THE 2022 GENERAL ELECTIONS DATA ANALYSIS

THE PRESIDENTIAL SEAT

"In 2022, Kenyans participated in a highly contested election dubbed 'THE FIFTH', featuring intense presidential campaigns. The dynamic landscape saw opinion polls fluctuating daily, with the main contenders being Raila and Ruto. Two years later, this analysis aims to highlight key insights for future candidates facing formidable opponents."

In Kenya a candidate shall be declared elected as President if the candidate receives--

(a) more than half of all the votes cast in the election; (50% + 1) and
(b) at least twenty-five per cent of the votes cast in each of more than half of the counties.

"Ruto received 7,176,141 (50.49 percent) of the votes, **narrowly** beating his rival, Raila Odinga, who managed 6,942,930 votes (48.85 percent)"

ALJAZEERA NEWS

Raila Odinga
Orange Democratic Movement
/Azimio La Umoja Coalition

48.85%

6,942,930 VOTES

VS

50.49%

William Ruto
United Democratic Alliance/
Kenya Kwanza Alliance

7,176,141 VOTES

>25% in 35 of 47 Counties

>25% in 40 of 47 Counties

AREAS TO HIGHLIGHT:

--Registered voters that did not show up
       -Highlighting most affected counties

--Rejected cast ballots
       -Highlighting counties with most rejected ballots

--strongholds of contestants
       -Highlighting average value that qualifies a region as a stronghold

--strategies in acquiring votes in counties with most and least registered voters
       -How the winning contestant used this to their advantage

--The contribution of turnups to winning
       -Especially highlighting why low turnups may reduce the effect of a stronghold

etc.

I created a database called 2022_elections_db
With 5 normalized tables
1 general table elections_2022_ was also used(un-
normalized version of all 5 tables)-- necessary to
simplify this project

The 5 tables:

--aspirants
--registeredVoters
--spoiltVotes
--totalValidVotes
--votesperaspirant

```sql
• CREATE VIEW POLLS_VW AS
  SELECT SUM(registeredvoters) AS REGISTERED_VOTERS,
  SUM(totalvalidvotes) AS TOTAL_VALID_VOTES,
  SUM(REJECTEDBALLOTS) AS REJECTED_BALLOTS
  FROM elections_2022_;

• SELECT * FROM POLLS_VW;
```

creating a view to get:
-Total Registered Voters
-Total Valid Votes
-Rejected ballots

REGISTERED_VOTERS	TOTAL_VALID_VOTES	REJECTED_BALLOTS
22120458	14213137	113614

REGISTERED_VOTERS	TURNUP	PERCENTAGE_OF_TURNUP	FAILED_TURNUP	PERCENTAGE_OF_FAILED_TURNUP
22120458	14326751	64.77	7793707	35.23

35.23% did not turn up to vote

COUNTYNAME	TT_VOTES_CAST	REGISTEREDVOTERS	NO_TURNUP_PER_COUNTY
MOMBASA	281113	642362	56.24
KILIFI	288687	588842	50.97
GARISSA	110550	201513	45.14
KWALE	180368	328316	45.06
NAIROBI	1352236	2416551	44.04

Top 5 counties with the greatest no turnup to vote
The **coastal region** was strongly highlighted with counties having even **more than half** of registered voters **not showing up** to vote.

COUNTYNAME	TT_VOTES_CAST	REGISTEREDVOTERS	NO_TURNUP_PER_COUNTY
BOMET	301151	377023	20.12
WEST POKOT	174963	220042	20.49
KERICHO	336337	428126	21.44
ELGEYO/MARAKWET	166761	213904	22.04
NAROK	310029	398852	22.27

-Table 3 shows counties where **most** registered voters **turned up**.
--The query written asked for the bottom 5 counties when ranked based on no turn up of voters
-Aspirants should mark the **rift-valley region** as **promising** in regards to voters turn up.

```
45    SELECT
46    SUM(totalvalidvotes) AS TOTAL_VALID_VOTES,
47    SUM(REJECTEDBALLOTS) AS REJECTED_BALLOTS,
48    ROUND((SUM(REJECTEDBALLOTS)/SUM(totalvalidvotes)*100),2) AS REJECTED_VOTES
49    FROM elections_2022_;
50
51 •  SELECT * FROM POLLS_VW;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: IA

TOTAL_VALID_VOTES	REJECTED_BALLOTS	REJECTED_VOTES
14213137	113614	0.80

Insight:
Not a significant number of votes were rejected

