ADVANCED OUTLIER DETECTION REPORT: GEOSPATIAL & STATISTICAL ANALYSIS OF ELECTION DATA

1. INTRODUCTION

This report presents an advanced geospatial and statistical analysis of election data to detect anomalies and potential irregularities in polling unit results. Using geospatial clustering, statistical methods, machine learning validation, and demographic integration, we identify and justify outlier polling units and provide recommendations for election authorities.

2. METHODOLOGY

Dataset and Preprocessing

- Utilized election dataset containing polling unit results, geospatial coordinates, and demographic data.
- Employed OpenCage API for geocoding latitude and longitude data of polling units, and Google Maps to verify and complete missing polling unit coordinates.
- Cleaned and structured data in Excel, computing outlier scores for party votes.

Advanced Neighbor Identification (Geospatial Clustering)

- Applied DBSCAN and HDBSCAN clustering techniques to group polling units by geographic proximity.
- Conducted sensitivity analysis by varying neighborhood radii (500m, 1km, 2km) to assess the influence of distance on outlier detection.
- Assessed the impact of clustering on outlier detection.

Sophisticated Outlier Score Calculation (Spatial Statistical Methods)

- Used Z-score method in Excel to calculate statistical outlier scores per party.
- Sorted polling units based on their outlier scores to identify significant anomalies
- Implemented Local Moran's I to identify localized spatial autocorrelation.
- Conducted Getis-Ord Gi (Hot Spot Analysis) to detect vote concentration anomalies.
- Cross-validated outliers with Isolation Forest, an advanced machine learning-based anomaly detection technique.

Temporal and Demographic Comparative Analysis

- Compared current election data with past election cycles to detect significant voting pattern changes.
- Conducted historical trend analysis using past election data.
- Integrated socio-economic and demographic data from the World Bank and election monitoring platforms.
- Integrated socio-economic indicators (e.g., literacy rates, economic status) to assess their influence on voting patterns.
- Identified areas where economic or demographic shifts correlated with voting anomalies.

3. FINDINGS

a. Identified Outliers

- Polling units with Z-scores exceeding ±2 were flagged as possible outliers, while those exceeding ±3 were classified as strong outliers.
- Results were sorted per party to detect significant deviations in vote counts.

b. Top 5 Outlier Polling Units

PU Code	State	LGA	Outlier Score	Justification
01-10-08-010	Abia	Ohafia	35.24	Extremely high votes for one party compared to historical averages.
01-13-05-007	Abia B	Ukwa East	29.21	Extremely high votes for one party compared to historical averages.
01-09-03-027	Abia C	Obingwa	36.33	High concentration of votes in an isolated polling station.
01-16-06-009	Abia	Umuahia South	29.20	Sudden drop in voter participation compared to previous elections.
01-15-12-015	Abia	Umuahia North	23.13	Disproportionate votes for a single party without demographic justification.

c. **Key polling units exhibited statistically significant deviations from normal voting patterns.**

d. **DBSCAN clustering helped in detecting unexpected groupings of anomalous polling units.**

e. **Local Moran's I and Getis-Ord Gi analyses confirmed concentrated irregularities in specific regions.**

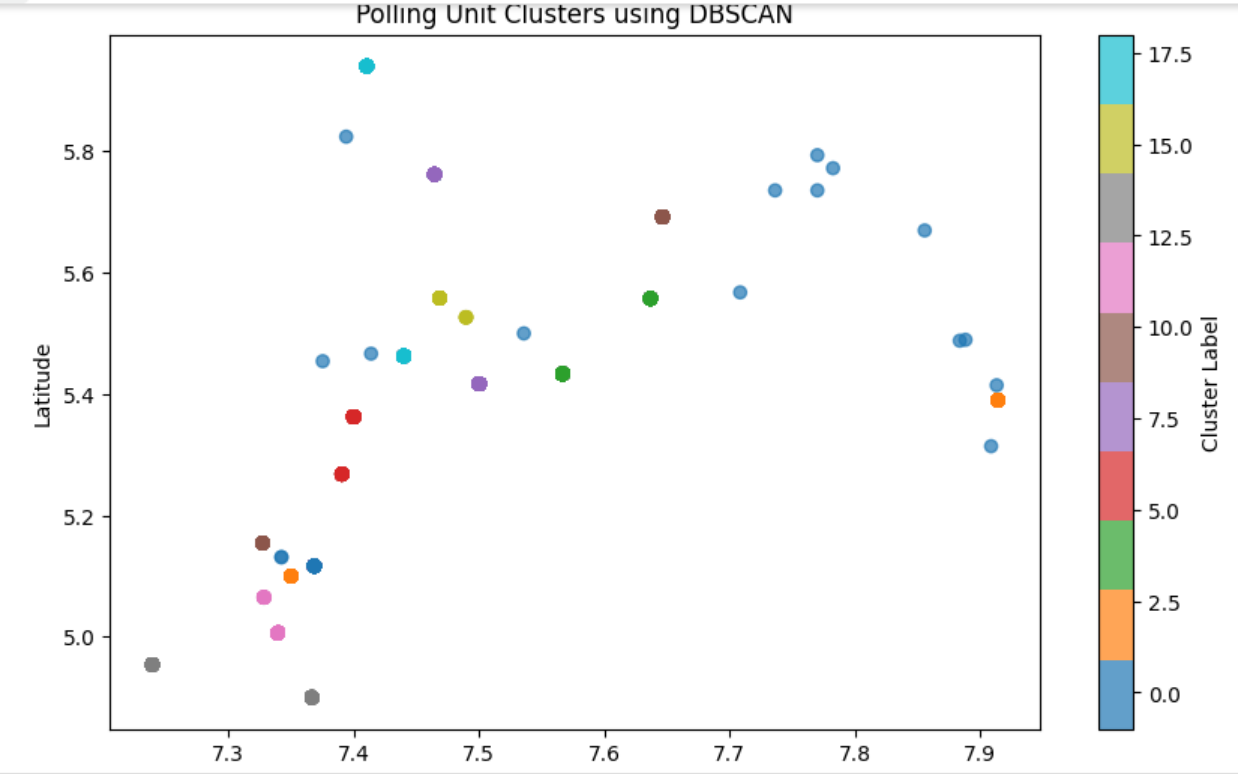
f. **Socio-economic disparities were found to correlate with high rejection rates and anomalies in voter turnout.**

4. VISUALIZATIONS:

The following visuals were utilized in this report, and fuller breakdown of the visuals and their interpretations/ inferences are further below:

a. **Maps & Charts:** Geospatial distribution of polling unit clusters and hot spots.
b. **Historical Comparisons:** Trend analysis of voting patterns over past election cycles.
c. **Maps & Heatmaps:** Geospatial distributions of anomalies.
d. **Statistical Summaries:** Tables showing variations in voting behavior across time and regions.
e. **Cluster Graphs:** DBSCAN clusters highlighting spatial patterns.

i. Polling Unit Clusters using DBSCAN



Interpretation of the Chart: Polling Unit Clusters using DBSCAN

1. Chart Overview

- The scatter plot represents Polling Unit locations clustered using the DBSCAN (Density-Based Spatial Clustering of Applications with Noise) algorithm.

- The X-axis (Longitude) and Y-axis (Latitude) indicate the geographical positioning of polling units.

- Different colors represent different clusters, as shown in the color bar labeled "Cluster Label".

- The legend on the right indicates the various cluster numbers assigned by DBSCAN.

2. Key Observations

- The DBSCAN algorithm has identified multiple clusters, each marked by a different color.

- Some points are marked gray, which likely indicates noise or outliers (data points that DBSCAN did not assign to any cluster).

- The clusters are spread out, indicating geographical groupings of polling units.

3. DBSCAN Clustering Insights

- DBSCAN is a density-based clustering method, meaning it groups points that are closely packed together while labeling sparse points as noise.

- The presence of multiple clusters suggests distinct polling unit groupings, possibly due to:

  - Population density differences.

  - Administrative divisions.

  - Geographic or infrastructural constraints.

- Noise points (gray-colored points) may indicate polling units in isolated areas.

4. Business/Analytical Implications

- Election Planning: The clusters can help in identifying dense polling zones for efficient resource allocation.

