### Detailed Report: Outlier Detection in Election Data Using Geospatial Analysis

#### **Introduction**

This report aims to detect outliers in election data based on geospatial analysis. The goal is to identify polling units with irregular voting patterns for four major political parties: APC, LP, PDP, and NNPP. The analysis helps pinpoint potential areas of concern regarding electoral anomalies or fraud.

**Methodology**

The analysis involves the following steps:

1. **Data Collection**: The dataset used in this analysis originates from the Taraba\_crosschecked spreadsheet, containing voting data for polling units. To facilitate geospatial analysis, longitude and latitude coordinates were obtained using the Google Maps Geocoding API. This allowed for precise geographical localization of each polling unit within the state.
2. **Outlier Detection**: Outlier scores are calculated based on deviations from expected voting patterns within geographical proximity.
3. **Geodesic Distance Calculation**: Geodesic distance was used to measure the distance between polling units, with a predefined radius of 1 km.
4. **Visualization**: Bar charts and pie charts are created to visualize the top outliers for each party.

#### **Summary of Findings**

The dataset includes outlier scores for various polling units, highlighting significant deviations for the four political parties. The top outliers for each party are identified and visualized.

**Top Outliers Summary**:

* **APC**: The top 3 polling units with the highest outlier scores.
* **LP**: The top 3 polling units with the highest outlier scores.
* **PDP**: The top 3 polling units with the highest outlier scores.
* **NNPP**: The top 3 polling units with the highest outlier scores.

#### **Visualization**

Incorporate visualizations such as maps, charts, and graphs to visually represent outlier distributions and voting patterns, enhancing the clarity and impact of the findings presented in this report.

##### **APC Outliers**

* **Bar Chart**: Displays the top 3 APC outliers by outlier score.
* **Pie Chart**: Shows the percentage distribution of outlier scores among the top 10 APC polling units.
* **It shows that the polling unit with the pu\_code 34-06-06-007** and pu\_Name **ISLAMIYA, ISLAMIYA PRIMARY SCHOOL** has the highest APC outliers.
* **Its neighbouring units are** ['34-06-05-019', '34-06-06-017'].

##### **LP Outliers**

* **Bar Chart**: Displays the top 3 LP outliers by outlier score.
* **Pie Chart**: Shows the percentage distribution of outlier scores among the top 3 LP polling units
* **It shows that the polling unit with the pu\_code 34-13-07-006** and pu\_Name **TUTUWA , TUTUWA PRIMARY SCHOOL**, has the highest LP outliers.
* **Its neighbouring unit is ['34-13-07-017']**

##### **PDP Outliers**

* **Bar Chart**: Displays the top 3 PDP outliers by outlier score.
* **Pie Chart**: Shows the percentage distribution of outlier scores among the top 3 PDP polling units.
* **It shows that the polling unit with the pu\_code 34-08-04-004,** pu\_Name **UNG. ARDIDO/DAN, ARDIDO/DAN**, has the highest PDP outliers.
* **Its neighbouring units is ['34-08-04-009'**].

##### **NNPP Outliers**

* **Bar Chart**: Displays the top 3 NNPP outliers by outlier score.
* **Pie Chart**: Shows the percentage distribution of outlier scores among the top 3 NNPP polling units.
* **It shows that the polling unit with the pu\_code 34-06-05-001, pu\_Name ALH. TAMMAH I, KOFAR MAITO and pu\_Code 34-06-05-003 with the pu\_Name DAN BURAM, KOFAR ABOKI DAN BURAM have the highest NNPP outliers.**
* **Their neighbouring units are ['34-06-05-003'] and ['34-06-05-001'] respectively.**

**Conclusion**

The geospatial analysis and outlier detection approach employed in this report have successfully identified polling units with potential irregularities in voting patterns.

The analysis successfully identifies and visualizes the top outliers in election data for four major political parties. These visualizations help in understanding voting irregularities and can be used to further investigate the causes of these anomalies. By focusing on the top outliers, electoral bodies can target their efforts to ensure the integrity of the voting process.

**Recommendations**

1. **Enhanced Monitoring:** Implement enhanced monitoring mechanisms to detect and prevent potential irregularities in future elections.
2. **Further Investigation:** Conduct further investigation into identified outliers to understand underlying causes and implications for electoral processes.
3. **Public Awareness:** Raise public awareness about the significance of monitoring and reporting irregularities to uphold election integrity.