

ElectViz Election Data Visualization for Media

Project: ElectViz – Multi-phase Indian election analysis with interactive Power BI dashboards.

Data Collection & Cleaning

Sources

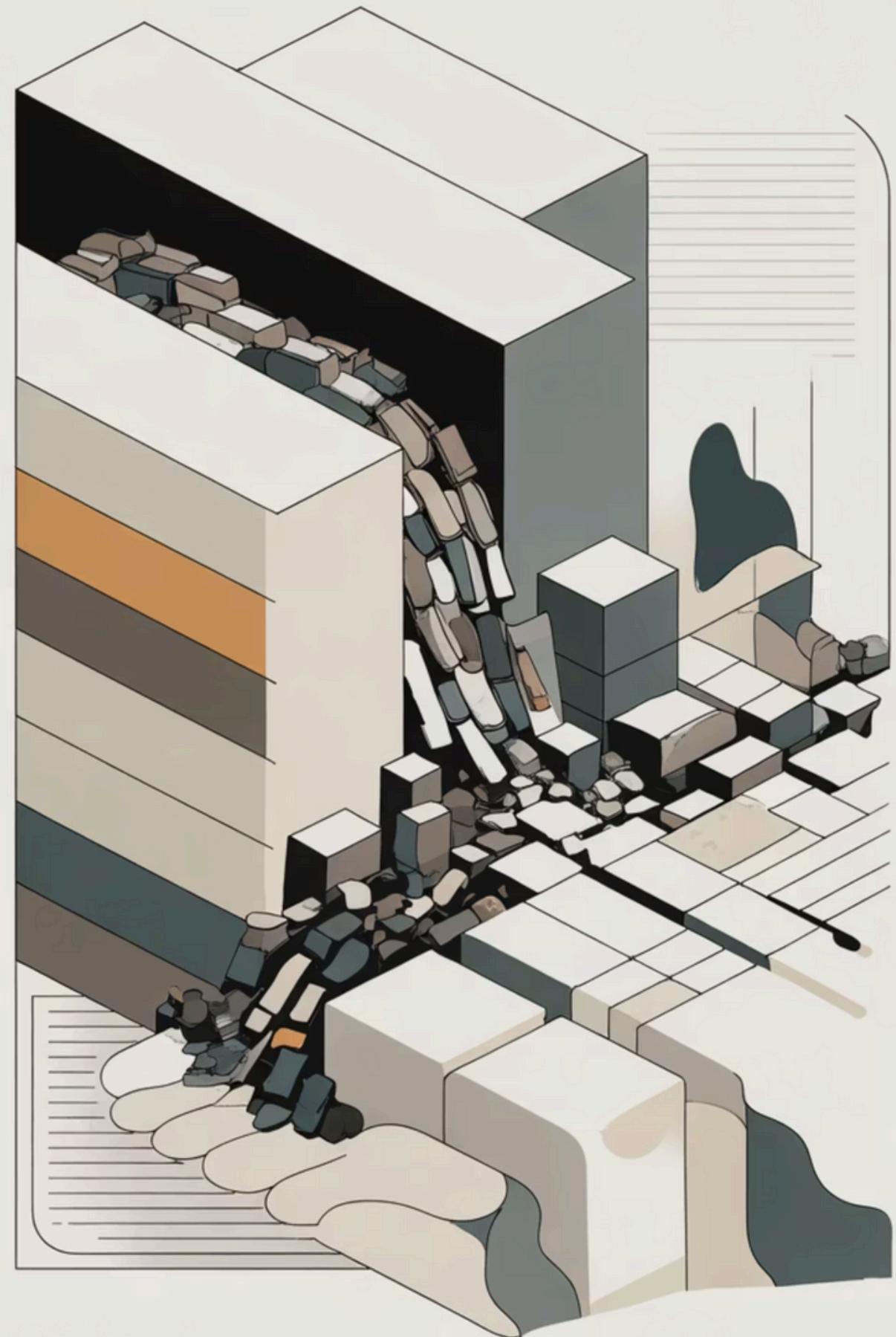
Kaggle election datasets providing state, constituency, party, candidate, votes, electors, phase mapping, and postal/EVM splits.

Power Query Steps

removed null/blank rows, trimmed/cleaned text, enforced data types, deduplicated on composite keys (State–Constituency–Party–Year–Phase), and filtered out totals/“Grand Total” records.

Normalization

standardized state/UT names, harmonized party abbreviations, and split EVM vs Postal votes into separate numeric columns.



Data Modeling

Star Schema

FactVotes (measures per state/constituency/phase) at grain State-Constituency-Party-Phase; dimensional tables for Date, State, Party, Candidate, Phase.

One big table in the center with numbers (the fact), surrounded by small descriptive tables (the dimensions).

Relationships

Every dimension has a primary key (e.g., StateKey, PartyKey). The fact stores those as foreign keys (StateKey, PartyKey, ConstituencyKey, PhaseKey, DateKey).

one-to-many from dimensions to FactVotes; active relationship on StateKey and PhaseKey; Date table marked as date for time intelligence.

Calculated Keys

Use integer surrogate keys in each dimension (e.g., StateKey = 101 for "Uttar Pradesh") instead of joining on names; this avoids spelling/casing mismatches and speeds joins.

Create these keys during Power Query/ETL and carry them into the fact.surrogate integer keys for stable joins; bi-directional filtering avoided except where slicer sync was required.

DAX (Calculated Columns & Measures)