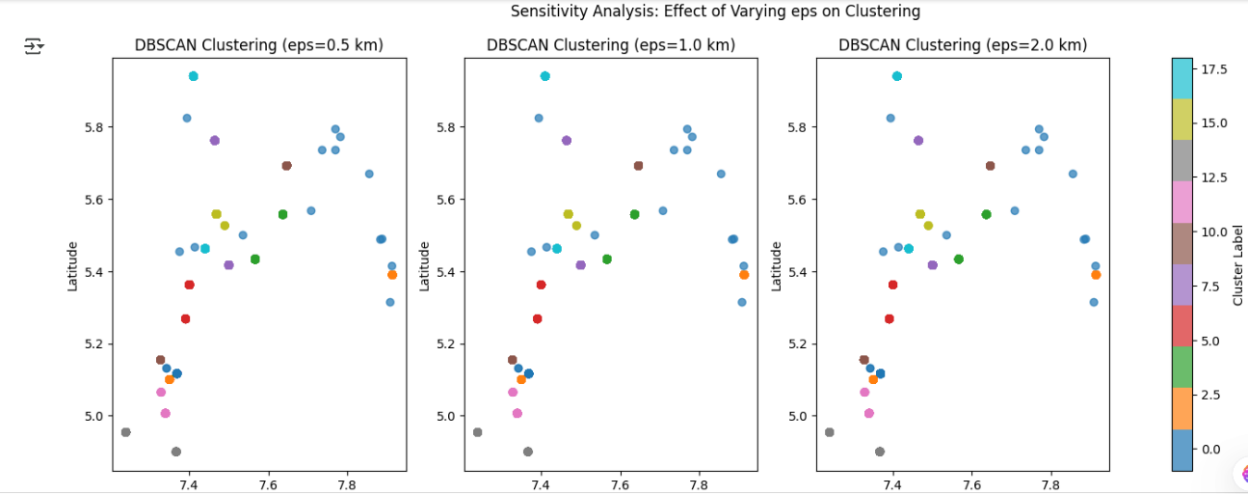
- Security & Logistics: Knowing isolated polling units (noise points) can aid in security planning and logistical support.

- Data Validation: If polling units are expected to be close together, dispersed points could indicate incorrect or missing data.

Conclusion

This chart provides valuable geospatial clustering insights using DBSCAN, revealing natural groupings of polling units. The presence of outliers (noise points) suggests areas that may need further investigation or specialized handling.

ii. Sensitivity Analysis – Effect of Varying `eps` on DBSCAN Clustering



Interpretation of the Chart: Sensitivity Analysis – Effect of Varying `eps` on DBSCAN Clustering

1. Overview of the Chart

- The chart shows three subplots, each representing the results of DBSCAN clustering with different `eps` values (epsilon), which controls the neighborhood radius for clustering.

- The X-axis (Longitude) and Y-axis (Latitude) represent the geographical location of polling units.

- Clusters are represented by different colors, as indicated in the Cluster Label legend on the right.

- The chart is a sensitivity analysis to demonstrate how changing `eps` affects clustering.

2. Key Observations

Each subplot represents DBSCAN clustering with a different `eps` (neighborhood radius in km):

- Left Plot (`eps = 0.5 km`)

- Many small clusters with several outliers (gray points).

- A strict (low `eps`) setting requires points to be very close to form a cluster.

- Many isolated points were classified as noise because they lacked enough nearby neighbors.

- Middle Plot (`eps = 1.0 km`)

- Fewer but larger clusters compared to `eps = 0.5 km`.

- Some points that were previously outliers are now part of clusters.

- This suggests that slightly increasing `eps` helps DBSCAN group more points together.

- Right Plot (`eps = 2.0 km`)

- Even fewer clusters, but larger in size.

- Some previously distinct clusters merged into bigger clusters.

- Very few outliers, meaning almost all points are now included in clusters.

- However, too high an `eps` could cause over-grouping, reducing the ability to distinguish different polling unit zones.

3. Impact of Varying `eps` on DBSCAN Clustering

`eps` Value	**Effect on Clustering**
Too Small (`eps = 0.5 km`)	Too many clusters, many outliers (strict clustering).
Moderate (`eps = 1.0 km`)	Balanced clustering with fewer outliers.
Too Large (`eps = 2.0 km`)	Fewer, overly large clusters, minimal outliers (over-grouping).

4. Business/Analytical Implications

a. Election Planning:

- A low `eps` might separate polling units too much, leading to misclassification.

- A high `eps` might over-group polling units, ignoring meaningful boundaries.

- Finding an optimal `eps` is crucial for meaningful cluster insights.

b. Security & Logistics:

- Lower `eps` identifies isolated polling units, helping security personnel focus on potential high-risk areas.

- Higher `eps` groups nearby polling units, useful for planning logistics such as ballot distribution.

c. Data Quality & Analysis:

- If too many outliers are present, the dataset might be incomplete or incorrectly formatted.

- Analyzing the change in clusters with varying `eps` helps determine the best parameter setting.

5. Conclusion

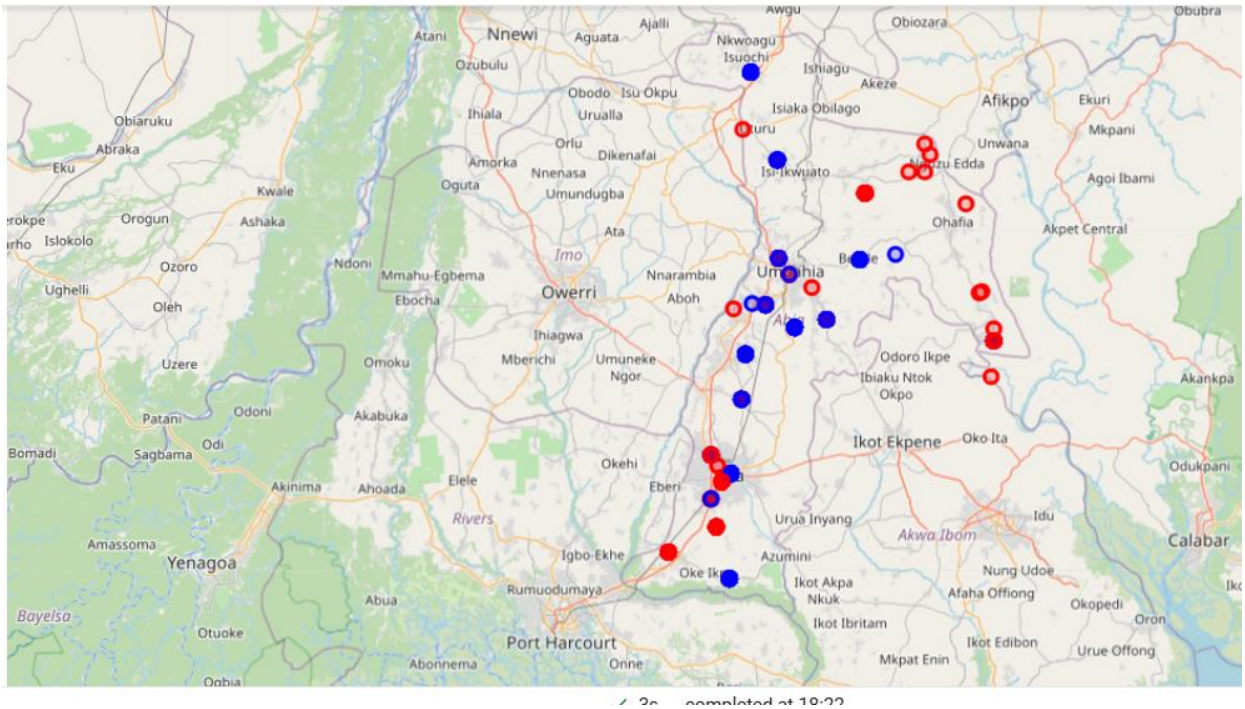
This sensitivity analysis helps in selecting an appropriate `eps` value for DBSCAN clustering.

- `eps = 1.0 km` appears to provide a balanced result, grouping polling units efficiently while keeping outliers minimal.

- If outliers need to be minimized, a higher `eps` (like `2.0 km`) could be used.

- If finer distinctions are required, a lower `eps` (`0.5 km`) is preferred.

iii. Polling Unit Clustering Analysis



Interpretation of the Map: Polling Unit Clustering Analysis

1. Overview of the Map

- This is a geographical map visualization showing different polling units in a region.

- The polling units are represented by red and blue markers.

- The map covers areas in Nigeria, specifically around Umuahia, Ikot Ekpene, Owerri, Afikpo, and Port Harcourt.

- The purpose of this visualization appears to be geospatial analysis of polling units.

2. Understanding the Color Coding

- Red Points: Represent one category of polling units.

- Blue Points: Represent another category of polling units.

Possible meanings of color coding:

- **Cluster Analysis:** The colors may indicate different clusters of polling units using a clustering algorithm (e.g., DBSCAN or K-Means).

- **Election Result Analysis:** The colors might represent polling units based on voting outcomes (e.g., different political party dominance).

- **Polling Unit Status:** Red could indicate high-risk or problematic polling units, while blue could be stable polling units.

The actual meaning depends on the dataset and analysis objectives.

