

VISUALIZING ELECTION

DATA FOR A NEWS COMPANY

PRESENTED BY:
TEAM 4
BHUMI KALA
SANTOSH SK
SRINIBAS MASANTA
VAIBHAV YADAV



TOPIC OUTLINE

INTRODUCTION

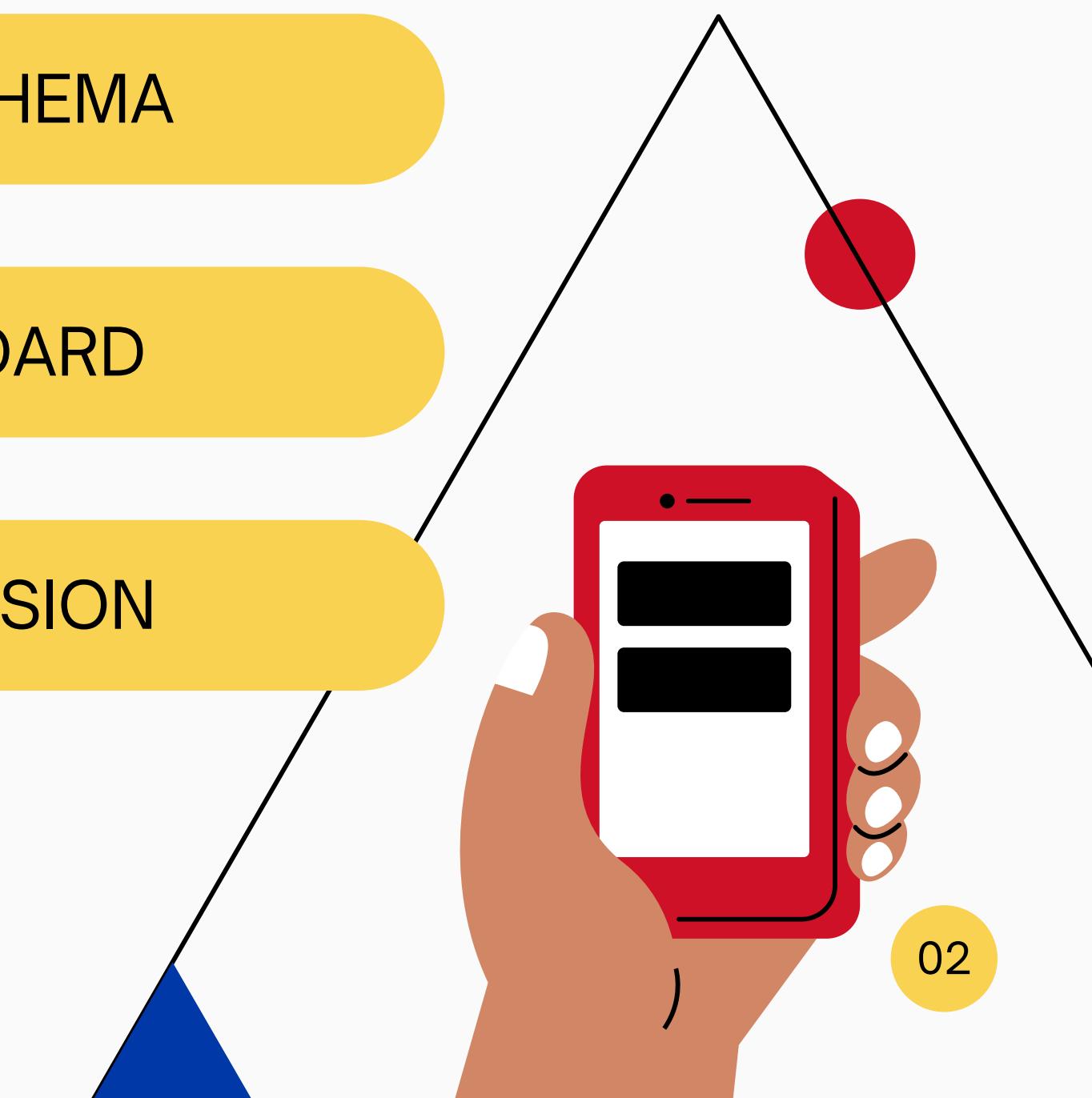
STAR SCHEMA

DATASET OVERVIEW

DASHBOARD

DATA CLEANING

CONCLUSION



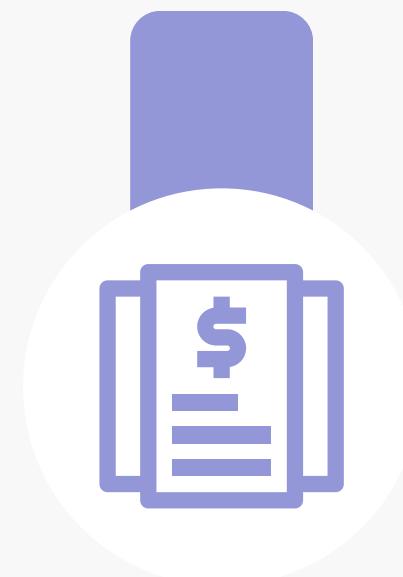
INTRODUCTION

In today's fast-paced election environment, news agencies need accurate and real-time visualizations to effectively report voting trends and results. This project addresses that need by integrating complex election data from multiple sources into Power BI dashboards. These dashboards provide clear insights into vote shares, turnout patterns, party trends, and constituency-level performance, enabling journalists and editorial teams to analyze and present election data with clarity and speed. The solution is designed to be scalable and adaptable, supporting real-time updates and detailed demographic analysis to enhance news coverage and public understanding.



WHY USE POWER BI?

Power BI is Microsoft's business intelligence tool that transforms raw data from multiple sources into interactive, real-time, and visually compelling insights for better decision-making.



Data Integration Power

- Multiple Sources: Connects to CSV, Excel, APIs, databases, and real-time feeds.
- Unified Data: Combines historical and live election data in one platform.



Visualization & Storytelling

- Interactive visuals for complex electoral data.
- Easy for journalists to interpret vote shares, turnout, and maps.



Cost & Usability Advantage

- Cost-effective alternative to enterprise BI tools like Tableau or Qlik.
- User-friendly with a low learning curve, ideal for non-technical staff.



Real-Time Analytics

- Live Data: Supports live data refresh for real-time election reporting.
- Timeliness: Crucial for high-stakes, time-sensitive environments like elections.



Data Modeling Capabilities

- Supports star schema design, relationships, hierarchies, and measures (DAX).
- Handles large-scale datasets efficiently, ensuring accuracy.

ABOUT THE DATASET

"The dataset contains detailed information on parliamentary and assembly constituencies, including IDs, names, and state mapping. It forms the backbone of the star schema, enabling constituency-level analysis of votes, candidates, and parties."

Constituency Information

- Contains details about each parliamentary and assembly constituency.
- Includes unique IDs, constituency names, and mapping to states.

State & Region Mapping

- Each constituency is linked with its state and region.
- Enables state-level comparisons and aggregated insights.

Electoral Boundaries

- Records boundary definitions to track constituency-level changes across elections.
- Important for historical consistency and trend analysis.