```
73 • SELECT COUNTYNAME,SUM(TOTALVALIDVOTES+REJECTEDBALLOTS) AS TT_VOTES_CAST,
74    ROUND((REJECTEDBALLOTS/SUM(TOTALVALIDVOTES+REJECTEDBALLOTS)*100),2)
75    AS PERCENT_REJECTS
76    FROM ELECTIONS_2022_
77    GROUP BY COUNTYCODE
78    ORDER BY PERCENT_REJECTS DESC
79    LIMIT 5;
```

sult Grid   Filter Rows: Export: Wrap Cell Content: $\underline{I}A$

COUNTYNAME	TT_VOTES_CAST	PERCENT_REJECTS
LAMU	50957	1.74
MOMBASA	281113	1.36
BUNGOMA	410667	1.34
VIHIGA	186448	1.13
TANA RIVER	94573	1.13

Even with the total rejected % being insignificant;a deeper analysis of what percentage of **CASTED** votes ended up being rejected per county **highlighted** the **coastal** region(Lamu,Mombasa,Tana River) and the **western** region(Bungoma,Vihiga) most affected/as areas of concern.

COMPARING STRONGHOLDS

	COUNTYNAME	WILLIAM_RUTO	RAILA_ODINGA
▶	BOMET	75.7057	3.5497
	ELGEYO/MARAKWET	74.8153	2.2875
	KERICHO	74.4783	3.5160
	NANDI	69.0989	6.4061
	THARAKA-NITHI	62.5441	6.4932
	BARINGO	62.3144	14.6659

	COUNTYNAME	RAILA_ODINGA	WILLIAM_RUTO
▶	HOMA BAY	72.5415	0.6345
	SIAYA	69.5447	0.8096
	KISUMU	69.1358	1.6479
	MIGORI	62.7085	11.1981
	BUSIA	54.2304	11.7080
	MANDERA	48.9688	13.0629

Both contestants gathered strongly from their home regions.
The difference may then have been influenced by counties that supported them equally,or from their running mates or regions that were not neutral,or how well they gathered from their strongholds.

--winning contestant(Ruto) WOULD have gathered over 50% **FOR ALL** his top 6 strongolds.

Calculation is done based on registered voters; using only voters that showed up results in percentages close to 100% hence a standardization method **required**

```sql
34 •  SELECT COUNTYNAME,WILLIAMRUTO,SUM(TOTALVALIDVOTES+REJECTEDBALLOTS) AS TT_CASTED,
35     ROUND((WILLIAMRUTO/SUM(TOTALVALIDVOTES+REJECTEDBALLOTS)*100),2) AS WR
36     FROM ELECTIONS_2022_
37     GROUP BY COUNTYCODE
38     ORDER BY WR DESC
39     LIMIT 5;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: IA

COUNTYNAME	WILLIAMRUTO	TT_CASTED	WR
ELGEYO/MARAKWET	160033	166761	95.97
KERICHO	318861	336337	94.80
BOMET	285428	301151	94.78
NANDI	280813	309067	90.86
THARAKA-NITHI	145081	162558	89.25

TABLE 1:Ruto's stronghold

TABLE 2:Raila's stronghold

```sql
41 •  SELECT COUNTYNAME,RAILAODINGA,SUM(TOTALVALIDVOTES+REJECTEDBALLOTS) AS TT_CASTED,
42     ROUND((RAILAODINGA/SUM(TOTALVALIDVOTES+REJECTEDBALLOTS)*100),2) AS RAILAODINGA_
43     FROM ELECTIONS_2022_
44     GROUP BY COUNTYCODE
45     ORDER BY RAILAODINGA_ DESC
46     LIMIT 5;
47
48 •  SELECT COUNTYNAME,WILLIAMRUTO,REGISTEREDVOTERS,
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: IA

COUNTYNAME	RAILAODINGA	TT_CASTED	RAILAODINGA_
HOMA BAY	399784	406157	98.43
SIAYA	371092	378248	98.11
KISUMU	419997	433577	96.87
MIGORI	294136	349384	84.19
BUSIA	226042	279619	80.84

As was explained earlier;the strongholds remain but results seem unrealistic

```sql
112 •  SELECT COUNTYNAME,RAILAODINGA,WILLIAMRUTO,
113     ABS(WILLIAMRUTO-RAILAODINGA) AS DIFFERENCE
114     FROM ELECTIONS_2022_
115     WHERE ABS(WILLIAMRUTO-RAILAODINGA) <5000;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

