**Advanced SQL Project Questions for ACLED Dataset answer with your validated and running SQL statements**

**1. Create a view for high-fatality events in the last 5 years [Assumption: An event that led to more than one fatality is a high fatality event]**

SELECT \*

FROM TBL\_ACLED\_FINAL

WHERE YEAR >= 2019 AND FATALITIES > 1;

**2. Calculate the rolling average of fatalities over a 7-day window for 'Battles': [Should I add the daily fatalities or get the daily averages?] - Remember to round to whole number**

WITH DAILY AS(

SELECT

EVENT\_DATE,

*SUM*(FATALITIES) AS DAILY\_FATALITIES

FROM TBL\_ACLED\_FINAL

WHERE EVENT\_TYPE = 'Battles'

GROUP BY EVENT\_DATE

ORDER BY EVENT\_DATE)

SELECT

EVENT\_DATE,

*AVG*(DAILY\_FATALITIES) OVER (ORDER BY EVENT\_DATE ROWS BETWEEN 6 PRECEDING AND CURRENT ROW) AS SEVEN\_DAY\_MAVG

FROM DAILY;

**3. Generate a report showing the month-over-month change in event counts for 'Protests'**

SELECT

*TRUNC*(EVENT\_DATE,'MONTH') AS MONTH,

--COUNT(EVENT\_ID\_CNTY) AS EVENT\_COUNT,

--LAG(COUNT(EVENT\_ID\_CNTY),1) OVER(ORDER BY TRUNC(EVENT\_DATE,'MONTH')) AS PREV\_EVENT\_COUNT,

*COUNT*(EVENT\_ID\_CNTY) - *LAG*(*COUNT*(EVENT\_ID\_CNTY),1) OVER(ORDER BY *TRUNC*(EVENT\_DATE,'MONTH')) AS EVENT\_CHANGE

FROM TBL\_ACLED\_FINAL

WHERE EVENT\_TYPE = 'Protests'

GROUP BY *TRUNC*(EVENT\_DATE,'MONTH')

ORDER BY MONTH;

**4. Find the top 5 regions with the highest diversity of disorder types [Data has 5 regions each having 4 disorder types]**

SELECT

REGION,

*COUNT*(DISTINCT DISORDER\_TYPE)

FROM TBL\_ACLED\_FINAL

GROUP BY REGION;

**5. Analyze the trend of political violence over the years using a common table expression (CTE)**

WITH CTE AS (

SELECT

YEAR,

*COUNT*(\*) AS EVENT\_COUNT

FROM TBL\_ACLED\_FINAL

WHERE DISORDER\_TYPE LIKE '%Political violence%'

GROUP BY YEAR

ORDER BY YEAR)

SELECT

YEAR,

EVENT\_COUNT,

*LAG*(EVENT\_COUNT,1) OVER (ORDER BY YEAR) AS EVENT\_LAST\_YEAR,

EVENT\_COUNT- *LAG*(EVENT\_COUNT,1) OVER (ORDER BY YEAR) AS TREND

FROM CTE

GROUP BY YEAR, EVENT\_COUNT

ORDER BY YEAR;

**6. Create a materialized view summarizing events by country and disorder type**

SELECT

COUNTRY,

DISORDER\_TYPE,

*COUNT*(\*) AS EVENT\_COUNT

FROM TBL\_ACLED\_FINAL

GROUP BY GROUPING SETS((COUNTRY), (DISORDER\_TYPE),());

**7. Compare events from 2020 and 2021 involving 'Al Shabaab'**

SELECT

YEAR,

*COUNT*(\*)

FROM TBL\_ACLED\_FINAL

WHERE YEAR IN ('2020','2021') AND (ACTOR1 = 'Al Shabaab' OR ASSOC\_ACTOR\_1 = 'Al Shabaab'

OR ACTOR2 = 'Al Shabaab' OR ASSOC\_ACTOR\_2 = 'Al Shabaab')

GROUP BY YEAR

ORDER BY YEAR;

**8. Create an index to optimize queries for high-fatality events in 'Nigeria'**

**9. Rank events by fatalities within each country for 'Protests'**

SELECT

*DENSE\_RANK*() OVER(ORDER BY *SUM*(FATALITIES) DESC) AS FATALITIES\_RANK,

COUNTRY,

*SUM*(FATALITIES) AS COUNT\_FATALITIES

FROM TBL\_ACLED\_FINAL

WHERE EVENT\_TYPE = 'Protests'