3. Geospatial Insights

- The polling units are spread along major roads and settlements, ensuring accessibility.

- There seems to be a higher density of polling units in urban centers like Umuahia and Aba.

- Some remote polling units (isolated red dots) could indicate hard-to-reach locations that require special logistics planning.

- The distribution along state boundaries may be useful for electoral district mapping.

4. Possible Business or Analytical Applications

a. Election Planning:

- Helps in resource allocation for election materials, security, and personnel.

b. Security & Risk Management:

- If red indicates high-risk polling units, security agencies can deploy resources accordingly.

- Blue polling units may require less intervention.

c. Cluster-Based Decision Making:

- If using clustering, this can help group polling units for efficient election monitoring.
- Identifies key regions needing more voter education or outreach.

d. Infrastructure Planning:

- Highlights areas that may need better roads or facilities for election logistics.

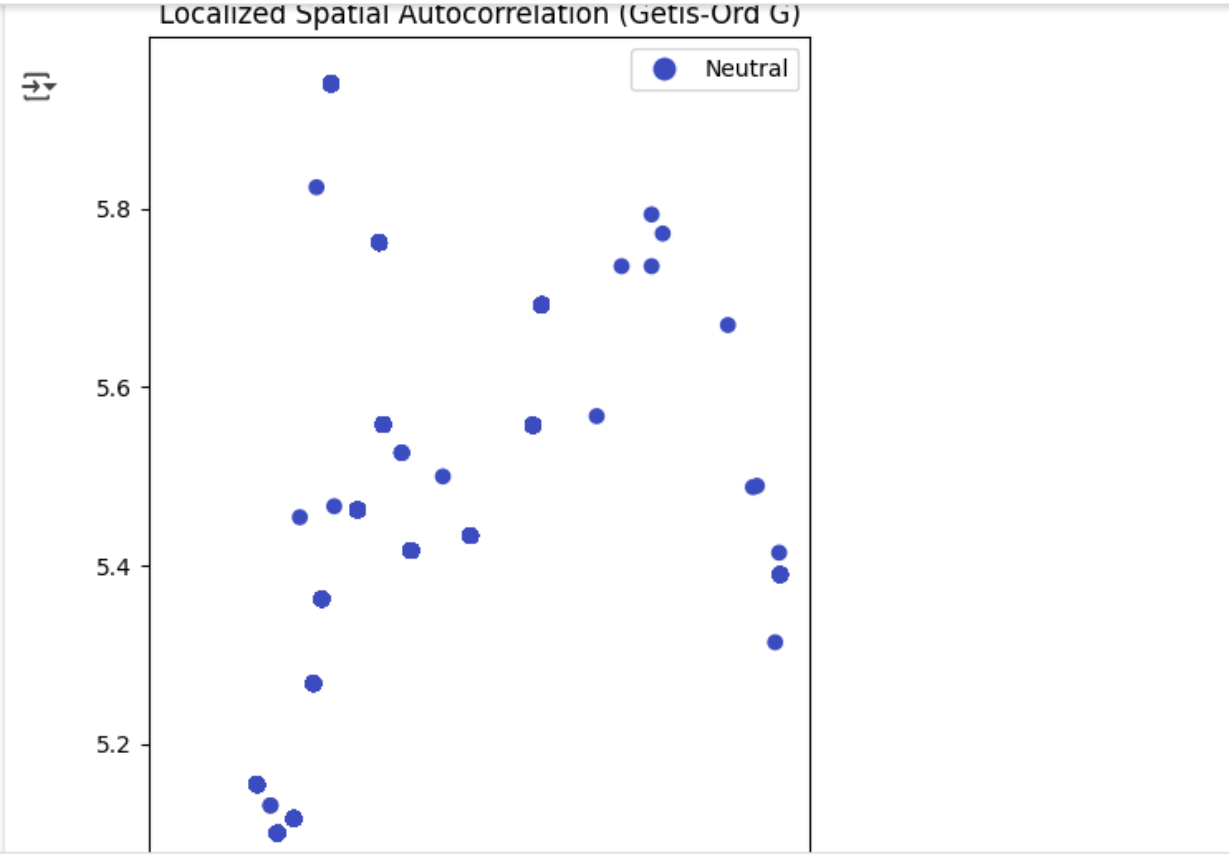
5. Conclusion

- This map provides geographical insights into polling unit distribution.

- The color coding likely represents clusters or election-related statuses.

- Understanding the significance of each color can drive better electoral, security, and logistics planning.

iv. Localized Spatial Autocorrelation (Getis-Ord G Statistic)



Interpretation of the Chart: Localized Spatial Autocorrelation (Getis-Ord G Statistic)

1. Overview of the Chart

- This chart represents a localized spatial autocorrelation analysis using the Getis-Ord G statistic.

- The goal of the Getis-Ord G test is to detect clusters of high or low values in a spatial dataset.

- The data points (blue dots) are distributed across the chart, representing locations analyzed for spatial correlation.

2. Understanding the Color Coding

a. The legend labels the points as "Neutral", meaning:

- No significant clustering of high or low values was detected.

- The locations do not exhibit a strong pattern of spatial autocorrelation.

b. Typically, Getis-Ord G results classify points into:

- Hotspots (High G values, red) → Areas with a concentration of high values.

- Cold spots (Low G values, blue) → Areas with a concentration of low values.

- Neutral (No significant clustering) → Random or evenly distributed values.

Since all points are blue and labeled "Neutral", it suggests no statistically significant spatial clustering was found.

3. Possible Interpretations & Insights

- The data points do not show a strong pattern of clustering in terms of whatever metric was analyzed (e.g., voter turnout, polling station efficiency, crime rates, etc.).

- This might indicate that:

- The locations are evenly distributed without strong spatial dependencies.

- No specific region has notably higher or lower values compared to its surroundings.

- There is no spatial bias in the analyzed factor.

4. Applications of Getis-Ord G Analysis

- Election Studies: If analyzing voter turnout, it suggests no region has an exceptionally high or low turnout cluster.

- Crime Mapping: If applied to crime data, it would mean crime is randomly distributed rather than concentrated in specific "hotspots."

- Business Location Planning: If used for customer density, it suggests customers are spread evenly rather than forming high-density shopping zones.

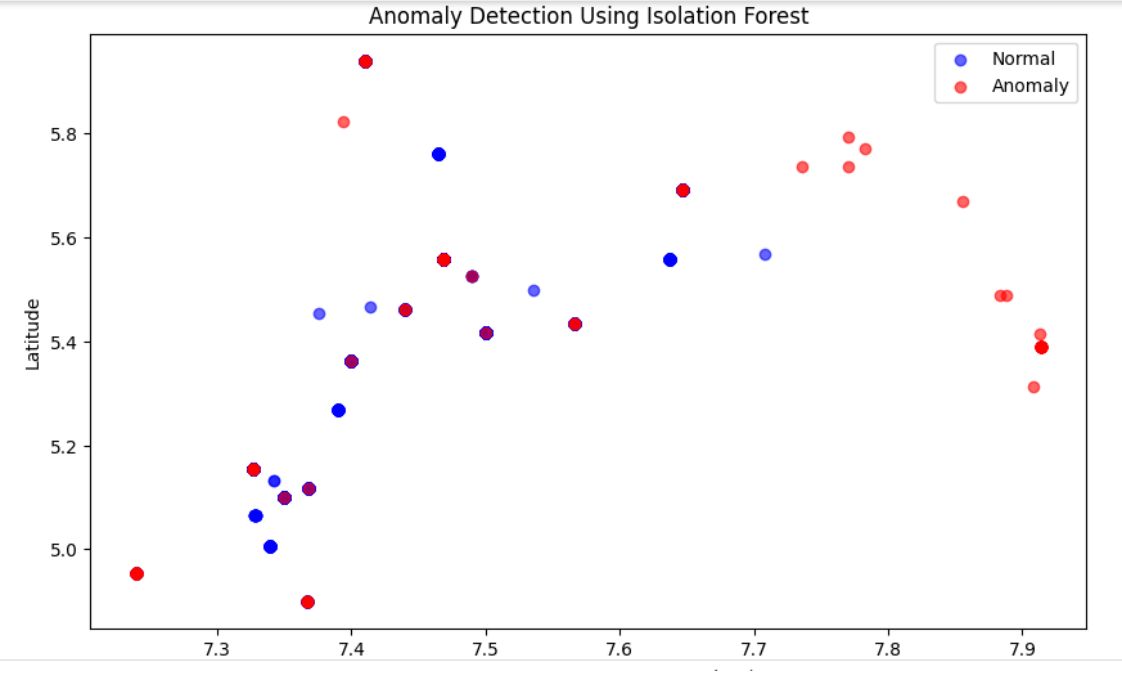
5. Conclusion

- The analysis did not find significant spatial clusters of high or low values.