Demographic Coverage

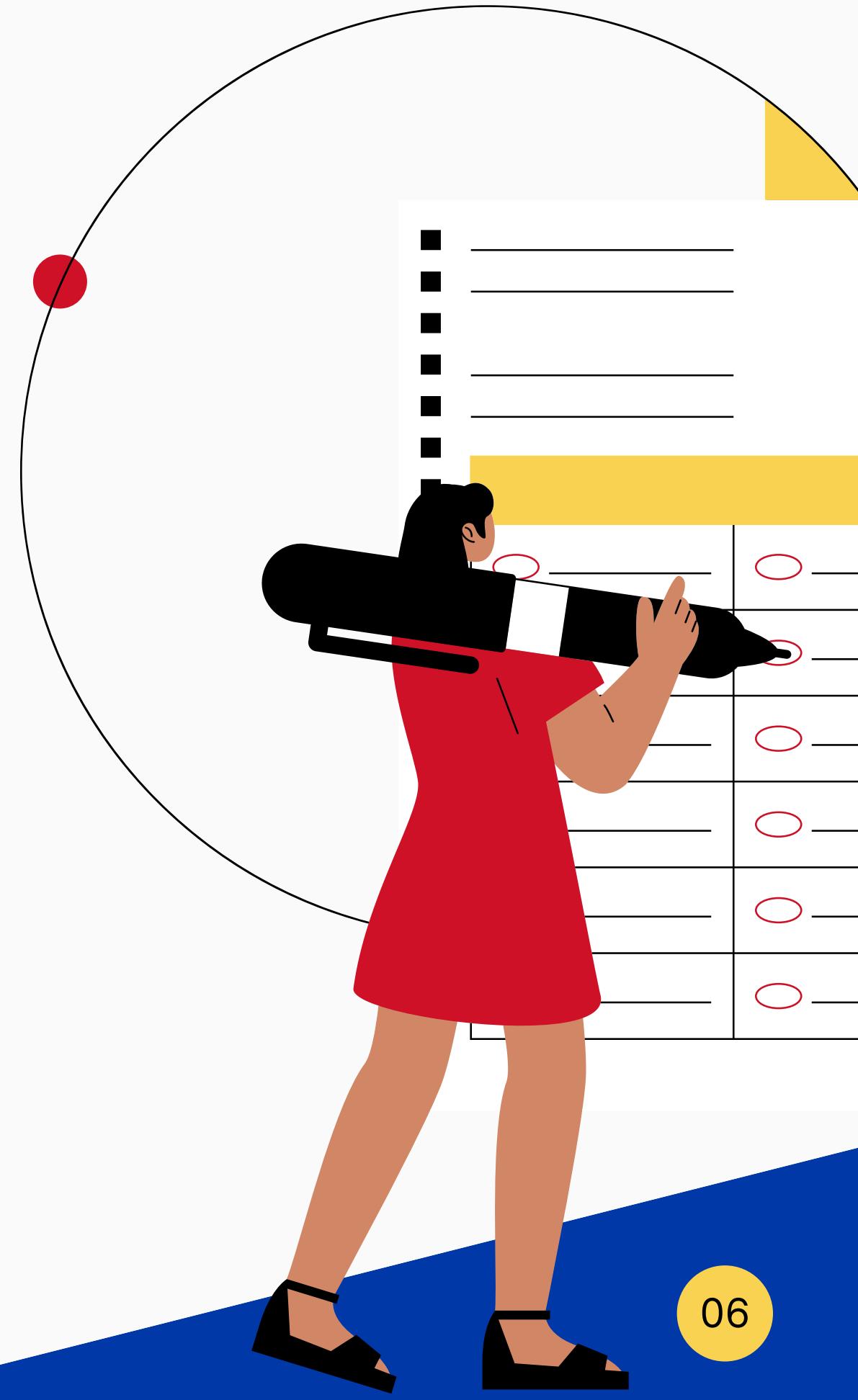
- Captures urban–rural distribution and other constituency-level attributes.
- Supports deeper insights into voter behavior and turnout trends.

Analytical Relevance

- Serves as a dimension table in the Power BI star schema.
- Forms the backbone for linking votes, candidates, and parties to their respective constituencies.

DATA CLEANING

Data cleaning ensures the constituency dataset is accurate, consistent, and standardized by removing duplicates, handling missing values, and validating mappings before integration into Power BI.



DATA CLEANING & TRANSFORMATION IN POWER BI (POWER QUERY)

Loksabha_1962-2019 (1).csv

The screenshot shows the 'File Origin' tab of the Power Query Editor. It displays the file path '1252: Western European (Windows)', the delimiter as 'Comma', and the data type detection set to 'Based on first 200 rows'. Below these settings, a preview of the data table is shown, containing columns such as 'Pc_name', 'no', 'type', 'state', 'candidate_name', 'party', and 'elector'. The preview shows several rows of data from the CSV file.

- Set correct data types for all columns (Year → Number, Votes → Whole Number, etc.)
- Created derived columns such as Vote Share (%) and Result (Winner/Loser)
- Validated final data model consistent, complete, and ready for dashboard integration.
- Renamed columns for clarity (e.g., Cand_name → Candidate Name)
- Removed duplicates & null entries to ensure accuracy.

- Imported & inspected raw CSV (Loksabha 1962–2019) in Power Query Editor
- Standardized text using Trim, Clean, and Capitalize Each Word
- Unified naming for parties and constituencies (e.g., “BJP”, “Bharatiya Janata Party” → “Bharatiya Janata Party”)
- Set correct data types for all columns (Year → Number, Votes → Whole Number, etc.)
- **Created derived columns** such as Vote Share (%) and Result (Winner/Loser)
- Validated final data model consistent, complete, and ready for dashboard integration.

The screenshot shows the 'Transform' tab of the Power Query Editor. The 'Queries' pane on the left lists nine queries: 'lok_sabha_pc_metrics', 'lok_sabha_results_clean', 'dim_party', 'dim_candidate', 'dim_constituency', 'dim_election_date', 'fact_votes', 'assembly', and 'Loksabha_1962-2019 (1)'. The main workspace shows a table with four columns: 'Pc_name', 'no', 'type', and 'state'. The data in the table corresponds to the preview shown in the first screenshot.

DATA INTEGRATION THROUGH API

- API = Application Programming Interface
- Acts as a bridge between data source and application
- Enables real-time, structured, and automated access to datasets
- Industry value: APIs reduce manual data handling and ensure consistency.

Source: [Datameet India Election Data API]

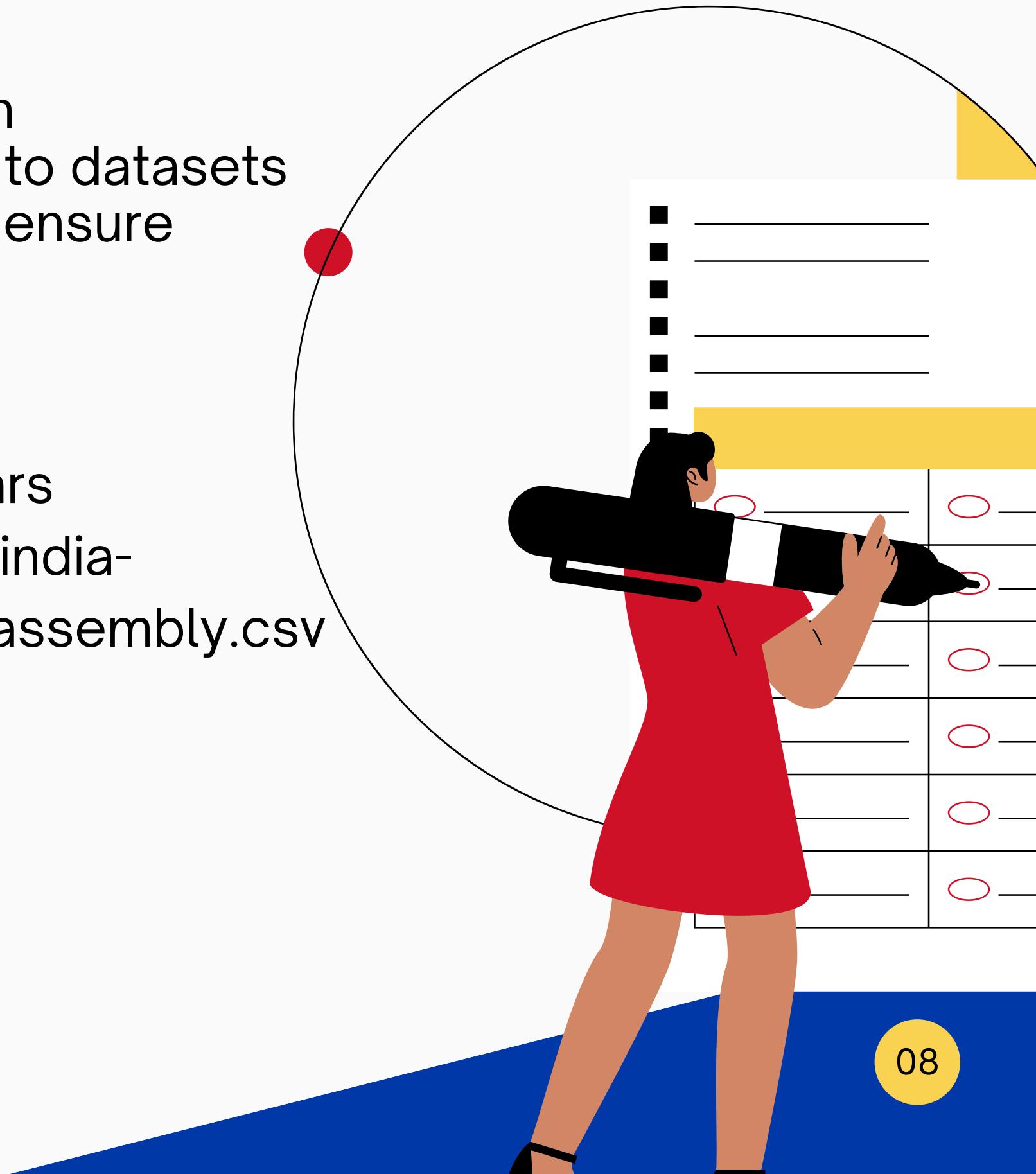
Taken from Github

Provides assembly election data across states and years

API Link: <https://raw.githubusercontent.com/datameet/india-election-data/refs/heads/master/assembly-elections/assembly.csv>

