed the ability to analyze complex election data and extract meaningful patterns and trends.
- Interpreted state-wise, region-wise, and year-wise voting behaviors to support data-driven conclusions.
- Solved visualization and performance challenges by optimizing data models and selecting appropriate chart types.
- Translated raw numerical data into clear analytical insights that support comparative and trend analysis.
- Improved logical thinking by validating insights against multiple filters and scenarios.

8.3 Soft Skills and Team Collaboration

- Enhanced **communication skills** by explaining analytical findings through visual dashboards.
- Gained experience in **team collaboration**, coordinating tasks and integrating individual contributions into a unified project.
- Improved **time management** by completing weekly milestones within project timelines.
- Developed **adaptability** by incorporating feedback and making iterative improvements to dashboards.

8.4 Domain Knowledge and Application

- Learned key election metrics such as **voter turnout ratio, party performance, and candidate participation**.
- Applied domain knowledge to design dashboards that reflect real-world electoral scenarios.

- Understood how data analytics and visualization can support **political analysis, media reporting, and public awareness.**
- Connected theoretical concepts of data analytics with practical, real-world election data applications.
- Gained deeper understanding of the structure of the Indian electoral system, including constituencies, regional classifications, and multi-party competition dynamics.
- Interpreted how socio-regional variations influence voting behavior and party dominance across different states.

9. Testimonials from Team

Jyoti Pathak -“Working on the ElectViz project provided us with a practical understanding of how raw, real-world datasets can be systematically transformed into meaningful and actionable insights using Power BI. Throughout the project, we learned how to approach data analytically, starting from data exploration and cleaning to modeling and visualization. It strengthened our technical proficiency in DAX, Power Query, and dashboard design, while also enhancing our analytical thinking. Additionally, collaborating closely as a team helped us improve coordination, idea-sharing, and collective problem-solving skills.”

Jasleen Kaur -“This internship offered valuable hands-on experience in managing and analyzing large-scale datasets. We gained practical exposure to data cleaning, preprocessing, creating calculated columns and measures, and designing interactive dashboards in Power BI. The experience helped us understand the importance of data accuracy and structured modeling in delivering reliable insights. Working on real election data significantly boosted our confidence in using Power BI as a professional data visualization and business intelligence tool.”

Nikhitha -“Collaborating as a team throughout the ElectViz project enhanced our communication, coordination, and task management skills. We learned how to divide responsibilities effectively, track

milestones, and meet deadlines efficiently. Beyond technical learning, the project deepened our understanding of Indian election data, voter behavior, and gender participation trends. It allowed us to connect theoretical knowledge of data analytics with practical, real-world applications, making the learning experience both insightful and impactful.”

Bhavishya Priyadarshini V -“The ElectViz project provided valuable exposure to structured project planning, milestone-based execution, and collaborative development. From data collection and preprocessing to dashboard deployment, we followed a systematic workflow that improved our organizational and analytical capabilities. Working together on complex data analysis tasks enhanced our ability to think critically, adapt to challenges, and resolve issues effectively. Moreover, designing interactive dashboards helped us develop the skill of presenting complex datasets in a clear, concise, and visually compelling manner, ensuring that insights are easily understandable for diverse audiences.”

10. Conclusion

The **Infosys Springboard Virtual Internship 6.0** project successfully demonstrated the practical application of data analytics and visualization techniques to analyze Indian General Election data in a structured and meaningful way. By leveraging **Microsoft Power BI**, the team transformed raw and complex electoral datasets into interactive and visually engaging dashboards. These dashboards effectively highlighted key aspects such as voter turnout trends over the years, comparative party performance, state-wise and regional voting behavior, and gender participation patterns.

Through systematic data cleaning, preprocessing, and transformation using Power Query, the dataset was refined to ensure accuracy and consistency. Calculated fields and measures were created to derive meaningful insights, enabling deeper analysis beyond basic statistics. The use of slicers, filters, and drill-down features allowed users to interact with the data dynamically, making the dashboards more insightful and user-friendly.

The project provided valuable insights into electoral patterns, including shifts in voter turnout across different regions, dominance of political parties in specific states, and variations in male and female participation. These insights highlighted the power of **data-driven storytelling**, especially in media and public communication, where complex political data needs to be presented in a clear, concise, and impactful manner.

In addition to technical learning, the project significantly strengthened analytical thinking, problem-solving abilities, and collaborative skills among team members. Working as a team encouraged effective communication, task coordination, and shared decision-making. The experience also enhanced proficiency in Power BI, data interpretation, and dashboard design principles.

Overall, this project stands as a strong example of how data analytics and visualization tools can simplify large and complex datasets, transform them into meaningful insights, and support informed decision-making in real-world scenarios. It reflects the growing importance of data literacy in today's digital era and demonstrates how structured analysis can contribute to transparency and better public understanding.

11. Acknowledgement

We would like to express our heartfelt gratitude to:

- **Infosys Springboard Team**

We would like to express our heartfelt gratitude to the **Infosys Springboard Team** for organizing and curating such a valuable and enriching internship opportunity. The structured learning environment, well-designed resources, and practical exposure provided through this program have significantly enhanced our understanding of real-world industry practices. This initiative has truly bridged the gap between academic learning and professional application.

- **Our Mentor**

We are especially thankful to **our mentor - Mrs. Nityashree S J** for her continuous guidance, constructive feedback, and unwavering support throughout the internship. Her insights, technical expertise, and encouragement helped us approach challenges with confidence

and clarity. The regular reviews and suggestions not only improved the quality of our work but also strengthened our analytical and problem-solving skills.

- **My Teammates**

We also extend our sincere appreciation to **our teammates (Bhavishya Priyadarshini V, Nikhitha, Jyoti Pathak, Jasleen Kaur)** for their collaboration, teamwork, and shared learning experience. The exchange of ideas, mutual support, and collective efforts made this journey more productive and enjoyable. Working together has enhanced our communication skills and taught us the importance of coordination and adaptability in a professional setting

This internship has been a significant milestone in our professional growth and technical journey. It has provided us with hands-on experience in data analysis, problem-solving, and project execution, while also helping us develop essential soft skills such as teamwork, time management, and effective communication. We are truly grateful for this opportunity and look forward to applying the knowledge and experience gained in our future endeavors.