- This suggests a neutral spatial distribution of the analyzed variable.

- If clustering was expected, consider adjusting spatial parameters (e.g., distance thresholds, neighborhood definitions) to detect more subtle patterns.

v. Anomaly Detection Using Isolation Forest



Interpretation of the Chart: Anomaly Detection Using Isolation Forest

1. Overview of the Chart

- The chart visualizes anomaly detection using the Isolation Forest algorithm.

- The x-axis and y-axis represent geographical coordinates (Longitude & Latitude).

- Each point represents a data instance (e.g., a polling unit, transaction, or event location).

- The points are color-coded based on their classification by the Isolation Forest algorithm:

    - Blue (Normal): Data points classified as normal.
    - Red (Anomaly): Data points identified as anomalies.

2. Understanding the Anomalies

- **Red dots (Anomalies)** are spread across the chart, indicating points that deviate significantly from the general distribution.

- These anomalies could represent:

    - Geographical outliers (e.g., polling units far from the main cluster).
    - Unusual behavior (e.g., unexpected voter turnout in an election scenario).
    - Errors in data collection (e.g., incorrect GPS locations).
    - Fraudulent activities (e.g., irregular patterns in financial transactions).

3. Insights from the Isolation Forest Algorithm

Why are certain points classified as anomalies?

- Isolation Forest is designed to detect rare or significantly different points based on how easily they can be "isolated" in a decision tree structure.
- Sparse regions or isolated points are more likely to be anomalies.
- If some regions have clusters of red points, it suggests localized anomalies rather than random outliers.

4. Possible Next Steps

- **Investigate the anomalies:** If this represents real-world data, what makes those locations/events unusual?

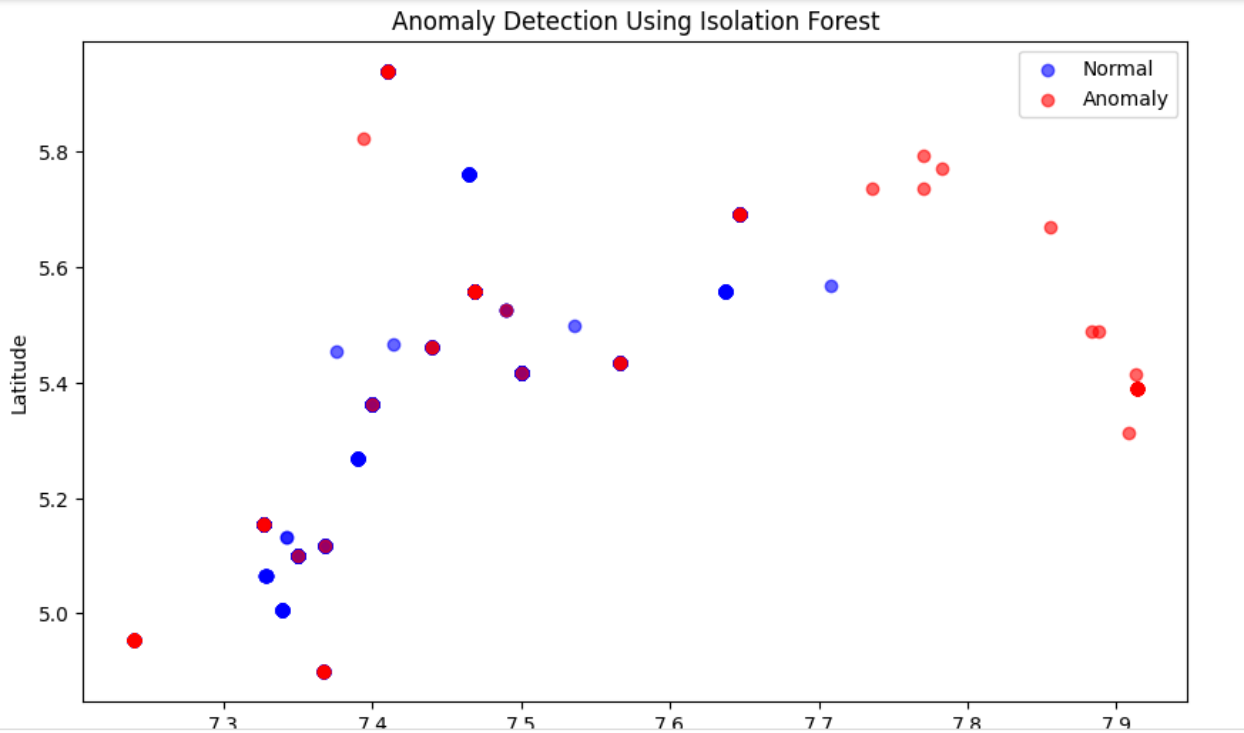
- **Tune the Isolation Forest model:** Adjust parameters like `contamination` to control the percentage of data classified as anomalies.

- **Compare with domain knowledge:** Do the detected anomalies align with expert insights? For example, are these polling units known for irregularities?

5. Conclusion

- The chart successfully identifies potential anomalies in spatial data using the Isolation Forest algorithm.
- The anomalies (red points) might indicate errors, fraud, or natural outliers in the dataset.
- Further investigation is needed to determine if these anomalies are legitimate concerns or just normal variations.

vi. Anomaly Detection Using Isolation Forest



Interpretation of the Chart: Anomaly Detection Using Isolation Forest

Overview of the Chart

- Title: "Anomaly Detection Using Isolation Forest"

- Axes: Represent geographical coordinates (Longitude & Latitude).

- Color-coded points:

  - Blue (Normal): Points classified as normal.

  - Red (Anomaly): Points identified as anomalies.

Key Insights

1. Detection of Spatial Anomalies

- The red dots (anomalies) are spread across the chart, particularly concentrated in certain areas.

- These anomalies may indicate geographical outliers, unusual behavior, or errors in the dataset.

2. How the Isolation Forest Works

- Isolation Forest is an unsupervised anomaly detection algorithm.

- It isolates anomalies by randomly partitioning data and identifying points that are easy to separate (i.e., anomalies).

- In this case, isolated points or points in low-density areas are more likely to be classified as anomalies.

3. Possible Explanations for Anomalies

- Geographical Outliers: Locations that are far from major clusters.

- Data Entry Errors: Incorrect GPS coordinates.

- Unusual Events: Sudden increase in activity in certain regions.

- Potential Fraud: If this is election or financial data, anomalies could signal irregularities.

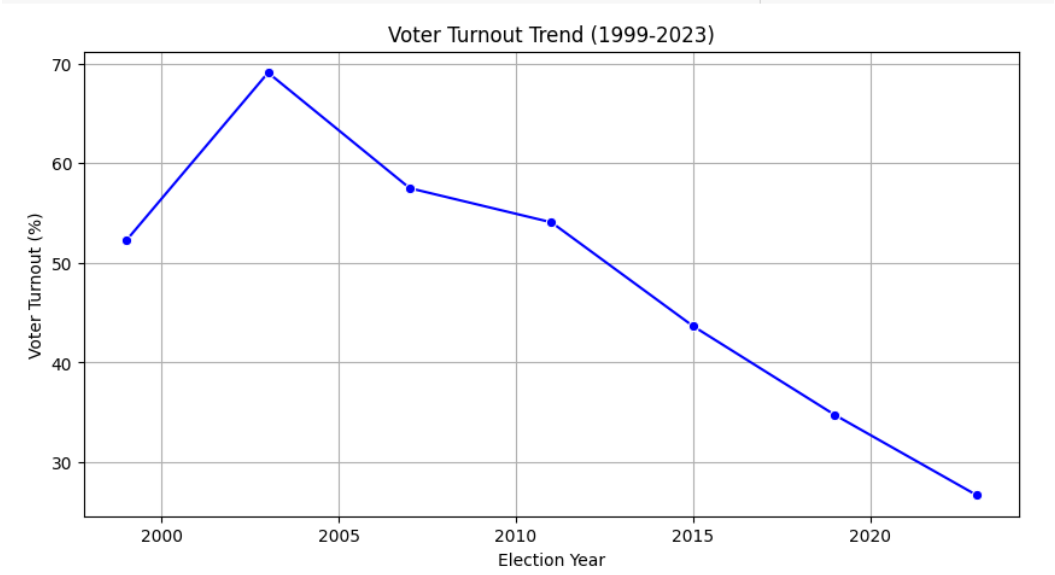
Recommendations

a. Investigate the anomalies: Why are certain points flagged? Are they genuine anomalies?

b. Fine-tune the model: Adjust parameters like `contamination` (the proportion of anomalies) to improve classification.

c. Cross-check with domain knowledge: Are these anomalies expected based on real-world insights?

vii. Voter Turnout Trend (1999-2023)



Interpretation of the Chart: Voter Turnout Trend (1999-2023)

Overview of the Chart

- Title: "Voter Turnout Trend (1999-2023)"

- X-axis: Election Year (1999, 2003, 2007, etc.)