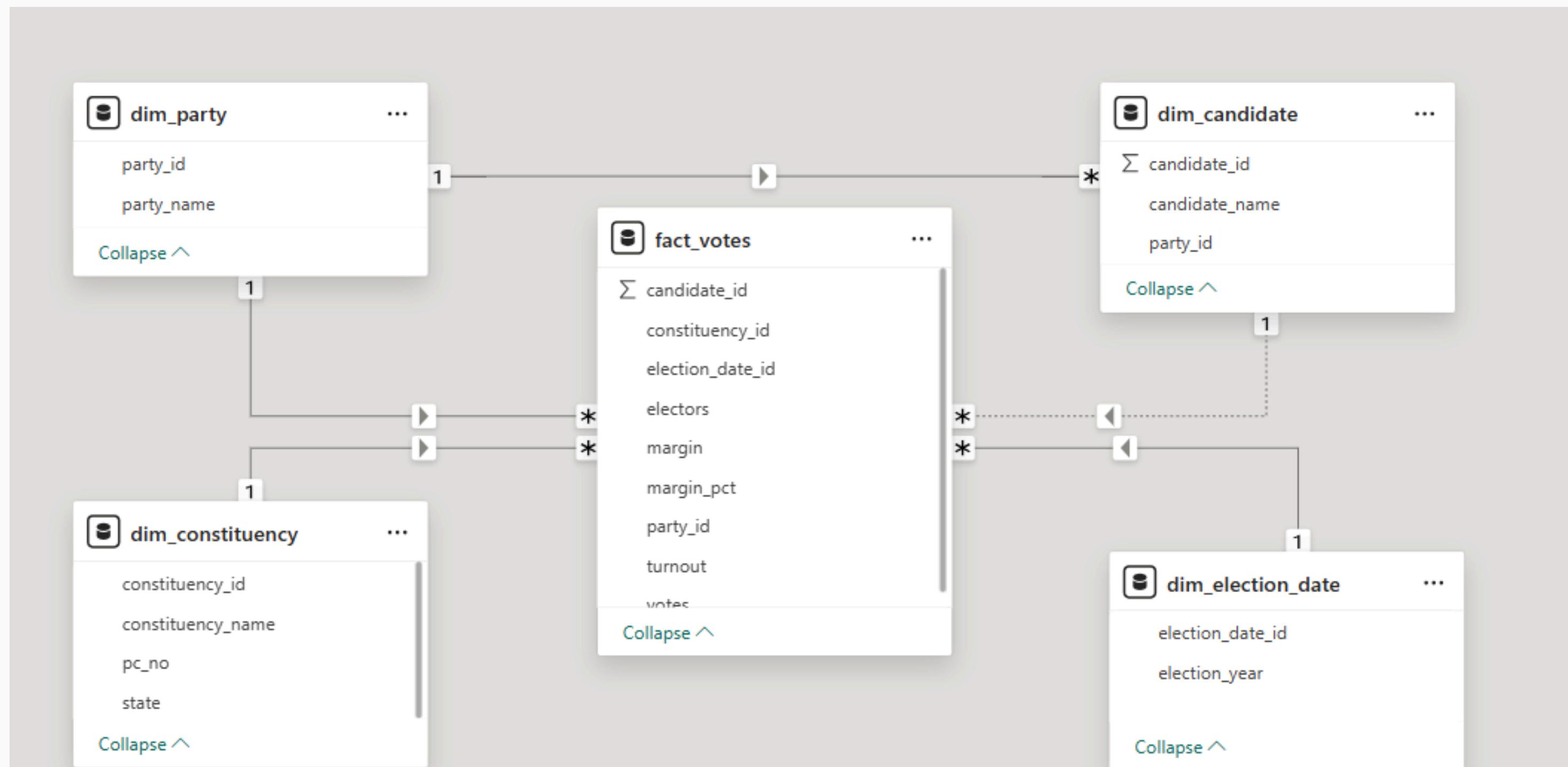
Data points include:

- State & constituency
- Candidate & party
- Votes received
- Election year



STAR SCHEMA

- A star schema is a data warehousing model with a central fact table linked to multiple dimension tables, making data easier to analyze and query.
- The fact table (`fact_votes`) sits at the center and stores measurable data like votes, turnout, and margins, with references to dimension tables.
 - Dimension tables (`dim_party`, `dim_candidate`, `dim_constituency`, `dim_election_date`) describe context such as parties, candidates, constituencies, and election dates.
 - Each dimension table connects to the fact table using foreign keys, enabling flexible filtering and grouping.



HISTORICAL TRENDS DASHBOARD



- Elections are not just numbers; they are stories of shifting public trust.
- Historical vote shares and turnout patterns reveal:
 - What drives voter behavior?
 - Where do parties rise or decline?
 - Which regions decide outcomes?
- Our dashboard turns raw data into narratives of democracy.