Calculated Columns

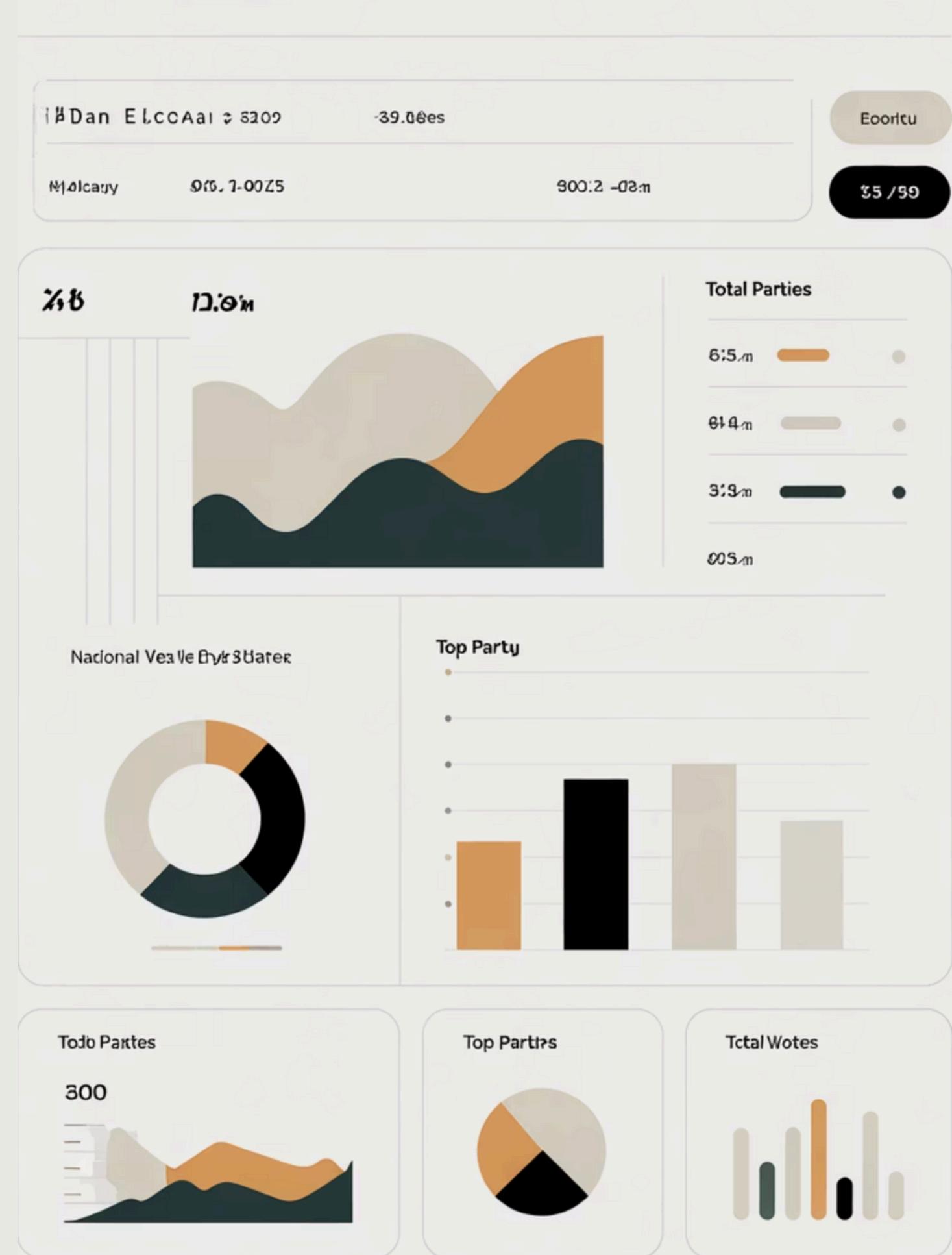
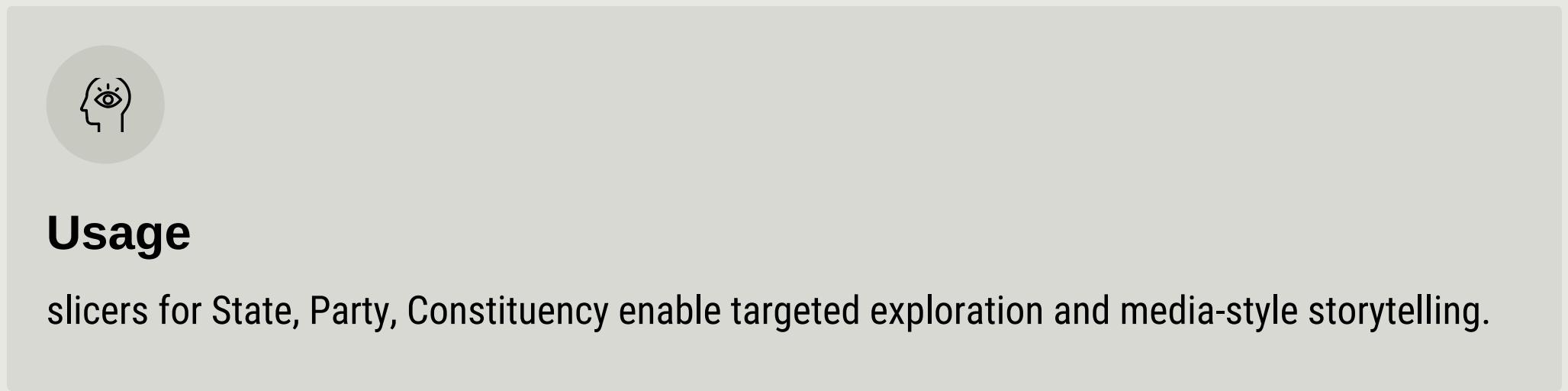
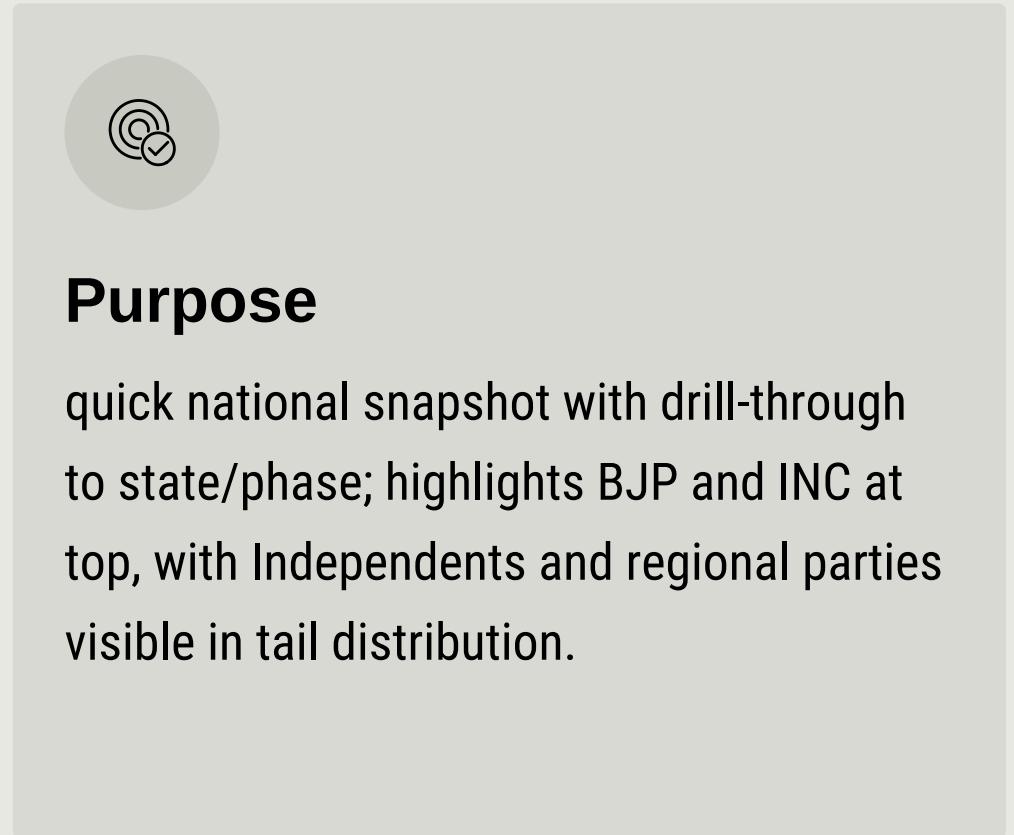
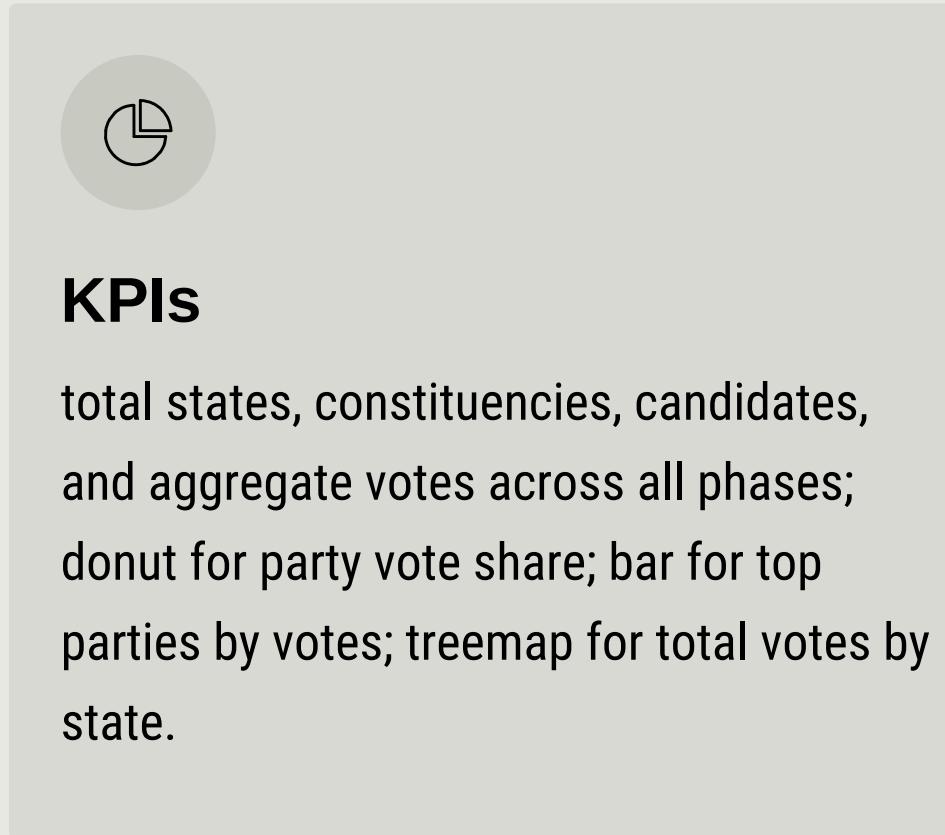
- Margin
- total states in different phase
- VoteType (EVM/Postal)
- CleanParty for grouping and top-N.