GROUP BY COUNTRY

ORDER BY COUNT\_FATALITIES DESC;

**10. Generate a report showing the cumulative number of 'Violence against civilians' events over time**

SELECT

EVENT\_DATE,

*COUNT*(\*),

*SUM*(*COUNT*(\*)) OVER(ORDER BY EVENT\_DATE) AS CUMM\_SUM

FROM TBL\_ACLED\_FINAL

WHERE EVENT\_TYPE = 'Violence against civilians'

GROUP BY EVENT\_DATE

ORDER BY EVENT\_DATE;

**11. Analyze the distribution of event types in 'Kenya' using a pivot table**

SELECT \*

FROM (SELECT EVENT\_TYPE,

YEAR

FROM TBL\_ACLED\_FINAL

WHERE COUNTRY = 'Kenya')

PIVOT

(

*COUNT*(\*)

FOR EVENT\_TYPE IN ('Protests','Violence against civilians','Battles','Strategic developments','Riots',

'Explosions/Remote violence'))

ORDER BY YEAR;

**12. Find the frequency of high-fatality events (over 100 fatalities) involving specific actors**

SELECT

ACTOR1,

*COUNT*(FATALITIES) AS HIGH\_FATALITY\_COUNT

FROM TBL\_ACLED\_FINAL

WHERE FATALITIES > 100

GROUP BY ACTOR1

ORDER BY HIGH\_FATALITY\_COUNT DESC;

**13. Identify co-occurrences of 'Violence against civilians' and 'Protests' events on the same day and location**

WITH VAC\_LIST AS(

SELECT \*

FROM TBL\_ACLED\_FINAL

WHERE EVENT\_TYPE = 'Violence against civilians'),

PROTEST\_LIST AS(

SELECT \*

FROM TBL\_ACLED\_FINAL

WHERE EVENT\_TYPE = 'Protests'

)

SELECT VAC\_LIST.EVENT\_ID\_CNTY,

VAC\_LIST.EVENT\_DATE

FROM VAC\_LIST

LEFT JOIN PROTEST\_LIST

ON VAC\_LIST.LOCATION = PROTEST\_LIST.LOCATION

WHERE VAC\_LIST.EVENT\_DATE = PROTEST\_LIST.EVENT\_DATE

ORDER BY VAC\_LIST.EVENT\_DATE;

**14. Explore hierarchical relationships of events caused by 'Al Shabaab' [ One category]**

SELECT

EVENT\_TYPE,

ACTOR1,

ACTOR2,

*COUNT*(\*) AS EVENT\_COUNT

FROM TBL\_ACLED\_FINAL

WHERE ACTOR1 = 'Al Shabaab'

OR ACTOR2 = 'Al Shabaab' AND ACTOR2 IS NOT NULL

GROUP BY ROLLUP(EVENT\_TYPE,ACTOR1, ACTOR2);

**15. Perform advanced text search for events mentioning 'peaceful' and 'violence' in notes**

SELECT \*

FROM TBL\_ACLED\_FINAL

WHERE NOTES LIKE '%peaceful%violence%';

**16. Partition the table by year for performance improvement and analyze 2023 events**

**17. Analyze the relationship between event types and average fatalities using complex joins and subqueries [Fatalities to whole numbers)**

SELECT \*

FROM (SELECT

YEAR,

EVENT\_TYPE,

FATALITIES

FROM TBL\_ACLED\_FINAL)

PIVOT(

*AVG*(FATALITIES) AS AVG\_FATALITIES

FOR EVENT\_TYPE IN ('Riots','Explosions/Remote violence','Strategic developments',

'Protests','Violence against civilians','Battles')

)

ORDER BY YEAR;

**18. Detect anomalies in fatalities using statistical z-scores**

SELECT \*

FROM (

SELECT

EVENT\_DATE,

FATALITIES,

FATALITIES - AVG(FATALITIES) OVER() / STDDEV(FATALITIES) OVER() AS Z\_SCORE

FROM TBL\_ACLED\_FINAL)

WHERE Z\_SCORE > '2.576' OR Z\_SCORE < '-2.576';