DASHBOARD

Historical Trends Analysis Dashboard

Stronghold Seats Count

2944

Swing Seats Count

4165

Neutral Seats Count

894

No. of Parties Participated

145

Dominant Party

Indian National Congress

Year

All

State

All

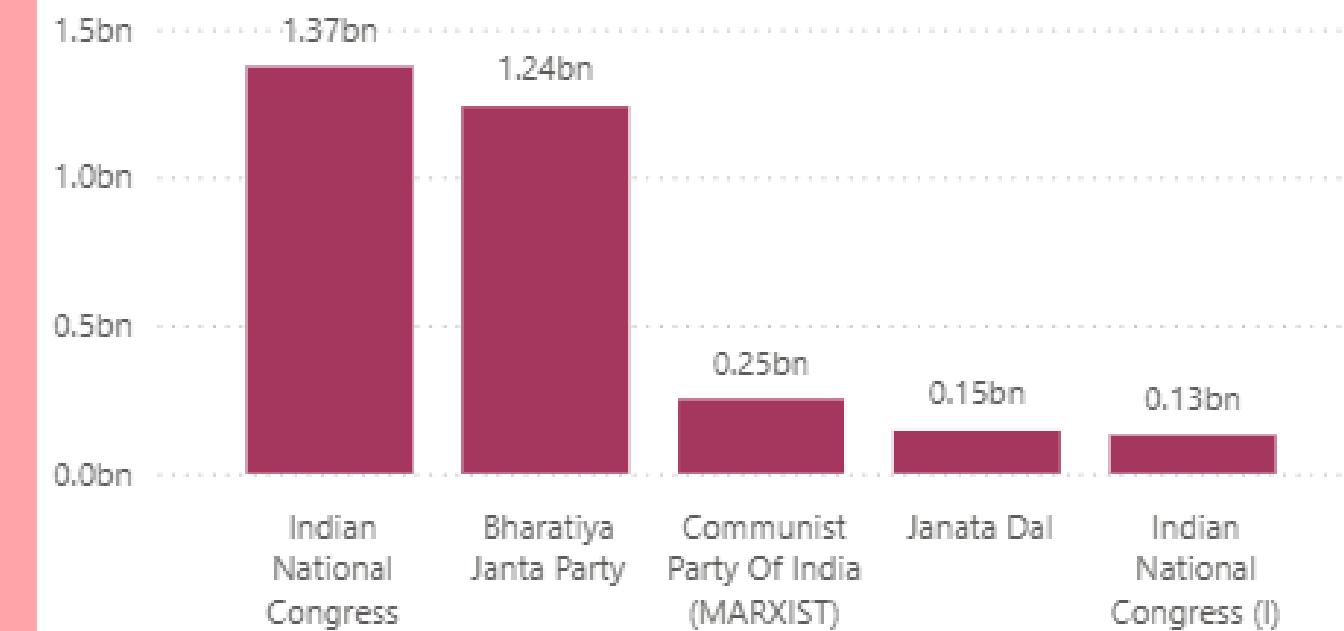
Region

All

Party

All

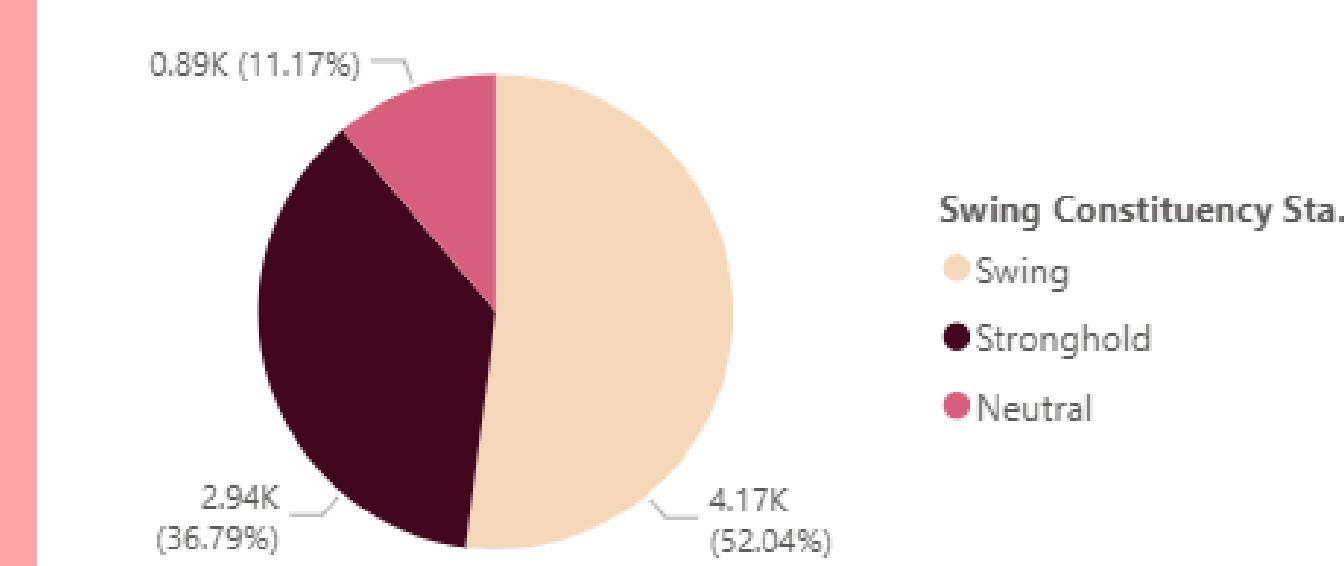
Sum of Votes by Party



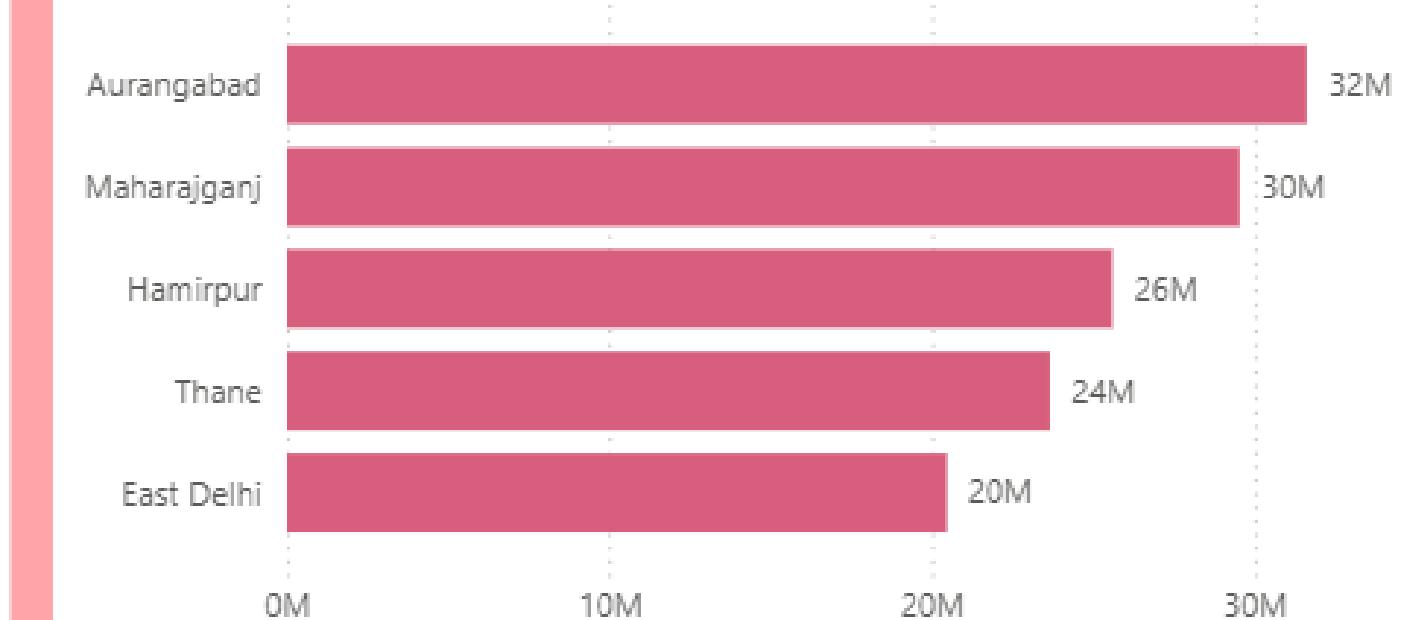
Average Turnout % by Year



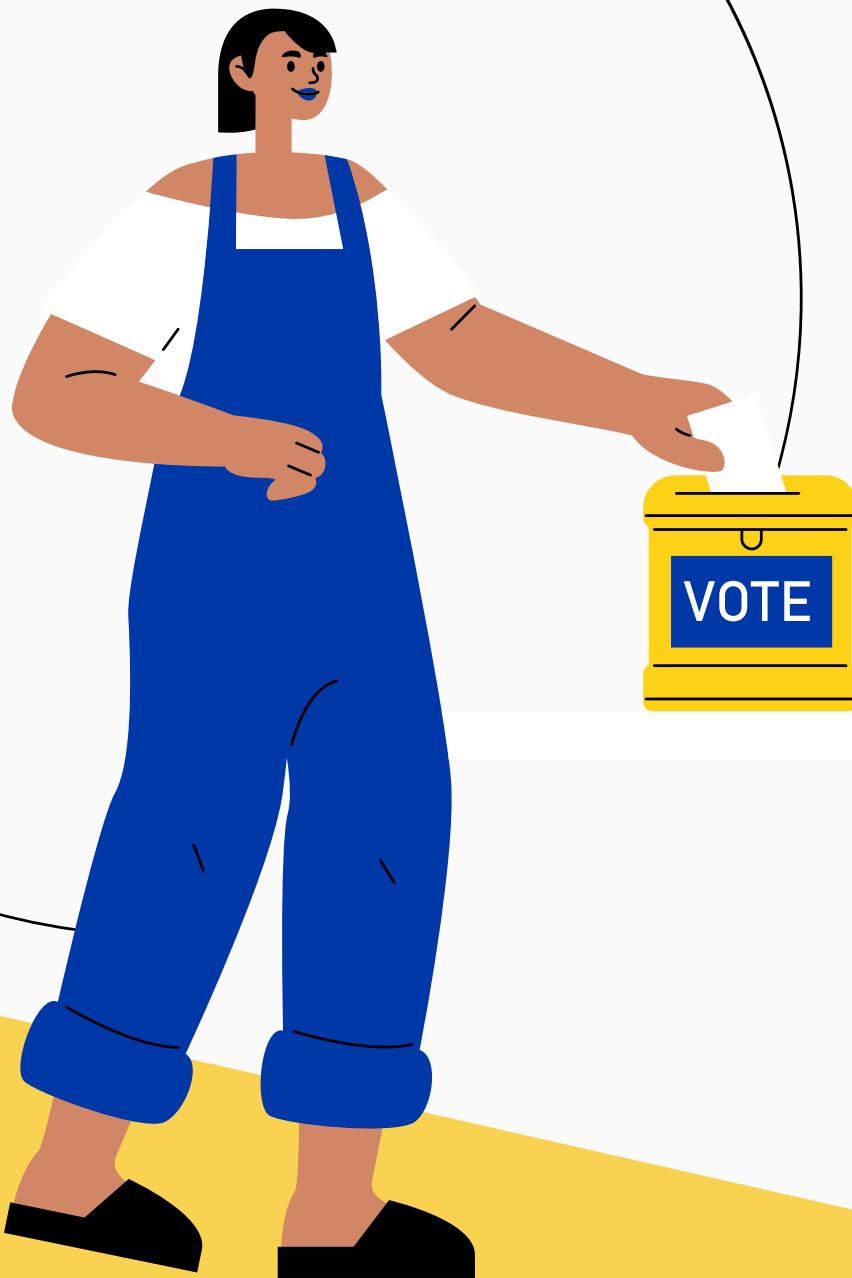
Swing Constituency Status Distribution



Electors Count by Region



DEMOGRAPHIC TRENDS



Lok Sabha Election Insights

- Scale: 834M electors, 8,047 candidates, 2,562 constituencies, 145 parties.
- Key Findings:
 - Mulayam Singh Yadav represents the largest electorate (12M+), followed by other political heavyweights.
 - Candidate gender breakdown reveals male dominance across parties; female representation remains low.
 - Seats won are distributed across key constituencies, allowing identification of major strongholds and local dynamics.