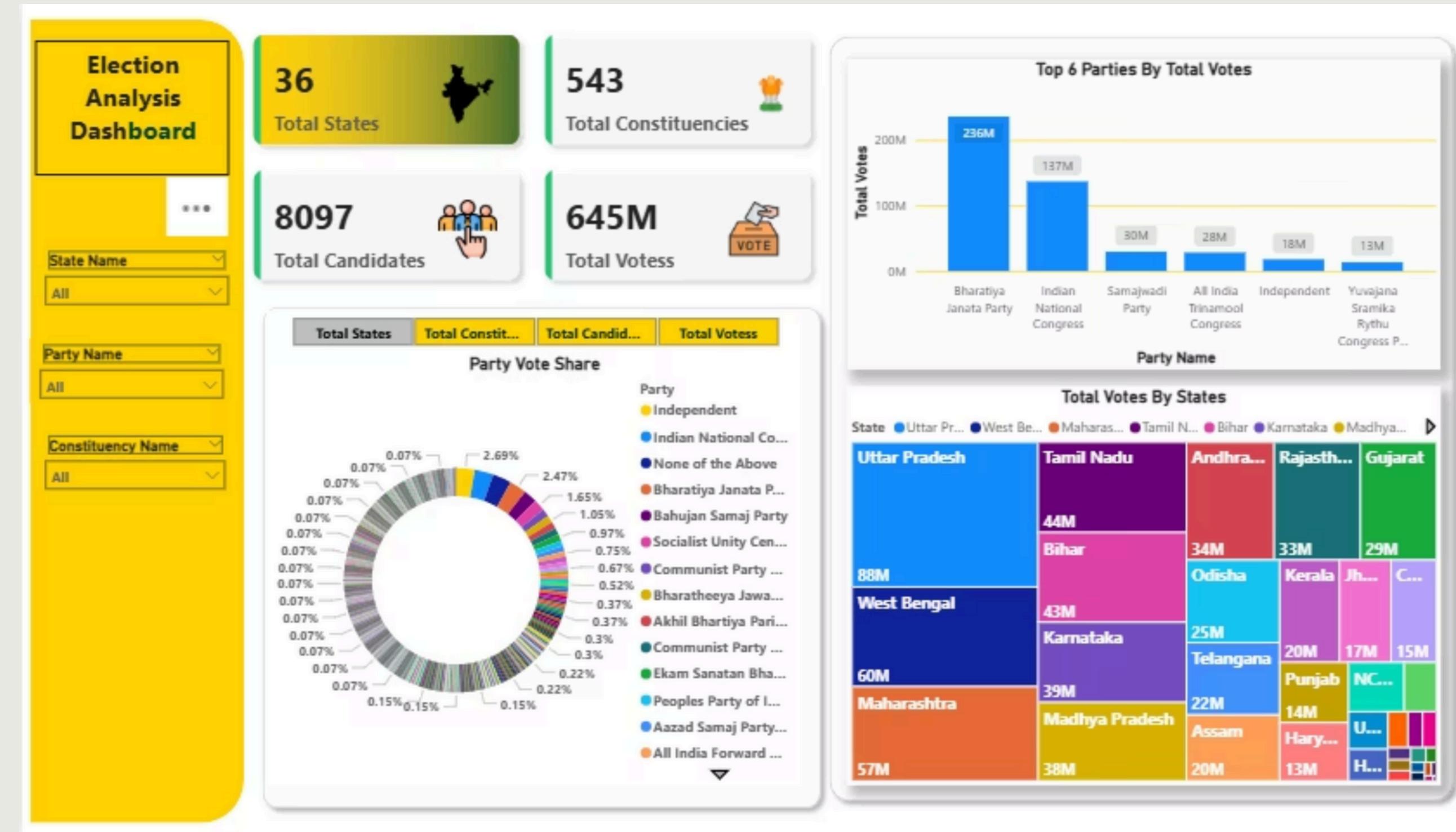
Core Measures

- **Total Difference = `SUM('Counted vs polled'[Difference])`**
- **Total EVM Votes Counted = `SUM('Counted vs polled'[EVM Votes Counted])`**
- **Total PC'S = `DISTINCTCOUNT('Counted vs polled'[PC Name])`**
- **Total Candidates = `DISTINCTCOUNT(eci_data_2024[candidate])`**
- **Total Constituencies = `DISTINCTCOUNT(eci_data_2024[constituency])`**
- **Top N Parties by Votes** using RANKX for leaderboard visuals.