COUNTYNAME	RAILAODINGA	WILLIAMRUTO	DIFFERENCE
LAMU	26160	22876	3284
MARSABIT	55675	58782	3107
DIASPORA	3727	2190	1537

There were only 2 counties where the difference of the top contestants
was very minimal/can be said that they almost tied.
It MAY be said that the contestants alternated with significantly
compensative margins in the counties that do not fall into the
strongholds or weakpoints(fall in the middle)
Diaspora can also be ommitted as a promising source of significant votes.

```
91 •  SELECT COUNTYNAME,REGISTEREDVOTERS
92    FROM ELECTIONS_2022_
93    GROUP BY COUNTYCODE
94    ORDER BY REGISTEREDVOTERS DESC
95    LIMIT 3;
96
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: TA

COUNTYNAME	REGISTEREDVOTERS
NAIROBI	2416551
KIAMBU	1275168
NAKURU	1055515

COUNTYCODE	COUNTYNAME	WILLIAMRUTO
22	KIAMBU	606429

COUNTYCODE	COUNTYNAME	RAILAODINGA
47	NAIROBI	767395

COUNTYCODE	COUNTYNAME	GEORGEWAJACKOYAH
47	NAIROBI	5807

COUNTYCODE	COUNTYNAME	WAIHIGAMWAURE
47	NAIROBI	4390

With Nairobi boasting the highest number of voters, it's notable that **all aspirants, except the eventual winner**, Ruto, received their maximum votes from the capital. In Kiambu, the county with the second-highest number of registered voters, Ruto secured the highest number of votes, hinting at the strategic advantage of focusing efforts beyond densely populated areas.

```sql
85 •  SELECT COUNTYCODE,COUNTYNAME,REGISTEREDVOTERS FROM ELECTIONS_2022_
86   ⊖ WHERE REGISTEREDVOTERS IN(SELECT MIN(REGISTEREDVOTERS)
87   └ FROM ELECTIONS_2022_ WHERE COUNTYCODE !=48);
88
```

COUNTYCODE	COUNTYNAME	REGISTEREDVOTERS
5	LAMU	81468

```sql
89 •  SELECT COUNTYNAME,
90  ⊖ CASE
91    WHEN GREATEST(RAILAODINGA, WILLIAMRUTO, WAIHIGAMWAURE, GEORGEWAJACKOYAH) = RAILAODINGA THEN 'Raila Odinga'
92    WHEN GREATEST(RAILAODINGA, WILLIAMRUTO, WAIHIGAMWAURE, GEORGEWAJACKOYAH) = WILLIAMRUTO THEN 'William Ruto'
93    WHEN GREATEST(RAILAODINGA, WILLIAMRUTO, WAIHIGAMWAURE, GEORGEWAJACKOYAH) = WAIHIGAMWAURE THEN 'Waihiga Mwaura'
94    WHEN GREATEST(RAILAODINGA, WILLIAMRUTO, WAIHIGAMWAURE, GEORGEWAJACKOYAH) = GEORGEWAJACKOYAH THEN 'George Wajackoyah'
95       END AS WINNER
96    FROM ELECTIONS_2022_
97    WHERE COUNTYNAME = 'LAMU';
98
99
100
```

COUNTYNAME	WINNER
LAMU	Raila Odinga

Lamu, emerged with the highest number of rejected votes, it is also situated in the coastal region illustrated as having low voter turnout.
Raila emerged top in Lamu.On the other hand,Lamu was a result in the regions with minimum difference in votes among top contestants
The significance of this result is clouded by the fact that Lamu has only around 8,000 voters, rendering the analysis of a 5,000-vote margin difference inconclusive. This underscores the complexity of Lamu as a region for contestants, requiring careful navigation despite its potential as both promising and challenging.

```
58 •  SELECT F1.COUNTYNAME
59     FROM F1
60     INNER JOIN F2
61     ON F1.COUNTYNAME = F2.COUNTYNAME;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