DASHBOARD

Demographics Dashboard

Total Electors

834M

Total Candidates

8047

Total Constituencies

2562

No. of Parties Participated

145

State

All

Region

All

Party

All

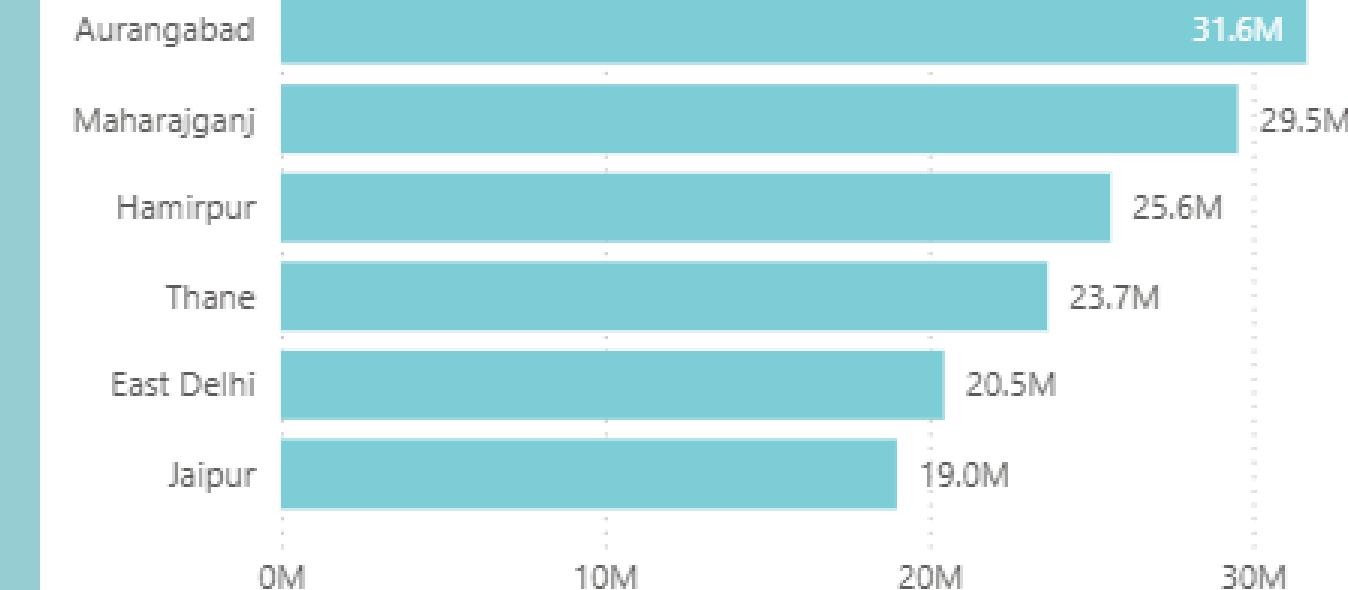
Year

All

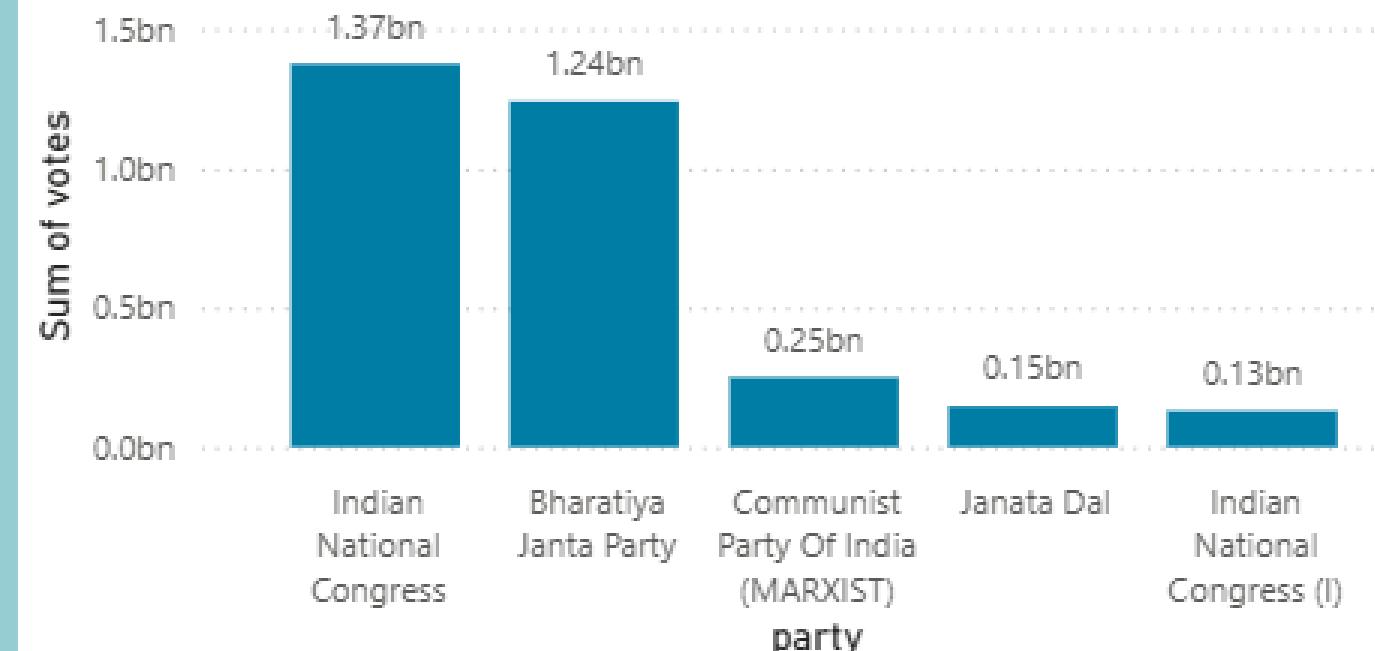
Average Margin % by Year



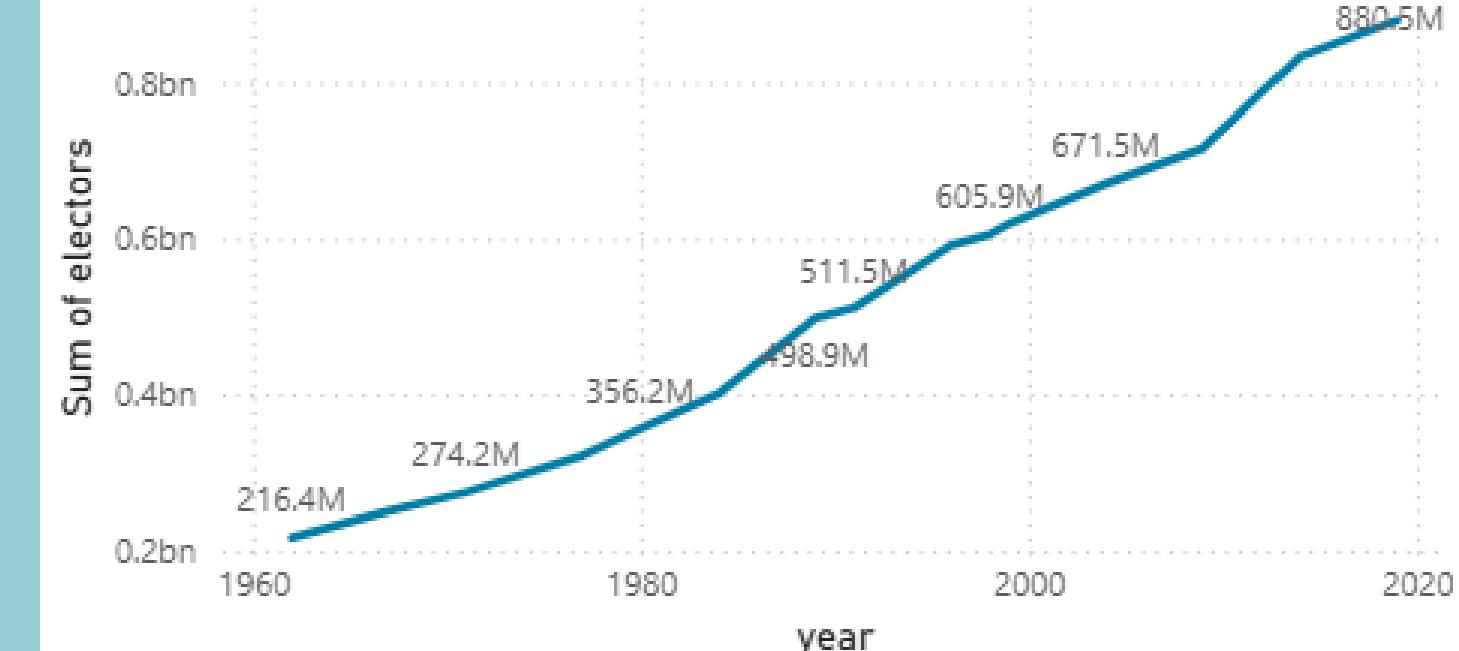
Top 6 Region by Total Electors



Top 5 parties by Total Votes



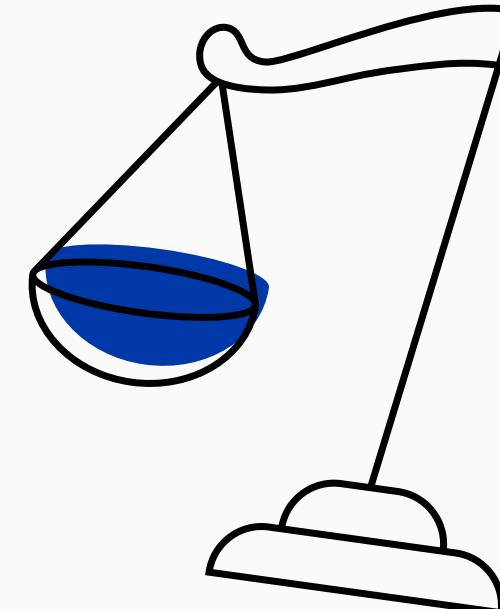
Electors Trend by Year



RURAL VOTER DYNAMICS

Why Is This Important?

Rural voters often determine election outcomes in India due to their significant population share and potential for high engagement. Some states like Uttar Pradesh, Bihar, and Maharashtra have rural voter numbers in tens of millions, shaping national and local election strategies. High turnout percentages in states such as Nagaland and Lakshadweep reveal areas with exceptional voter mobilization.



WHAT IS DRILL THROUGH?

Drill Through enables navigation from a summary dashboard to a detailed report page focused on a specific item like a region, candidate, or product.

When viewers right-click or select a data point, Power BI automatically opens a target page filtered to show granular details for that item only.