Visual logic: phase-aware measures using TREATAS/ISFILTERED to reflect slicer selection and persist KPI cards.

Main Dashboard (National Overview)





Second Dashboard (Counting & Margins)

Focus: EVM votes polled vs counted by state, postal votes distribution, and margin by state heat/category bars.

EVM

Votes Polled vs Counted

Total EVM Votes Polled vs Counted, Total Difference.

- ❑ Insight: surfaces states with larger deltas between polled and counted and where victory margins cluster, aiding operational and results-night narratives.

Postal

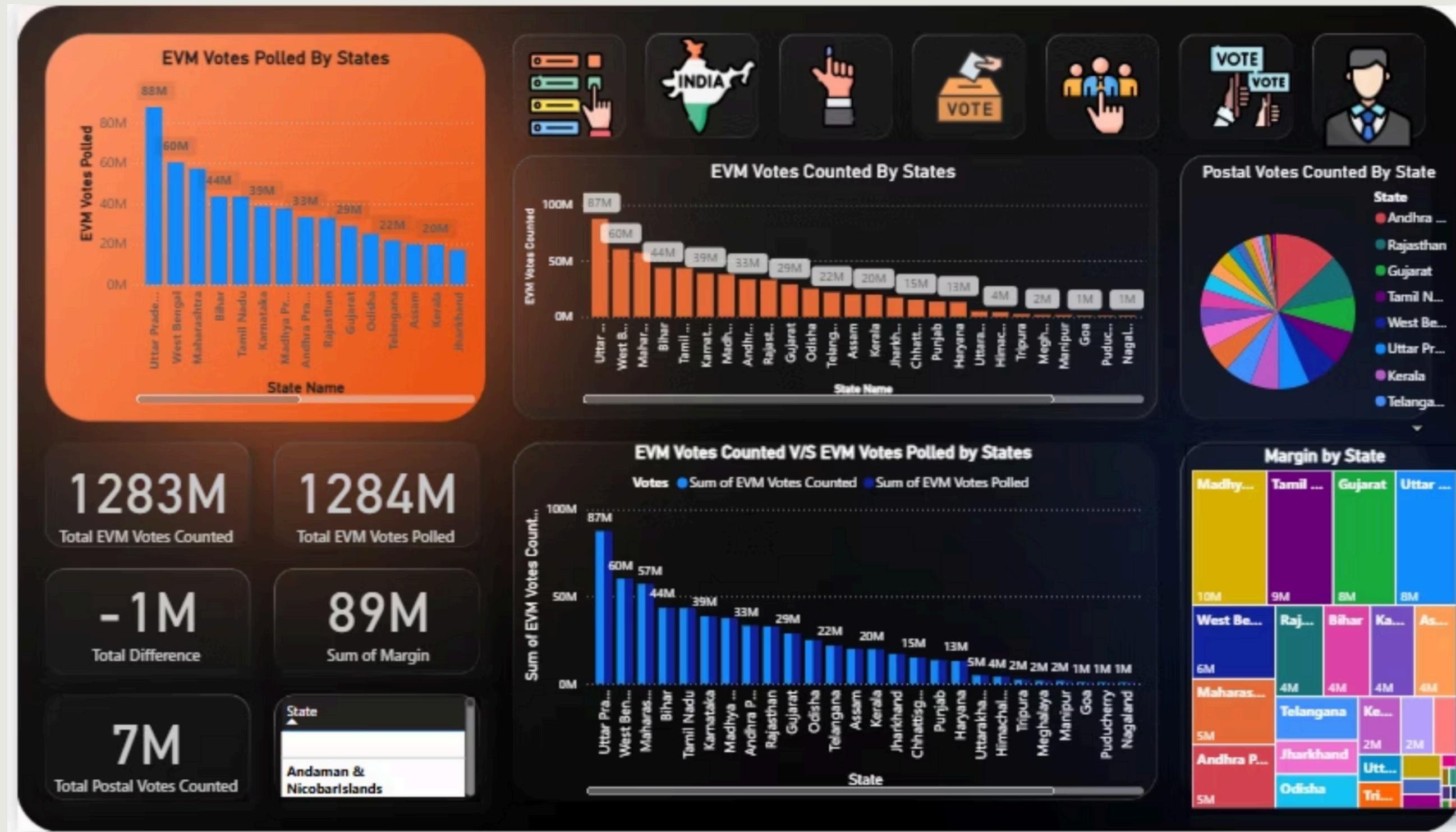
Votes Distribution

Total Postal Votes.

