Project Title:

Data Visualization of Electoral Trends Using ElectViz

Abstract

This project focuses on analyzing electoral data to understand voting trends, party performance, and constituency-level patterns over multiple election years. The goal is to transform raw election data into meaningful and interactive dashboards using data cleaning, modeling, and visualization techniques. The project ensures accurate insights through structured planning, sprint-based execution, and validation of data quality.

Objectives

- To explore and understand the structure of the election dataset

- To clean and preprocess data for better accuracy

- To create interactive dashboards for electoral trend analysis

- To calculate key performance metrics using DAX

- To provide data-driven insights for decision-making

Milestone 1: Dataset Exploration, WBS & Backlog

Dataset Exploration

The dataset includes:

- Election Year

- State and Constituency

- Candidate Names

- Political Parties

- Total Votes and Total Electors

- Vote Share Percentage

- Winner and Runner-Up details

- Winning Margin

While exploring the dataset, the following issues were identified:

- Inconsistent political party names (e.g., "INC" vs "Indian National Congress")

- Constituency name variations across election years

- Numeric fields stored as text values

- Incorrect vote share calculations in some records

Work Breakdown Structure

The project was divided into the following stages:

1. Requirement Gathering

2. Data Collection

3. Data Cleaning & Preprocessing

4. Data Modeling

5. Creating Page Dashboards

6. Developing DAX Measures

7. Testing & Review

8. Documentation Preparation

Backlog Creation

The product backlog included:

- Removing duplicate records

- Standardizing party names

- Creating fact and dimension tables

- Designing visual layouts

- Creating year-wise trend charts

- Adding slicers and interactivity

- Formatting the dashboard

- Preparing milestone documentation

Milestone 2: Data Modeling & Transformation – Detailed Overview

Milestone 2 focused on converting raw and unstructured election data into a clean, structured, and analytics-ready dataset. The main objective was to ensure that the data was accurate, consistent, and optimized for visualization in Power BI. This phase played a critical role in building a strong foundation for meaningful insights and dashboard creation.

During this milestone, extensive work was carried out on data cleaning, standardization, relationship building, and model optimization to support fast and reliable reporting.

Key Activities Completed in Milestone 2

The following tasks were successfully completed:

- Conversion of text-based numeric fields (such as vote counts and margins) into correct numeric data types

- Standardization of political party names, state names, and constituency names to maintain uniformity

- Removal of duplicate, null, and irrelevant records to improve data quality

- Creation of calculated columns and DAX measures for KPIs such as vote share, winning margin, and turnout percentage

- Design and implementation of a **star schema data model** with proper fact and dimension tables

- Establishment of relationships between tables to enable efficient filtering and slicing

- Validation checks to ensure data accuracy before moving to visualization

Tools & Technologies Used

- Microsoft Power BI

- Microsoft Excel

- Power Query Editor

- DAX (Data Analysis Expressions)

Requirement Gathering

Based on project needs, the following requirements were refined and implemented:

- Visualization of election trends across multiple election years

- Analysis of party-wise and state-wise performance

- Identification of winning margins and vote share patterns

- Interactive filtering using slicers (year, state, party, constituency)

- High data accuracy and consistency for reliable insights

Data Transformation Process

The structured transformation process included:

- Removing duplicate and irrelevant records

- Fixing inconsistent text values and spelling variations

- Converting incorrect data types (Text → Numeric / Date)

- Handling missing values using filters, conditional columns, and logical rules

- Creating new calculated columns to enhance analytical capability

- Optimizing the data model for better query performance

Sprint 1 – Data Cleaning & Preparation

- Imported raw datasets into Power BI and Excel

- Identified and resolved data quality issues (null values, duplicates, format mismatches)

- Standardized column names and removed irrelevant fields

- Performed initial data validation and consistency checks

Sprint 2 – Data Modeling & Initial Dashboard Development

- Designed and implemented fact and dimension tables

- Built a star schema data model

- Created relationships between tables for filtering and drill-down

- Developed key calculated columns and DAX measures such as:

  o Total Votes

  o Winning Margin

  o Vote Share Percentage

- **Created initial report pages in Power BI**, including:

- o Overview Dashboard

- o Year-wise Trends Page

- o Party-wise Performance Page

- Added basic slicers and filters for interactivity

Sprint 3 – Planned Activities

- Creation of additional dashboard pages such as:

  - o State-wise Performance Page

  - o Constituency-level Analysis Page

  - o Voter Turnout Analysis Page

- Enhancement of existing pages with advanced interactivity (drill-through, tooltips, bookmarks)

- Optimization of visuals for performance and user experience

- End-to-end testing and validation of dashboards

DAX Formulas Used

1. Total Votes Measure

Total Votes = SUM(ElectionData[TotalVotes])

2. Vote Share Percentage

Vote Share % =

DIVIDE(

   SUM(ElectionData[CandidateVotes]),

   SUM(ElectionData[TotalValidVotes])

) * 100

3. Winning Margin

Winning Margin =

SUM(ElectionData[WinnerVotes]) -

SUM(ElectionData[RunnerUpVotes])

4. Total Seats Won

Total Seats Won =

CALCULATE(

    COUNT(ElectionData[Constituency]),

    ElectionData[Position] = "Winner"

)