This makes analysis deeper and interactive, letting users explore the underlying reasons behind summary insights without cluttering the main dashboard.

Example: From overall state results, drill through to see details of a single state's voter turnout, candidate breakdown, or sales figures.

Why use it?

- Enables focused, actionable insights.
- Keeps dashboards clean while letting users investigate details as needed.

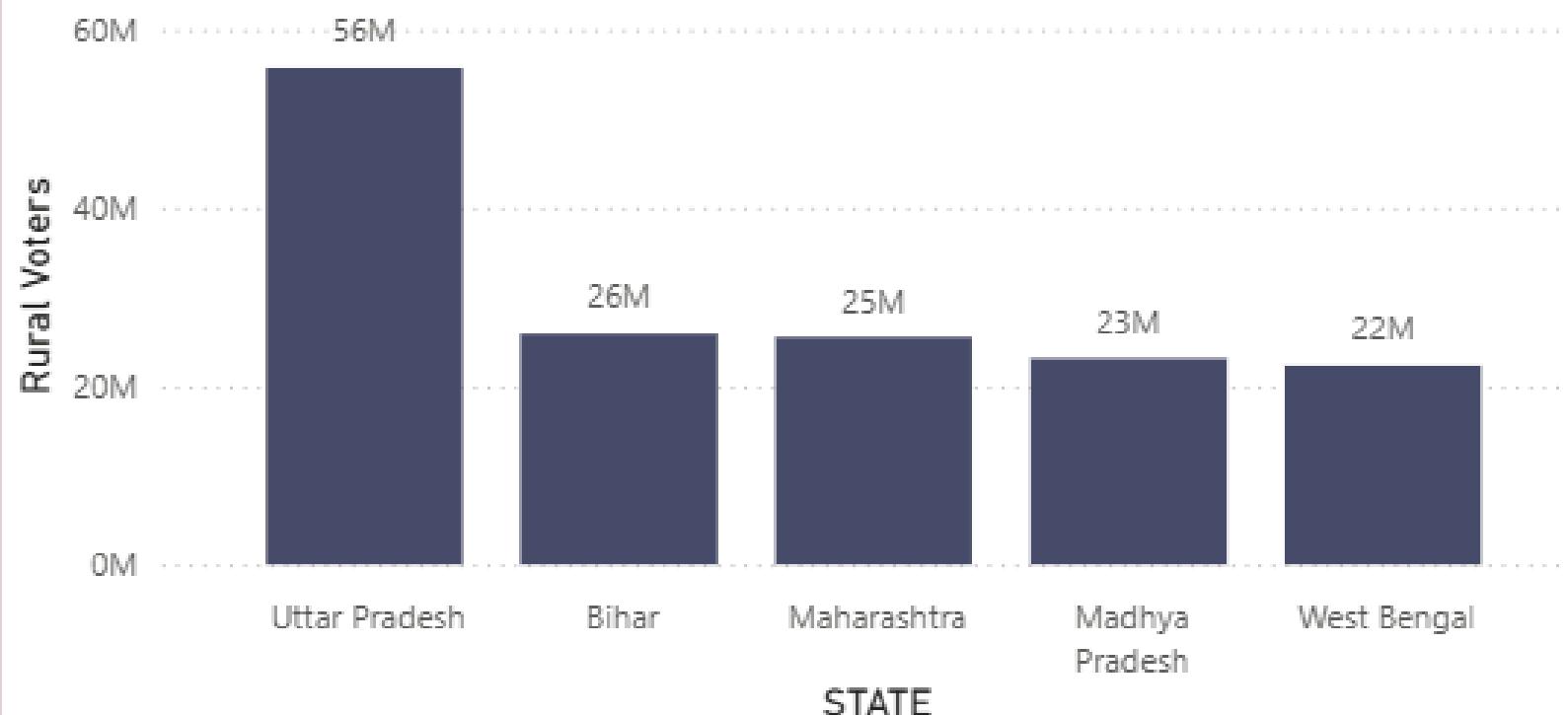




Rural Region Analysis

STATE	Sum of Total voters	Rural Voters	Rural Turnout %
Andaman & Nicobar Islands	190346	190346	0.71
Andhra Pradesh	15766113	15766113	0.76
Arunachal Pradesh	600828	600828	0.79
Assam	9976538	9976538	0.81
Bihar	25843709	25843709	0.57
Chhattisgarh	8694109	8694109	0.70
Dadra & Nagar Haveli	165324	165324	0.84
Daman & Diu	87233	87233	0.78
Goa	410495	410495	0.75
Gujarat	1101623	1101623	0.72
Haryana	3446801	3446801	0.74
Himachal Pradesh	835744	835744	0.67
Jharkhand	9743729	9743729	0.64
Karnataka	1133326	1133326	0.73
Kerala	820267	820267	0.71
Lakshadweep	43242	43242	0.87
Madhya Pradesh	23107981	23107981	0.62
Maharashtra	25478792	25478792	0.64
Manipur	773817	773817	0.84
Mizoram	434962	434962	0.62
Nagaland	1039962	1039962	0.88
Odisha	14557364	14557364	0.75
Punjab	6397188	6397188	0.72
Rajasthan	16977410	16977410	0.62
Sikkim	310095	310095	0.84
Total	267915173	267915173	0.66