Margin

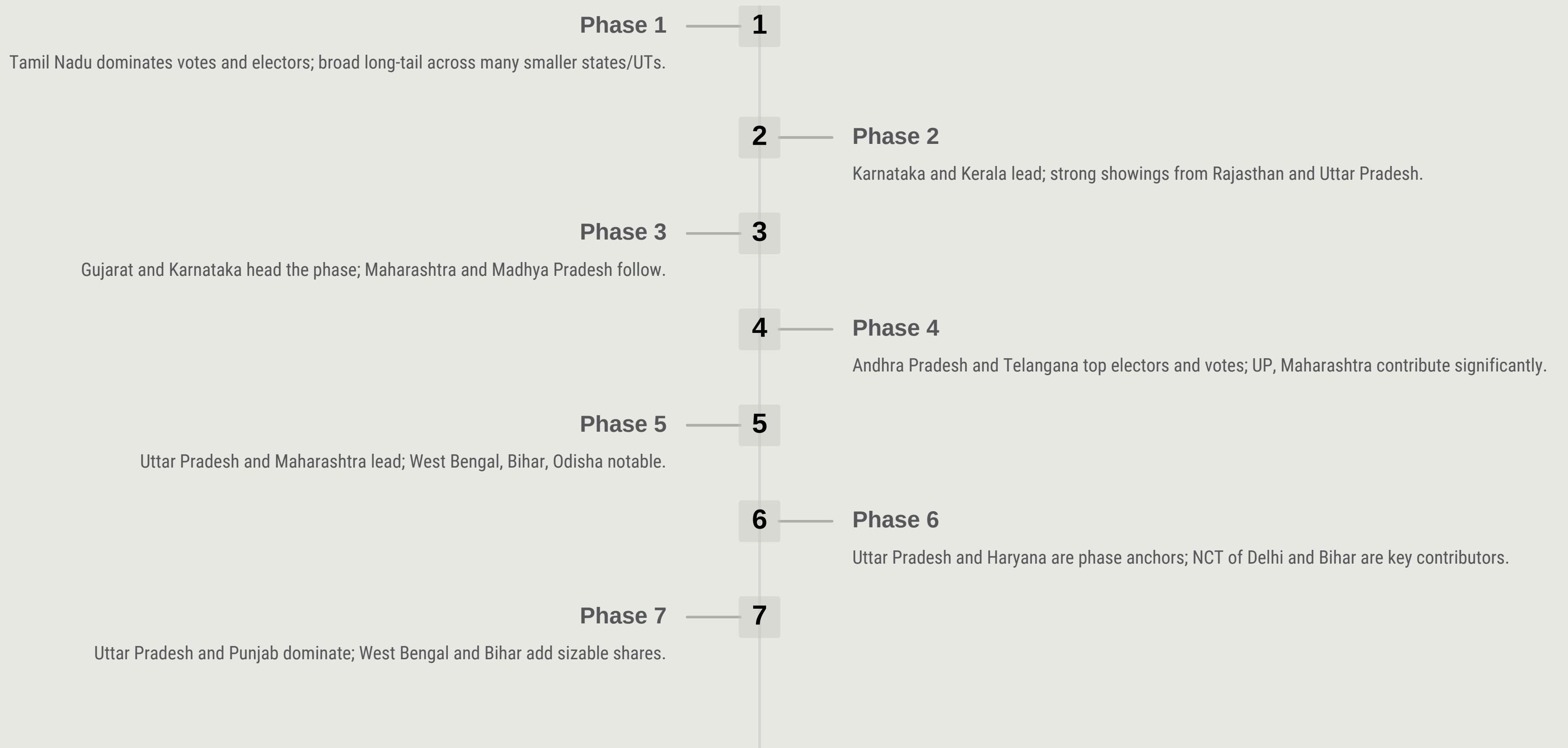
Analysis

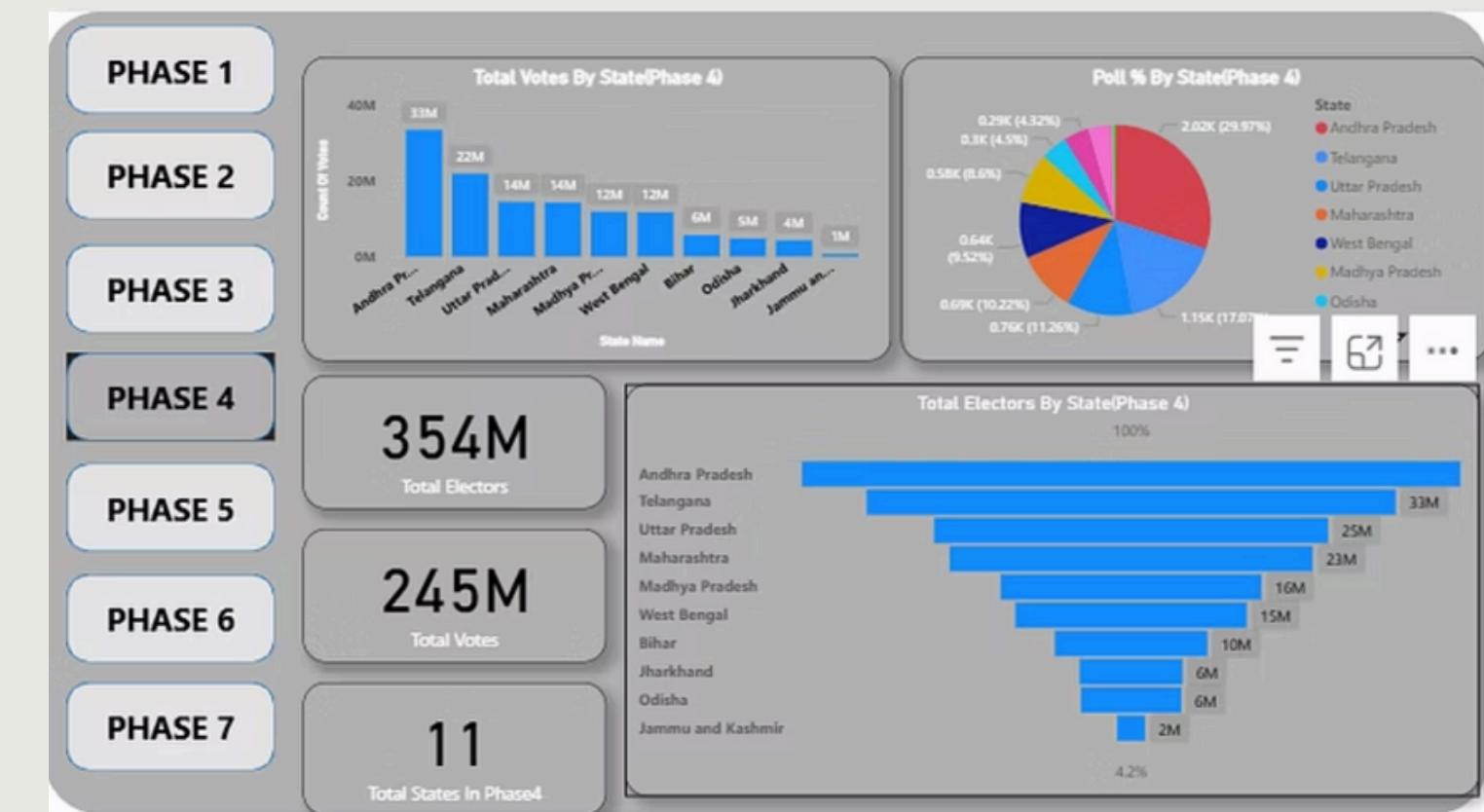
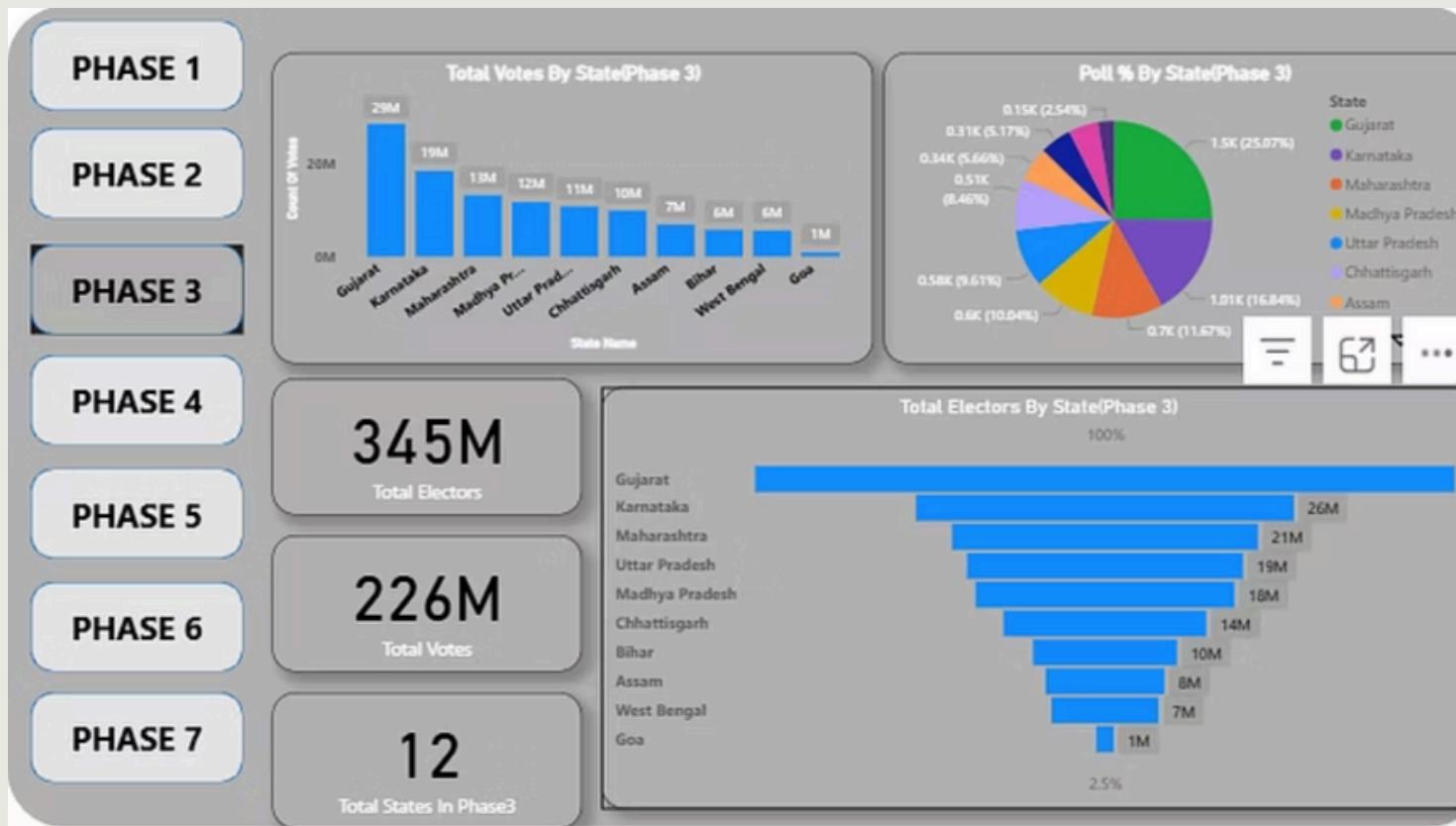