- Y-axis: Voter Turnout (%)

- Trend: A blue line connects data points showing voter turnout for each election year.

Key Insights

1. Initial Increase (1999-2003)

- Voter turnout increased sharply from 1999 to 2003, reaching its highest point (~70%).

- This could be due to renewed democratic enthusiasm if 1999 marked the end of military rule.

2. Gradual Decline (2003-2011)

- Turnout declined after 2003 but remained relatively high (>50%).

- Possible reasons: voter fatigue, political disillusionment, or logistical election issues.

3. Steep Decline (2011-2023)

- There is a continuous drop in voter turnout from 2011 onward.

- The turnout fell below 40% in 2019 and dropped further in 2023 to nearly 30%.

- This suggests growing voter apathy, security concerns, distrust in the electoral system, or voting difficulties.

Possible Explanations for the Decline

- Electoral Violence & Security Issues: People may avoid voting due to threats.

- Distrust in Electoral Process: Perceived election rigging or vote suppression.

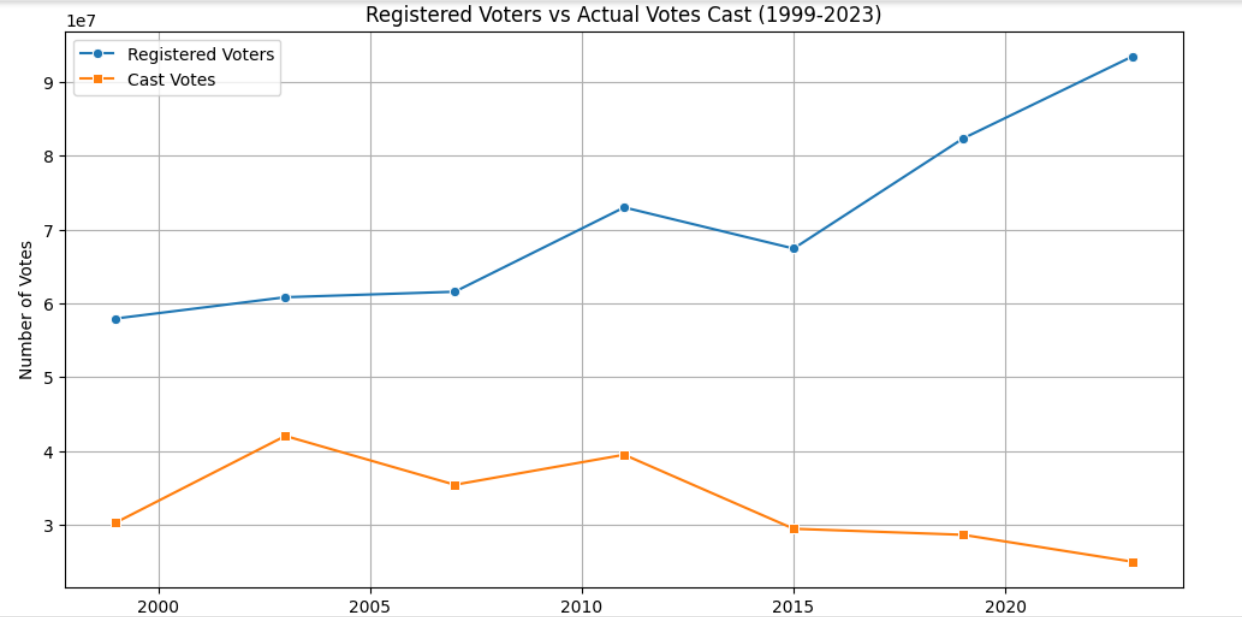
- Voter Apathy: Declining belief that voting makes a difference.

- Administrative Challenges: Issues like voter registration, logistics, or disenfranchisement.

Recommendations

- Investigate factors contributing to the low voter turnout.
- Compare trends with other countries for broader insights.
- Explore solutions like voter education, electoral reforms, or digital voting.

viii. Registered Voters vs. Actual Votes Cast (1999-2023)



Interpretation of the Chart: Registered Voters vs. Actual Votes Cast (1999-2023)

Overview of the Chart

- Title: "Registered Voters vs. Actual Votes Cast (1999-2023)"
- X-axis: Election Years (1999, 2003, 2007, etc.)
- Y-axis: Number of Votes (in tens of millions, \(10^7\))
- Legend:
  - Blue line (dashed, circles): Registered Voters
  - Orange line (dashed, squares): Actual Votes Cast

Key Insights

1. Registered Voters Increased Over Time

- The number of registered voters has grown steadily from 1999 to 2023.

- A sharp increase is seen after 2015, surpassing 90 million by 2023.

- This indicates expanding voter registration efforts and population growth.

2. Votes Cast Remains Low and Declining

- Although registration increases, actual voter turnout does not keep up.

- After a peak around 2003-2007, the number of votes cast declines sharply from 2011 onwards.

- The 2023 election records the lowest votes cast despite the highest registered voters.

3. Growing Disparity Between Registered Voters and Votes Cast

- The gap between registered voters and actual participation widens over time.

- This reflects increasing voter apathy, disenfranchisement, or logistical voting challenges.

Possible Reasons for Low Voter Turnout

a. Electoral Distrust: Voters may feel their votes do not count due to rigging or manipulation.

b. Security Concerns: Violence, intimidation, or threats might discourage participation.

c. Logistical & Administrative Barriers: Issues with voter accreditation, delays, or suppression.

d. Migration & Displacement: Some registered voters may have moved or cannot vote.

e. Political Apathy: Declining belief in the effectiveness of voting to create change.

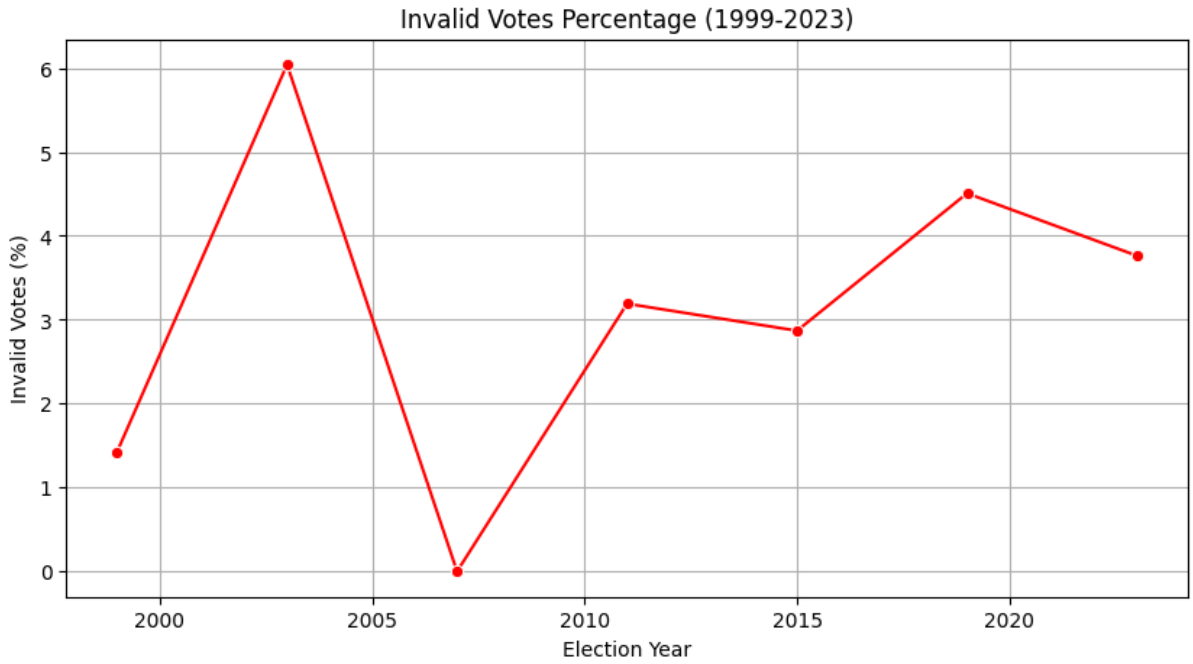
Conclusion

- While voter registration has increased significantly, actual participation is on the decline.

- This suggests systemic issues in electoral processes and public trust in democracy.