Top 5 States By Rural Voters



Top 5 States By Rural Turnout %



URBAN VOTER LANDSCAPE: SCALE AND ENGAGEMENT



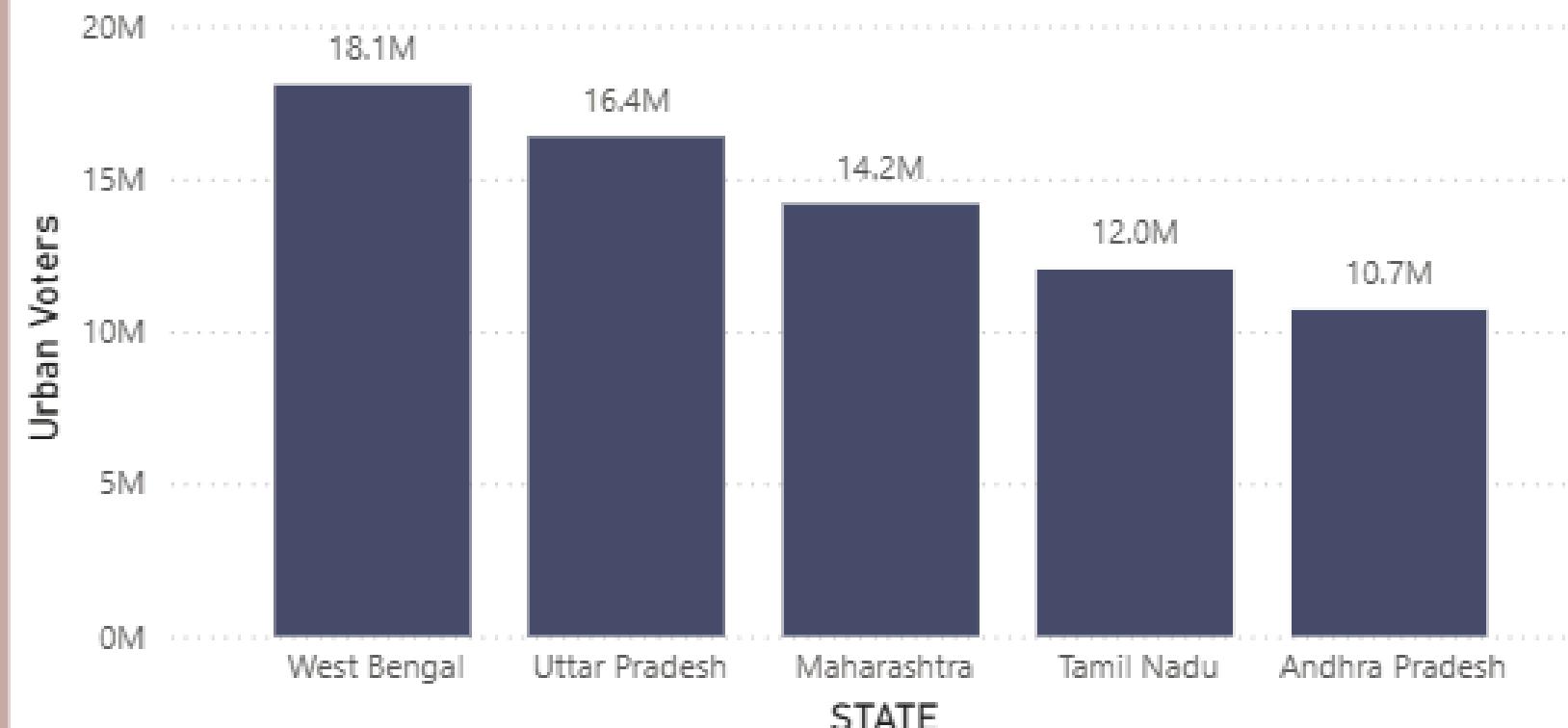
Urban regions are critical for electoral victory due to their dense populations and issue-driven voting patterns. West Bengal, Uttar Pradesh, and Maharashtra top the chart in urban voter numbers, but turnout remains modest at 65% lower than rural India's rate. Despite higher literacy and awareness, urban apathy is evident; however, places like Puducherry and West Bengal show turnout above 80%. Boosting urban participation is essential for fair representation and unlocking the full power of India's democracy.



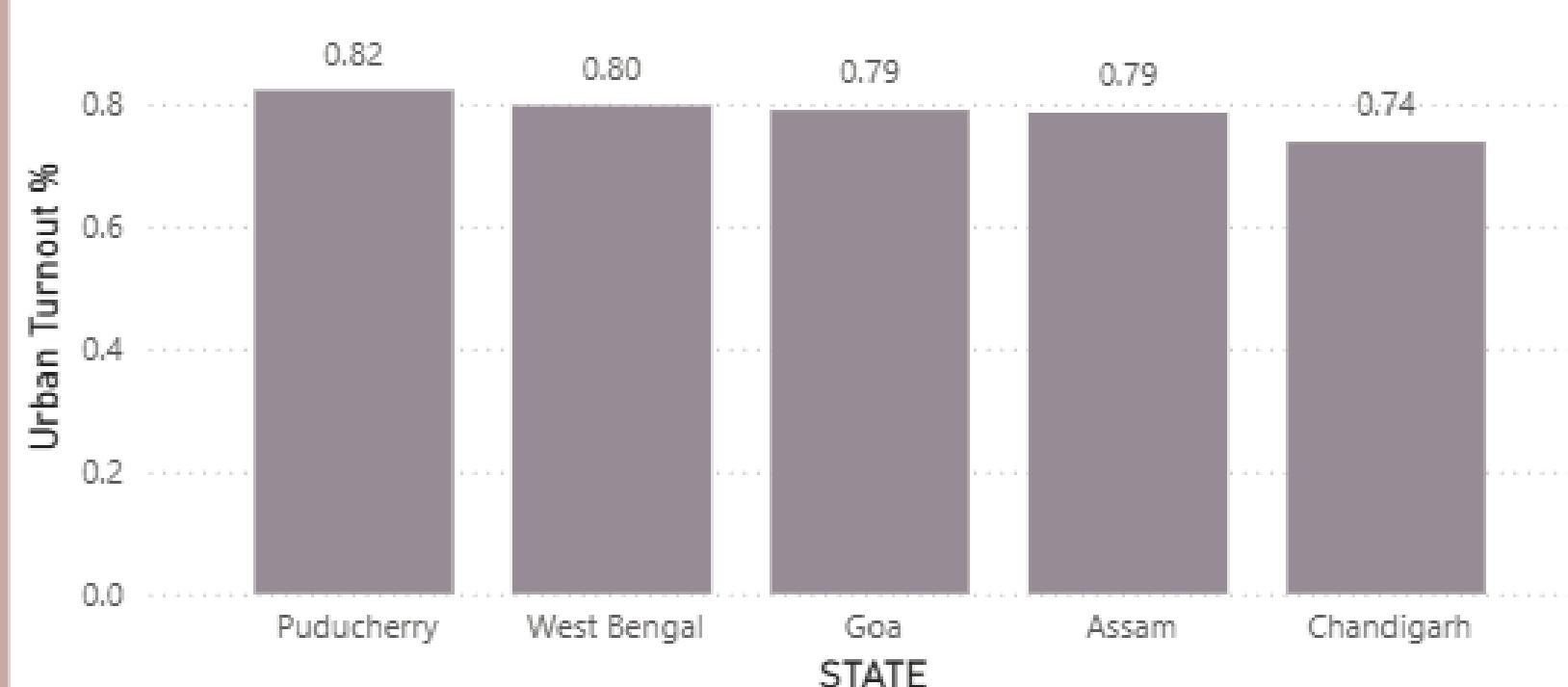
Urban Region Analysis

STATE	Sum of Total voters	Urban Voters	Urban Turnout %
Andhra Pradesh	10683453	10683453	0.66
Assam	4315791	4315791	0.79
Bihar	2647905	2647905	0.51
Chandigarh	453462	453462	0.74
Chhattisgarh	2510014	2510014	0.67
Goa	406945	406945	0.79
Gujarat	6257434	6257434	0.65
Haryana	8054450	8054450	0.70
Himachal Pradesh	738047	738047	0.64
Jammu & Kashmir	1568774	1568774	0.51
Jharkhand	3242896	3242896	0.63
Karnataka	4706142	4706142	0.58
Kerala	6192557	6192557	0.73
Madhya Pradesh	5456993	5456993	0.59
Maharashtra	14165893	14165893	0.55
Meghalaya	620041	620041	0.63
NCT OF Delhi	8275146	8275146	0.65
Odisha	3763144	3763144	0.68
Puducherry	740053	740053	0.82
Punjab	5447998	5447998	0.71
Rajasthan	6989560	6989560	0.65
Tamil Nadu	12018608	12018608	0.69
Uttar Pradesh	16390040	16390040	0.60
West Bengal	18072208	18072208	0.80
Total	143717554	143717554	0.65