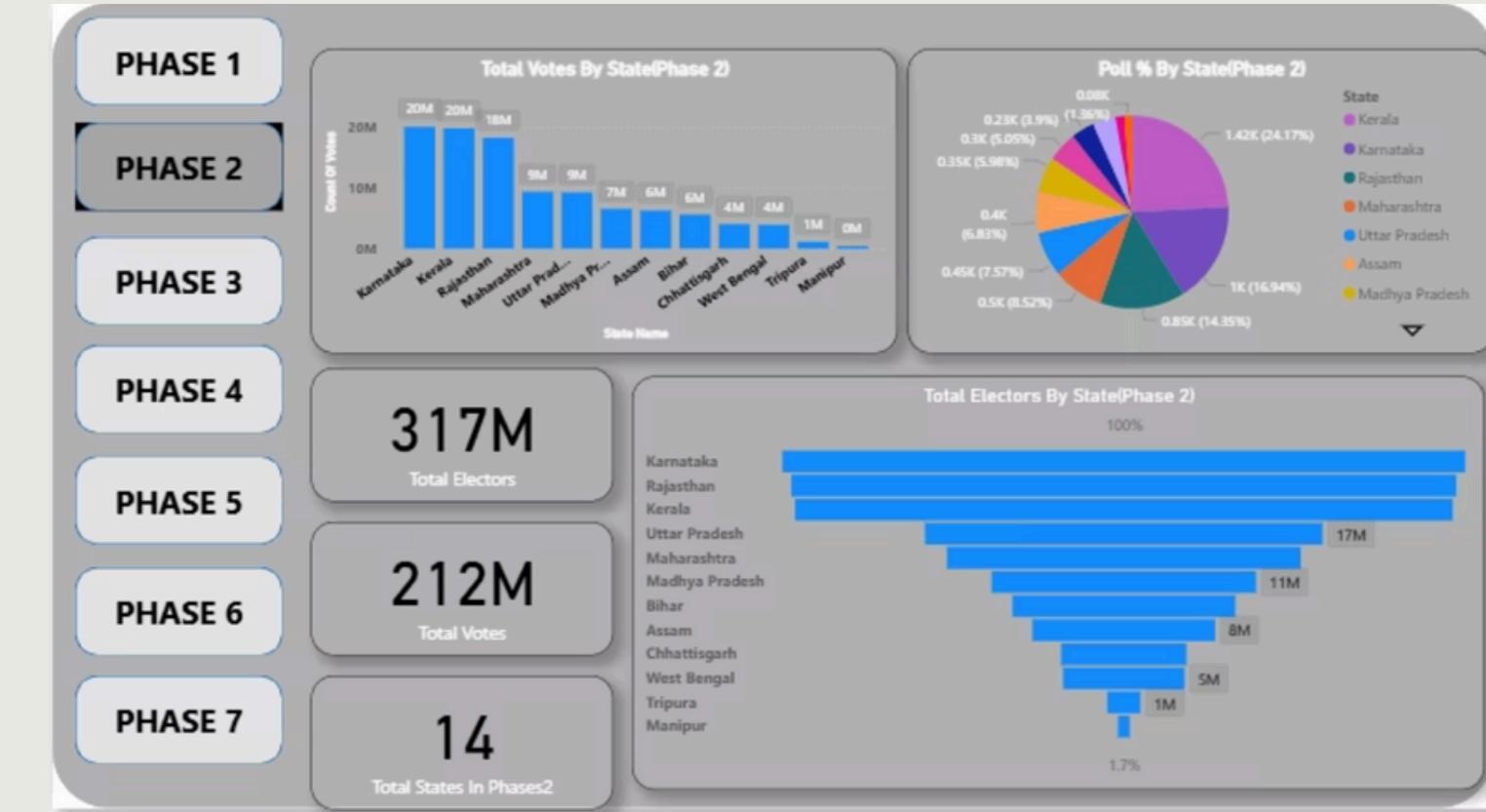
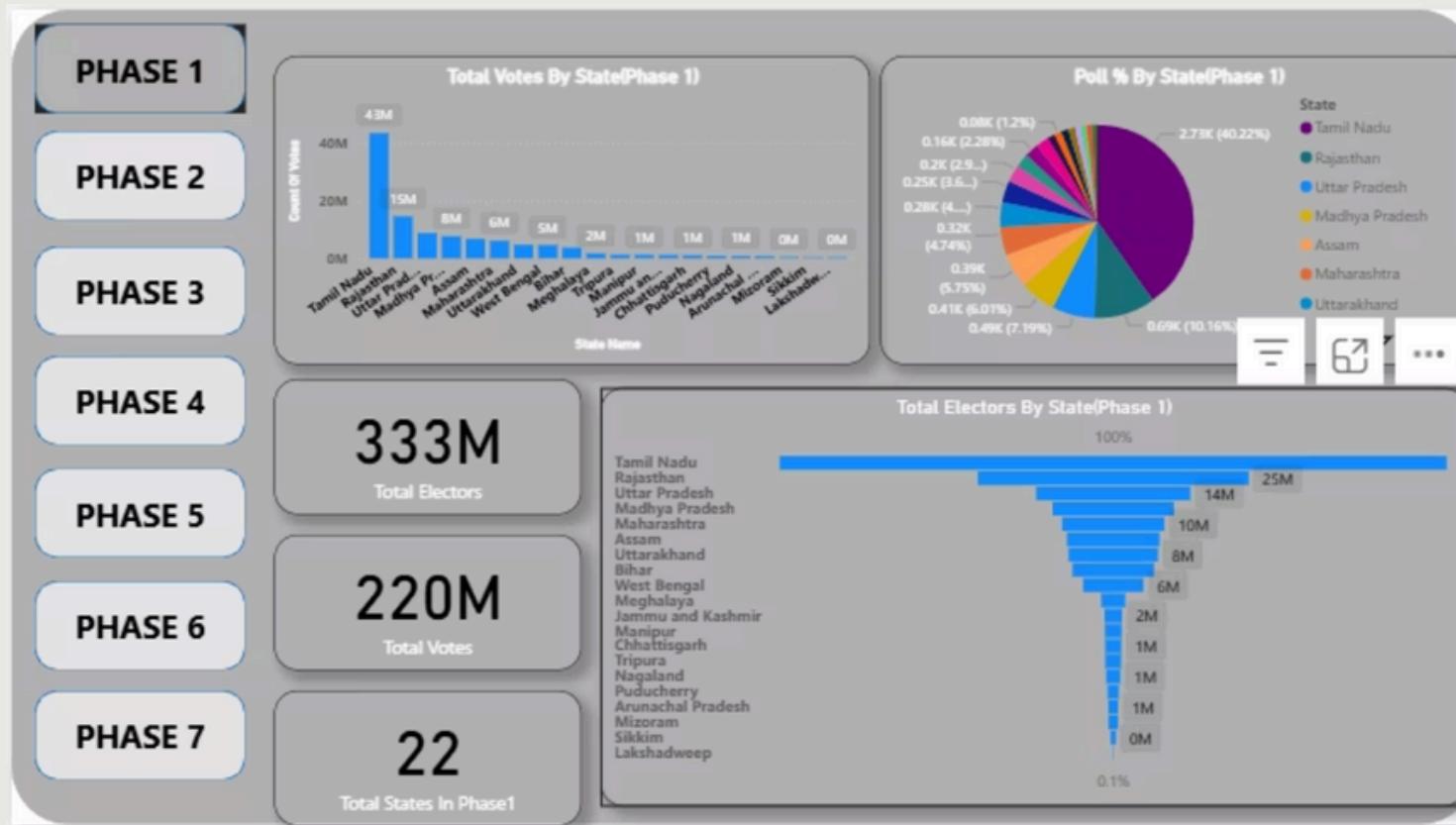
Sum of Margin and margin by state heat/category bars.