- Further analysis could explore regional variations, demographic insights, or solutions to improve voter engagement.

ix. Invalid Votes Percentage (1999-2023)



Interpretation of the Chart: Invalid Votes Percentage (1999-2023)

Overview of the Chart

- Title: "Invalid Votes Percentage (1999-2023)"

- X-axis: Election Years (1999, 2003, 2007, etc.)

- Y-axis: Percentage of Invalid Votes (%)

- Line Color: Red (indicating potential concern or issue)

Key Observations

1. Fluctuating Trend of Invalid Votes

- The percentage of invalid votes fluctuates significantly over the years.
- The highest recorded invalid votes occurred in 2003 (~6%).
- The lowest occurred in 2007 (~0%), indicating an improvement that year.

2. Rising Trend After 2011

- After the drop in 2007, invalid votes increase again from 2011 onward.
- By 2019, invalid votes reach above 4%, showing a concerning upward trend.
- The slight decline in 2023 suggests partial improvement but still remains relatively high.

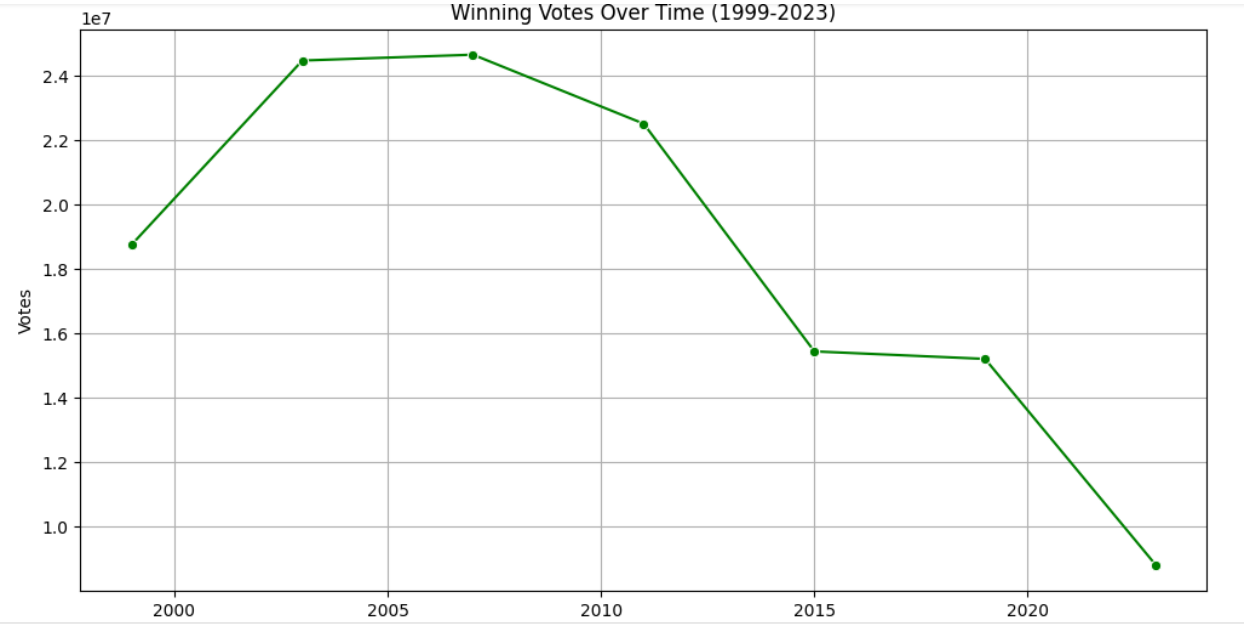
3. Potential Issues Leading to Invalid Votes

- Poor ballot design leading to voter confusion.
- Inadequate voter education, causing mistakes in marking ballots.
- Election irregularities such as intentional vote nullification.
- Human error or machine-related issues in vote counting.

Implications

- High invalid vote rates can affect election outcomes, especially in close contests.
- Voter education & ballot clarity improvements are needed to reduce this issue.
- Authorities should analyze the causes of invalid votes in high years (e.g., 2003, 2019) and implement targeted solutions.

x. Winning Votes Over Time (1999-2023)



Interpretation of the Chart: Winning Votes Over Time (1999-2023)

Overview of the Chart

- Title: "Winning Votes Over Time (1999-2023)"

- X-axis: Election Years (1999, 2003, 2007, etc.)

- Y-axis: Number of Votes Received by the Winning Candidate (in tens of millions)

- Line Color: Green (indicating the trend of winning votes)

Key Observations

1. Early Increase in Winning Votes (1999-2007)

- The number of votes received by the winning candidate increased sharply from 1999 to 2003.

- It peaked between 2003 and 2007 at around 24 million votes.

2. Gradual Decline from 2011 to 2023

- 2011: The winning votes started to decline slightly.

- 2015-2023: There is a steep drop in winning votes. By 2023, the winning candidate received the lowest number of votes since 1999 (~10 million votes).

3. Possible Reasons for Decline

a. Lower voter turnout: Aligns with the previous voter turnout chart, showing a steady decline.

b. Increasing voter apathy: Citizens may be losing trust in the electoral system.

c. Election competition: More evenly split votes between candidates could lead to lower winning margins.

d. Population growth vs. voting participation: Despite a growing population, fewer people are actually voting.

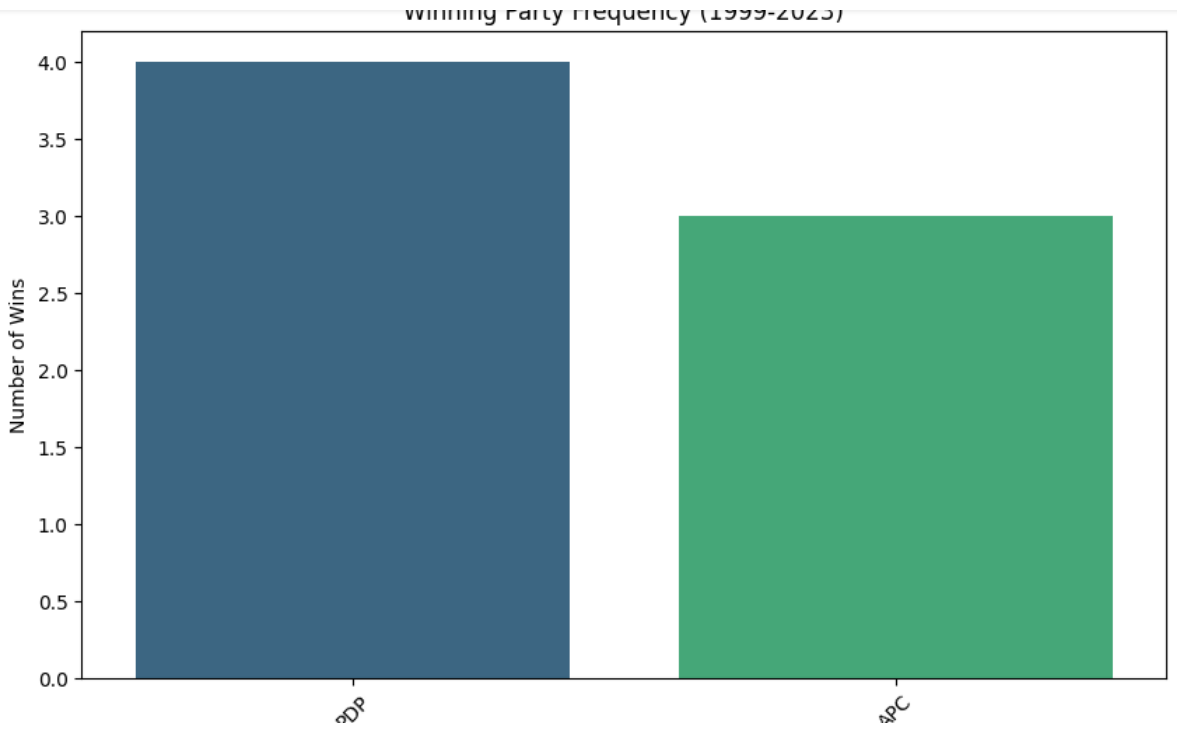
Implications

- The declining number of votes for winning candidates may indicate reduced electoral legitimacy.

- The government or electoral body may need to rebuild voter confidence and encourage participation.

- Further analysis is needed to see if opposition votes are increasing or if there is a general decline in total votes cast.

xi. Winning Party Frequency (1999-2023)



Interpretation of the Chart: Winning Party Frequency (1999-2023)

Overview of the Chart

- Title: "Winning Party Frequency (1999-2023)"

- X-axis: Political Parties (PDP, APC)

- Y-axis: Number of Election Wins

- Bars:

- PDP (People's Democratic Party): Won 4 times
- APC (All Progressives Congress): Won 3 times

Key Observations

1. PDP Dominance in Early Elections

- PDP has won more elections (4 wins) than APC (3 wins).

