**Report on Outlier Detection in Election Data Using Geospatial Analysis**

**Introduction**

In the recently concluded election, the Independent National Electoral Commission (INEC) faced numerous legal challenges regarding the integrity and accuracy of the election results. Allegations of vote manipulation and irregularities prompted a thorough investigation. This report aims to uncover potential voting irregularities by identifying outlier polling units where the voting results deviate significantly from neighboring units.

**Data Preparation**

The dataset includes polling unit information and the votes each party received. The initial steps involved cleaning the data and adding geolocation data (latitude and longitude) for each polling unit.

**Neighbor Identification**

To identify neighboring polling units, we used geospatial analysis techniques. We calculated the geodesic distance between each polling unit and defined a radius of 1 km to determine neighboring units. The geopy library was used to perform these calculations efficiently.

**Outlier Score Calculation**

For each polling unit, we compared the votes received by each party with those of neighboring units. An outlier score was calculated based on the deviation of votes from neighboring units. The higher the score, the more significant the deviation, indicating potential irregularities.

**Results**

The polling units were sorted by outlier scores for each party. The top 3 outliers were identified and analyzed in detail. Visualizations were used to represent the results effectively.

**Top 3 Outliers**

1. Polling Unit A

- Significant deviation in votes for APC.

- Neighboring units showed consistent voting patterns, but Unit A had an unusually high vote count for APC.

2. Polling Unit B

- Large discrepancy in votes for LP.

- Neighboring units had lower votes for LP, suggesting potential manipulation.

3. Polling Unit C

- High outlier score for PDP.

- Neighboring units had a balanced vote distribution, while Unit C had an anomaly in favor of PDP.

Conclusion

The analysis revealed several polling units with significant deviations in voting patterns. These outliers indicate potential voting irregularities. Further investigation is recommended to ensure the transparency and integrity of the election results.