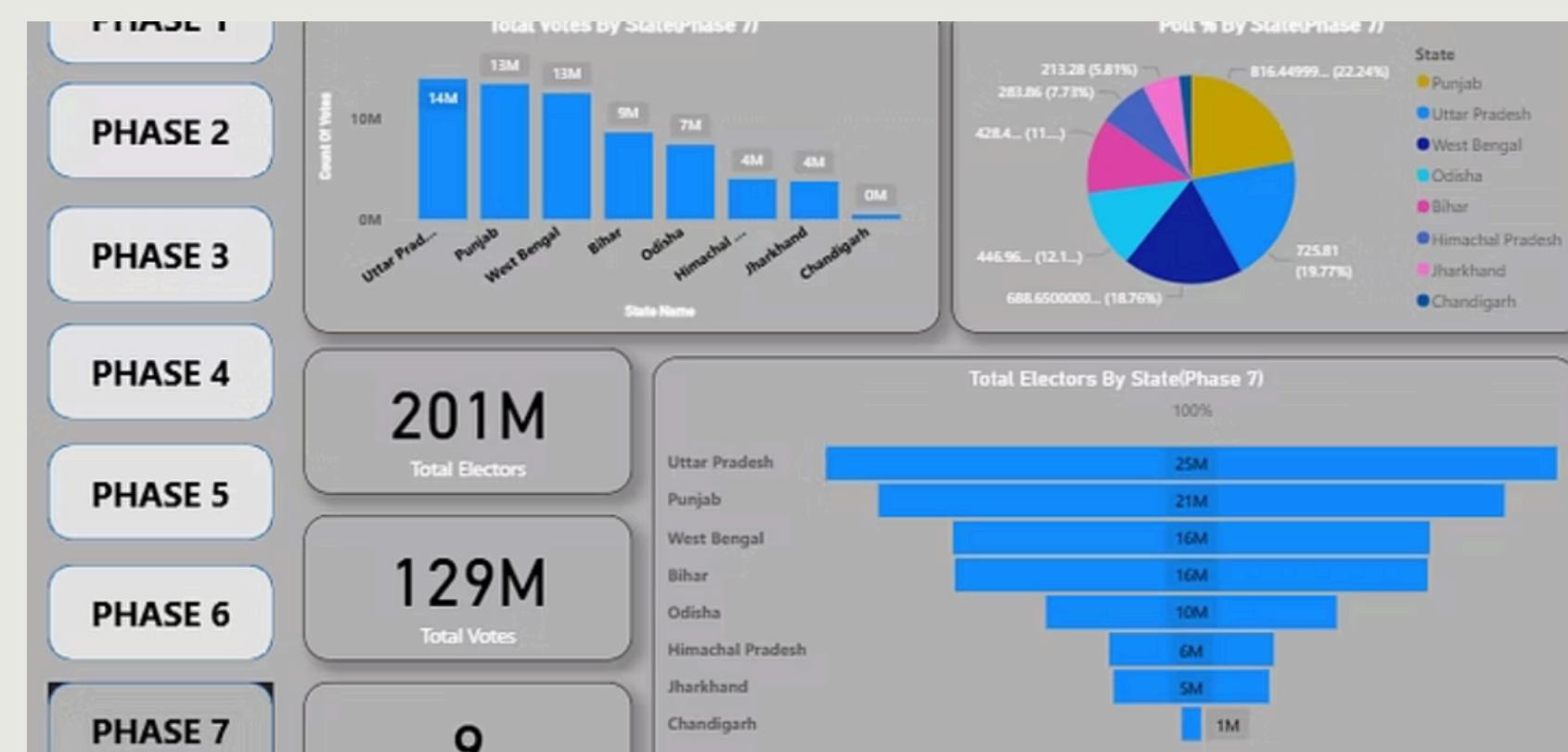
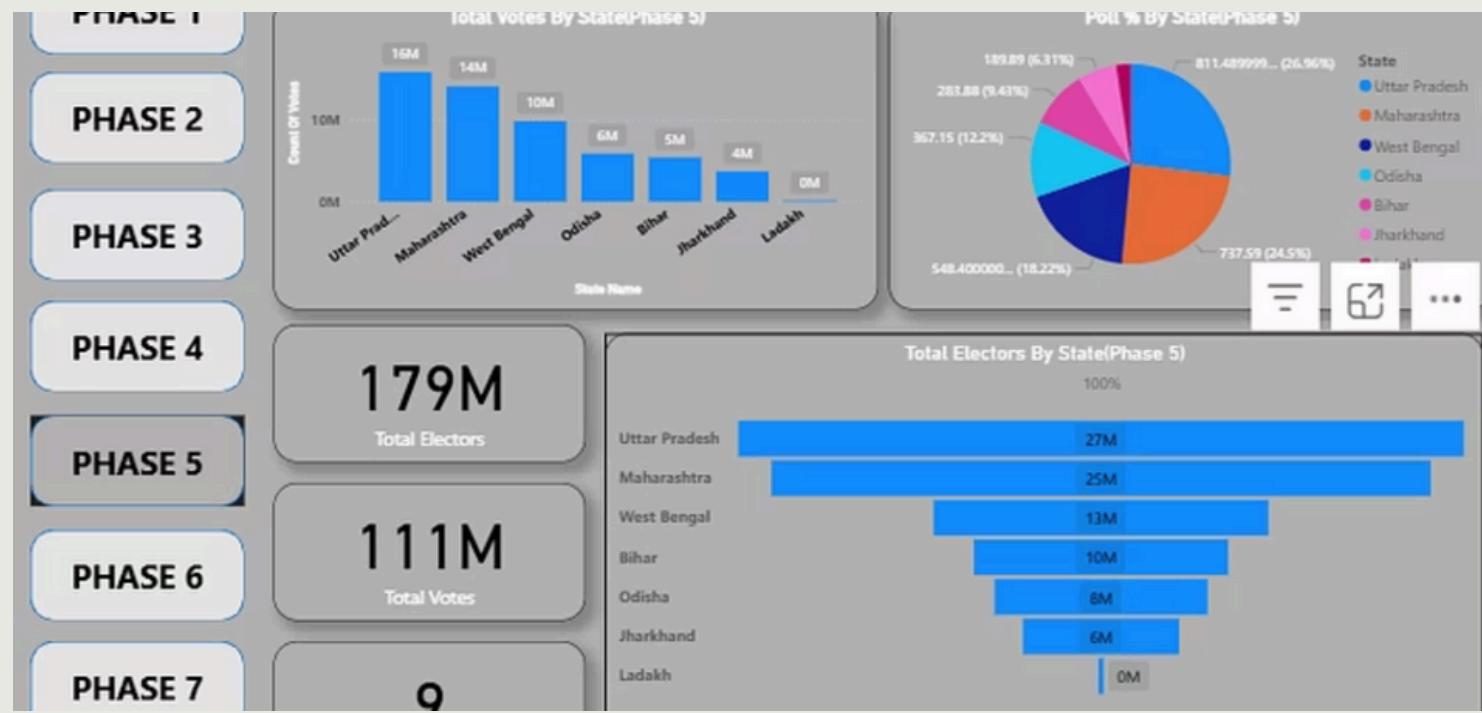