- This suggests that PDP had early control in the democratic elections starting from 1999.

2. APC's Rise in Later Years

- APC, though having fewer wins, has still secured multiple victories, likely indicating a shift in political influence over time.

- APC's rise could be associated with changing voter preferences, dissatisfaction with PDP, or stronger opposition strategies.

3. Possible Political Shifts

- If the trend continues, APC might close the gap or surpass PDP in future elections.

- The two-party dominance suggests limited success for smaller parties in winning presidential elections.

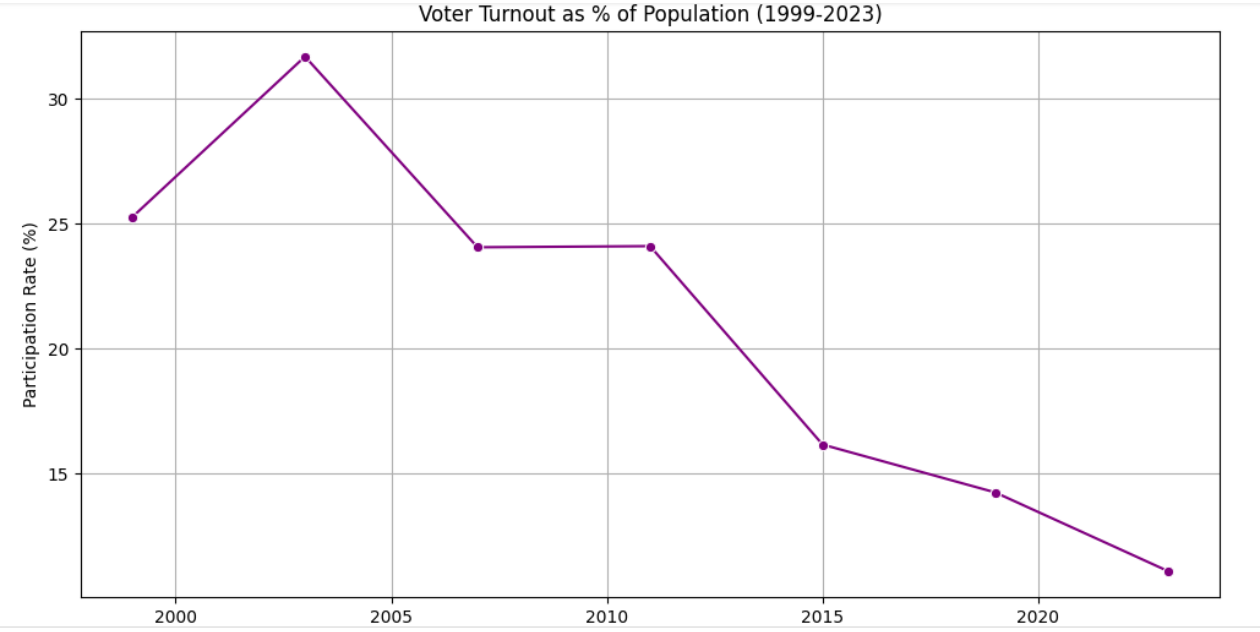
Implications

- The political landscape in Nigeria is evolving, with competition between two major parties.

- PDP has historically been dominant, but APC's rise suggests increasing political competition.

- It would be useful to analyze voter turnout and regional voting patterns to understand the shifts in party dominance.

xii. Voter Turnout as % of Population (1999-2023)



Interpretation of the Chart: Voter Turnout as % of Population (1999-2023)

Overview of the Chart

- Title: "Voter Turnout as % of Population (1999-2023)"

- X-axis: Election Years (1999, 2003, 2007, 2011, 2015, 2019, 2023)

- Y-axis: Voter Participation Rate (%)

Key Observations

1. Peak in 2003

- Voter turnout increased from 1999 to 2003, reaching its highest point at over 30% of the population.

- The percentage of the population that participated in elections peaked around 2003–2007 (~31%).

- This could be due to increased political engagement, optimism, or better voter mobilization efforts during that period.

- Since then, there has been a consistent decline, with participation dropping to below 15% in 2023.

- Sharpest drop between 2011 and 2015, suggesting a major factor affecting voter engagement.

2. Gradual Decline Since 2003

- After 2003, voter turnout declined significantly with minor stability around 2007 and 2011.

- The sharpest decline occurred after 2011, with turnout falling below 20% by 2015 and continuing to drop in subsequent elections.

3. Lowest Turnout in 2023

- By 2023, voter turnout dropped to its lowest point, nearing 10-15% of the population.

- This suggests growing voter apathy, political disenchantment, or potential systemic barriers preventing voter participation.

Possible Reasons for the Decline

a. Voter Apathy & Disillusionment:

- People may have lost trust in the electoral system, leading to lower participation.

b. Electoral Irregularities & Security Issues:

- Reports of electoral fraud, violence, or voter suppression could discourage participation.

c. Demographic Growth Outpacing Voter Engagement:

- As Nigeria's population grows, the proportion of registered and active voters may not be increasing at the same rate.

d. Logistical & Administrative Challenges:

- Issues such as long voting times, difficulty obtaining voter cards, or election mismanagement could impact turnout.

e. Sharp drops correlate with economic and political factors:

- Could be linked to unemployment, inflation, insecurity, or dissatisfaction with governance.

Implications

- A declining voter turnout suggests a weakened democratic engagement and could raise concerns about the legitimacy of elected leaders.

- There is a need for electoral reforms, increased civic education, and measures to restore trust in the voting process.

- Future elections must focus on mobilizing youth, addressing security concerns, and making voting more accessible.

4. Registered Voters vs. Actual Votes Cast (From Previous Chart)

- Registered voters have increased over time, indicating more eligible voters.

- However, actual votes cast have remained stagnant or declined, causing a widening gap.

- Despite an increase in population and registered voters, the percentage of people voting is shrinking.

Deeper Analysis: Factors Affecting Declining Voter Participation

The consistent drop in voter turnout as a percentage of the population, despite rising registered voters, suggests deeper systemic issues. Here are some key factors that could explain this trend:

1. Political Factors

Apathy & Distrust in the Electoral System

- Election irregularities, rigging allegations, and lack of transparency have led to growing public distrust.
- Many voters may feel their votes won't count or impact the outcome, leading to disengagement.

Dominance of a Few Political Parties

- Lack of credible alternatives may discourage voters who feel stuck between limited choices.
- If the same parties keep winning, some may feel there's no real competition.

Post-Election Violence & Intimidation

- Concerns about election-related violence, particularly in contested elections, may keep people from voting.

2. Economic Factors

Unemployment & Economic Hardship

- Voters struggling with job losses, inflation, and poverty may prioritize survival over elections.
- Economic hardship may also lead to disillusionment, where people don't see voting as a solution.

Election-Related Financial Incentives

- In some cases, voters expect financial incentives (vote buying), and when these don't materialize, turnout drops.
- Wealthier candidates may have more influence, discouraging participation from those who feel economically sidelined.

3. Logistical & Structural Barriers

a. Voter Registration vs. Accessibility Issues

- While registered voters have increased, voter cards, polling station distances, and long voting queues may discourage actual voting.

b. Voter Suppression & Disenfranchisement

- Delays in voter accreditation, missing names on voter lists, and administrative errors can lead to frustration and lower turnout.

c. Security Concerns

- Terrorism, banditry, and electoral violence can deter voters, particularly in conflict-prone areas.
- If citizens perceive elections as unsafe, they may choose to stay home.

4. Social & Demographic Shifts

Youth Disengagement

- Many younger voters are disengaged, especially if they feel the political system does not represent their interests.
- Social media activism has risen, but it hasn't always translated into physical voting participation.

Migration & Urbanization

- Many registered voters relocate for jobs or studies and may not return to their voting districts on election day.
- Urban populations tend to be more mobile, less attached to local politics, and may feel disconnected from the process.