COUNTYNAME
► BOMET
KERICHO
ELGEYO/MARAKWET

Using view it was determined that counties out of the 5 counties with the major turnup **3/5** contributed as strongholds of William Ruto

```
65
66 •  SELECT F4.COUNTYNAME
67     FROM F1
68     INNER JOIN F4
69     ON F1.COUNTYNAME = F4.COUNTYNAME;
```
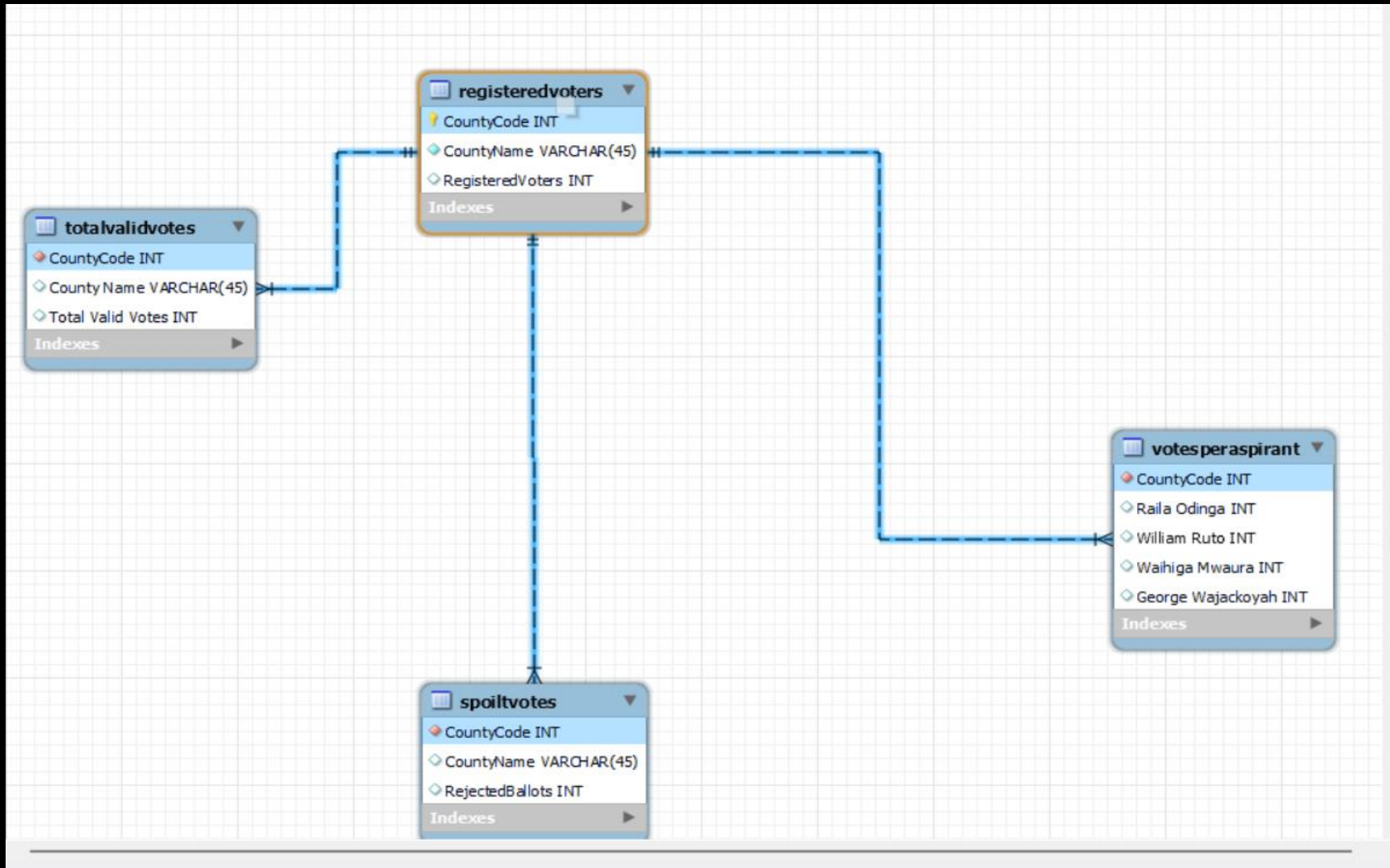
Result Grid | Filter Rows: | Export: | Wrap Cell Content:

COUNTYNAME

On the other hand Raila did not have a stronghold county in the top 5 counties that had the leading turnout

Contestants must prioritize their strongholds and ensure dependable turnout; otherwise, their efforts are futile

EER DIAGRAM

TOP 10 COUNTIES AND VOTES FOR RUTO VS RAILA

William Ruto	County	Raila Odinga
606,429	KIAMBU	210,580
561,775	NAIROBI	767,395
455,864	NAKURU	226,052
398,946	MERY	103,679
343,349	MURANGA	73,526
255,906	BUNGOMA	145,280
141,166	KAKAMEGA	357,857
101,456	MACHAKOS	304,809
10,011	KISUMU	419,997
3,497	HOMA BAY	399,784

https://public.tableau.com/app/profile/ann.wanjiku1711/viz/Elections_2022/Sheet4?publish=yes

```sql
ALTER TABLE aspirants
MODIFY COLUMN aspirantname VARCHAR(45);
select * from aspirants;
```

No project is truly complete without reminders. So,to the trigger and alter that rescued the project!

```sql
DELIMITER //
CREATE TRIGGER aspirants_before_insert
BEFORE INSERT ON ASPIRANTS
FOR EACH ROW
BEGIN
SET NEW.ASPIRANTNAME = CONCAT(UPPER(SUBSTRING(NEW.ASPIRANTNAME,1,1)),
SUBSTRING(NEW.ASPIRANTNAME FROM 2));
END//
DELIMITER ;
```