Phase 1–7 Snapshots

Phase cards show: Total Electors, Total Votes, States in phase, bar of Total Votes by State, funnel of Total Electors by State, and pie of Poll % by State.







Key Insights



Turnout Dynamics

phases with UP, TN, AP/Telangana lift national totals; turnout % varies meaningfully across phases tied to their state mixes.



State Contribution

a small set of large states (UP, MH, TN, AP/TG, WB) drive a majority of votes; dashboards quantify each state's phase-wise share.



Operational Signal

EVM polled vs counted gaps and postal vote patterns vary by state; margin distribution pinpoints competitive states for recount/dispute focus.



Party Landscape

national top-6 parties by votes evident, while treemap reveals regional dominance and long-tail parties contributing small but non-zero shares.



Project Overview

Tech Stack

Power BI Desktop,
Power Query (ETL),
DAX, star-schema
modeling; dataset from
Kaggle; visuals include
bar, funnel, treemap,
donut, pie, KPI cards.

Process

acquire → clean →
model → author
measures → build
national/phase
dashboards → validate
totals vs electors →
publish/report.

Deliverables

two dashboards
(national and
counting/margins),
seven phase pages, and
reusable DAX measures
for future election
cycles.

Conclusion and Next Steps

Outcome

a transparent, phase-aware election analytics suite enabling fast insights into turnout, state contributions, counting progress, and competitiveness.

Value

supports media coverage, academic analysis, and decision support with curated KPIs and drill-downs; modular model eases extension to new years/states.

Next Steps

add constituency-level maps, seat conversion and swing analysis, anomaly detection for EVM vs postal deltas, and automated data refresh pipelines.