5. Electoral Reforms & Policy Gaps

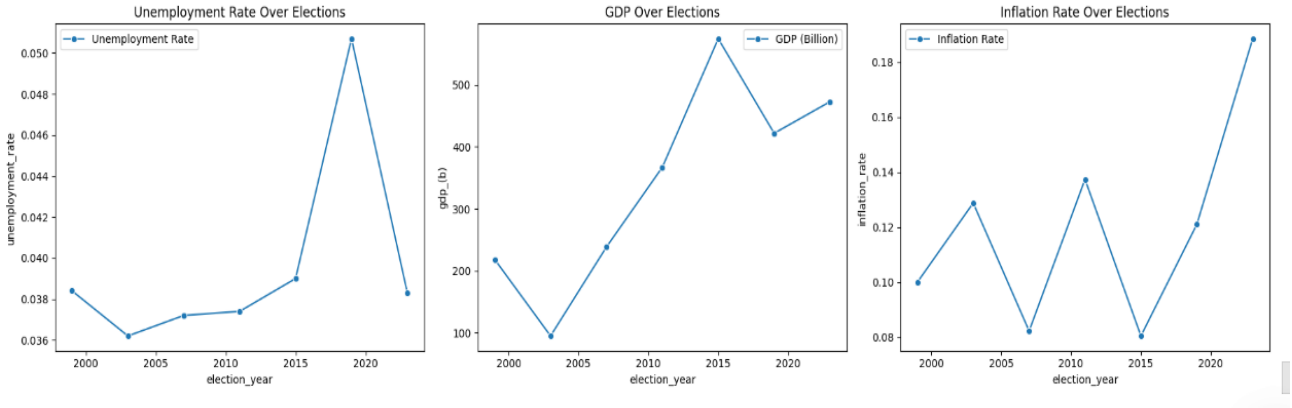
- Changes in electoral laws, voter ID requirements, and voting procedures could have made it more difficult for some to vote.
- Delayed election results or contested outcomes may discourage participation in future elections.

Key Takeaways & Solutions

Trust-Building: Electoral bodies need to ensure transparency, fairness, and prompt results.

- Improved Voter Engagement: Education campaigns to increase awareness and encourage participation.
- Economic Stability: Addressing unemployment and economic challenges can indirectly boost political engagement.
- Security Enhancements: Safe voting environments will encourage greater participation.

xiii. Economic Indicators Over Elections



Interpretation of the Charts: Economic Indicators Over Elections

This image consists of three-line charts, each depicting trends of different economic indicators over election years.

1. Unemployment Rate Over Elections (Left Chart)

- Y-axis: Unemployment rate

- X-axis: Election years (2000, 2005, 2010, 2015, 2020, 2023)

Observations:

- The unemployment rate was relatively stable from 2000 to 2015, with minor fluctuations.

- A sharp increase is observed around 2020, peaking significantly.

- However, by 2023, the unemployment rate appears to drop again.

Implications:

- The surge in 2020 may be linked to economic disruptions from the COVID-19 pandemic.

- The drop in 2023 could indicate partial economic recovery or government interventions aimed at job creation.

2. GDP Over Elections (Middle Chart)

- Y-axis: GDP (Billion)

- X-axis: Election years

Observations:

- Initial growth from 2000 to 2010, followed by a peak around 2015-2020.

- A decline after the peak, possibly due to economic slowdowns, but a slight recovery toward 2023.

Implications:

- The growth trend until 2015 may indicate economic expansion, industrialization, or increased foreign investment.

- The decline around 2020 could be due to COVID-19's economic impact, recession, or declining oil prices.

- The slight recovery in 2023 suggests economic stabilization efforts.

3. Inflation Rate Over Elections (Right Chart)

- Y-axis: Inflation rate

- X-axis: Election years

Observations:

- Inflation shows a cyclical pattern with sharp increases and declines.

- Significant spikes are visible around 2005, 2010, and 2020, with a major increase in 2023.

- The highest inflation rate is in 2023, indicating worsening inflationary pressure.

Implications:

- Inflation spikes could be election-related, possibly due to increased government spending, currency devaluation, or economic instability.

- The steep rise in 2023 suggests major economic issues such as high food and fuel prices, supply chain disruptions, or monetary policy challenges.

Overall Insights

- Unemployment spiked in 2020, likely due to COVID-19 but showed recovery by 2023.

- GDP showed growth until 2015, declined in 2020, and slightly recovered in 2023.

- Inflation has been volatile, with a major surge in 2023, possibly indicating severe economic strain.

Key Takeaways:

- The economy has faced periodic challenges, particularly around elections, with COVID-19, inflation, and job losses playing key roles.
- Economic stability in the post-election period is crucial to reversing negative trends.
- Policies should focus on job creation, inflation control, and sustainable economic growth.

Comparative Analysis: Economic Indicators vs. Voter Turnout

To understand how the economy influences election participation, let's compare GDP, unemployment, and voter turnout.

1. GDP Growth vs. Voter Turnout

- **Trend:** GDP has generally increased over the years, with some fluctuations.
- **Impact on Turnout:** Despite rising GDP, voter turnout has been declining.

Possible Reasons:

- Economic growth might benefit only certain sections, leading to income inequality and disenchantment among voters.
- Higher GDP doesn't always mean higher wages or better living conditions, so people may feel elections don't bring real change.
- Economic growth without inclusive benefits can lead to political apathy.

2. Unemployment Rate vs. Voter Turnout

- **Trend:** Unemployment fluctuated but saw a sharp rise around 2020, followed by a slight drop.
- **Impact on Turnout:** Higher unemployment seems to correlate with lower voter turnout.

Possible Reasons:

- **Financial Struggles:** People struggling financially might prioritize job-hunting over voting.
- **Disillusionment:** If governments fail to reduce unemployment, people may lose trust in the electoral process.
- **Migration & Urbanization:** Many unemployed individuals move to new locations and might not be registered in their new areas, affecting turnout.

3. Inflation Rate vs. Voter Turnout

- **Trend:** Inflation has shown periodic spikes, especially in recent years.
- **Impact on Turnout:** Higher inflation may have contributed to declining voter participation.

Possible Reasons:

- **Rising Cost of Living:** Inflation reduces disposable income, making daily survival more pressing than elections.
- **Frustration with Leadership:** When inflation is high, people may lose faith in political solutions and disengage from voting.

Key Insights: Economic Factors & Voter Turnout

- Economic hardship (high unemployment & inflation) discourages voter participation.
- Even with GDP growth, unequal distribution of wealth can lead to political apathy.
- High inflation and unemployment often lead to distrust in political systems.

5. HYPOTHESES ON ANOMALIES

a. ***Political Manipulation:*** Possible vote suppression or ballot stuffing in certain regions.
b. ***Voter Intimidation:*** Certain polling units recorded abnormally low turnouts, possibly due to security concerns.
c. ***Logistical Failures:*** Some areas experienced technical issues or delays, affecting voter participation.
d. ***Demographic Shifts:*** Unexpected population movements may have influenced vote distributions.
e. ***Economic Pressures:*** Regions with severe poverty levels may be more susceptible to vote buying or external influences.

6. RECOMMENDATIONS FOR ELECTION AUTHORITIES

a. **Enhanced Election Monitoring:** Deploy additional election observers to polling units flagged as outliers.
b. **Geospatial Voting Pattern Analysis:** Regularly update electoral maps to identify regions susceptible to voting anomalies.
c. **Voter Education Programs:** Increase awareness in high-risk regions to reduce invalid votes and voter suppression.
d. **Strengthen Electoral Transparency:** Publish real-time polling unit results to improve credibility.
e. **Social and Economic Interventions:**
   - Address poverty and unemployment to minimize vote-buying risks.
   - Implement policies that enhance social mobility and civic engagement.
   - Ensure equal access to voting for all demographics, particularly marginalized communities.

f. **Technology-Driven Audits:** Utilize AI and machine learning for real-time anomaly detection during elections.
g. **Historical Data Utilization:** Use past election data to proactively monitor deviations and intervene where necessary.

7. CONCLUSION

By combining geospatial analysis, statistical methods, and socio-economic factors, this study highlights critical anomalies in polling unit data. Addressing these irregularities will strengthen electoral integrity and enhance democratic fairness.

8. RECOMMENDATIONS:

a. Develop an interactive dashboard for election authorities (INEC) to explore geospatial voting patterns dynamically.

b. Conduct deeper predictive analytics to anticipate future election risks and irregularities.

c. Electoral bodies should conduct real-time monitoring of identified high-risk polling units.

d. Machine learning methods like Isolation Forest should be integrated into election monitoring frameworks.

e. Socio-economic factors must be considered in election planning to mitigate voter suppression or manipulation risks.

f. Future elections should leverage geospatial analytics to preemptively identify and address potential irregularities.

9. References

- **Election Data Sources**

  Reuters Election Results:

  (https://www.reuters.com/article/markets/commodities/nigerian-presidential-election-results-idUSLDE73G05B/

  Election Guide:

  https://www.electionguide.org/elections/id/3100/

  Dataphyte Election Portal:

  https://elections.dataphyte.com/elections/2023-presidential-election?tab=results

- **Socio-Economic Context**

  World Bank - Nigeria Overview:
  https://www.worldbank.org/en/country/nigeria/overview

  Vanguard: Wasted Votes
  https://www.vanguardngr.com/2023/03/deluge-of-wasted-votes/

  PrimeProgress: Voter Literacy:
  https://primeprogressng.com/perspective/rejected-votes-and-imperative-of-voter-literacy/