Top 5 States By Urban Voters



Top 5 States By Urban Turnout %



DASHBOARD

Urban v/s Rural Voting Trends Dashboard

Average Poll %

67.32

Urban Turnout %

0.65

Rural Turnout %

0.66

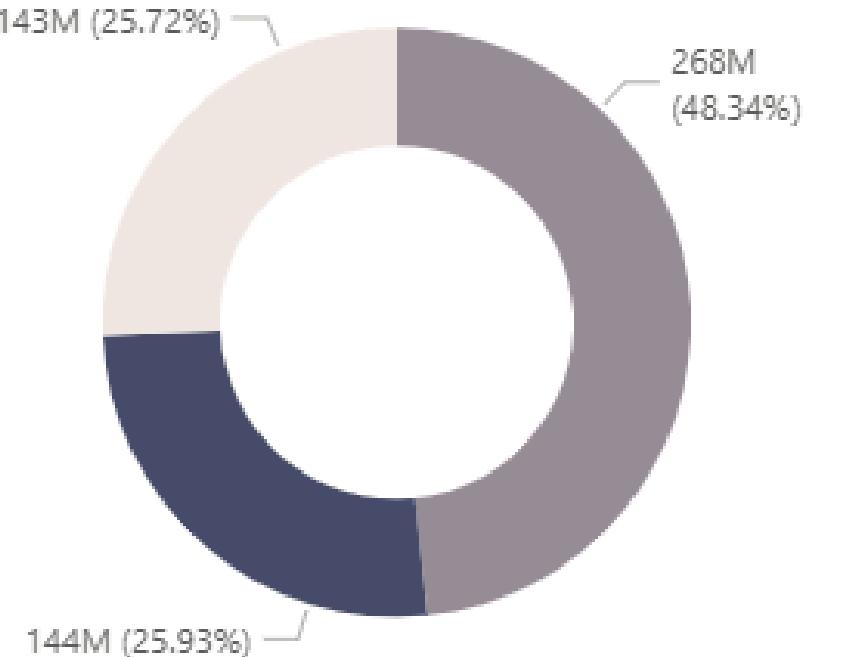
State & Constituency

- ✓ Andaman & Nicobar .
- ✓ Andhra Pradesh
- ✓ Arunachal Pradesh
- ✓ Assam
- ...

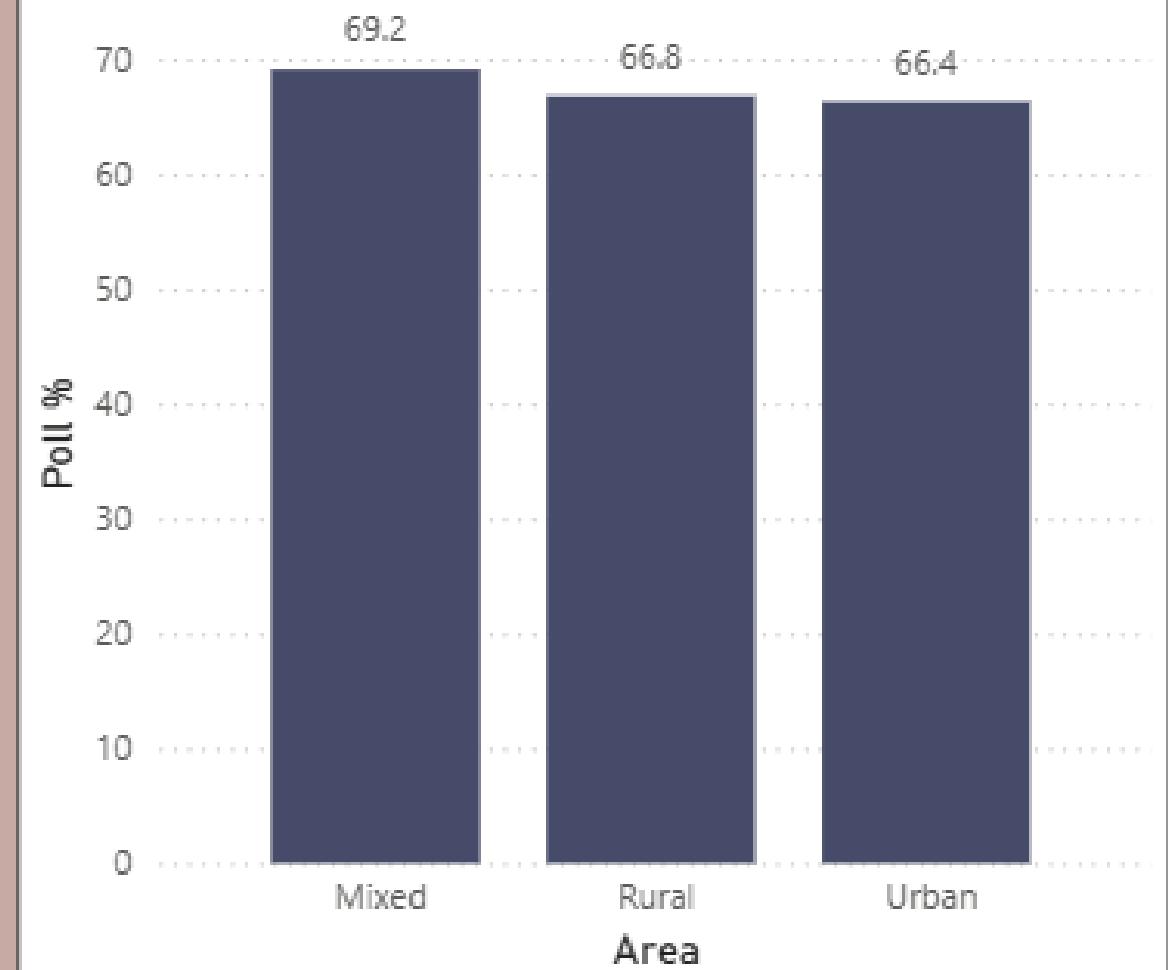
State



Total Voters by Area



Average Poll % by Area

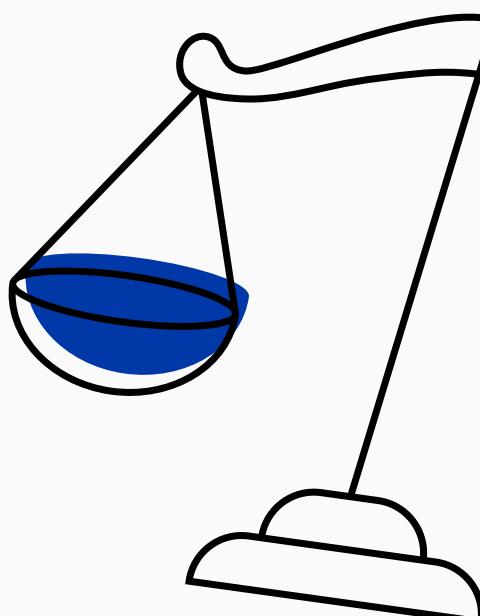


**YOUR VOTE
MATTERS**

CONCLUSION

This project demonstrates how advanced data analytics and visualization in Power BI can transform raw election data into actionable insights for policymakers, campaigners, and the public. By integrating demographic breakdowns, urban and rural region analyses, and interactivity features like drill-through, the dashboard reveals both macro and micro-level trends in Indian elections.

Key takeaways include the dominance of certain states and candidates in the electorate, the persistent gender gap, and the contrasting turnout dynamics between rural and urban India. The project highlights untapped opportunities to boost urban voter participation and better represent demographic diversity.



THANK YOU FOR LISTENING!

