Political Campaign Data Analysis: What Actually Wins Elections?

Joe Wiegert

Everyone talks about what "wins" elections. Is it the money? The outreach? The sheer number of doors knocked? Political consultants love to make claims, but what do the numbers actually say?

That's what I set out to analyze in this project. Using SQL, I built queries to cut through the noise and figure out which campaign factors had the biggest impact on election wins. I broke it down into four major areas:

Fundraising – Does more money equal more wins?

Voter Outreach – Do more phone calls and door knocks actually move the needle?

Event Attendance – Does drawing big crowds mean a campaign is more likely to succeed?

Voter Demographics & Turnout – Who is voting, and how much does turnout predict victory?

The results? Voter turnout was the strongest predictor of success. Not fundraising, not just outreach—actual votes cast. Campaigns that hit at least 62% voter turnout overwhelmingly won their elections. But beyond that, I uncovered patterns that challenge some of the conventional wisdom on campaign strategy.

This dataset was AI-generated, which meant I had to be hyper-aware of limitations in data distribution. But that actually made this a great exercise in debugging, refining queries, and thinking critically about data quality. More importantly, this project solidified my belief that data tells a story—but only if you know how to ask the right questions.

The following table schemas and result previews were used in this project:

Table: campaign_donations

Query:

SELECT COUNT(*)

FROM campaign_donations

;

Result: 7500

Table: campaigns

Query:

SELECT COUNT(*)

FROM campaigns

;

Result: 40

Table: election_results

result_id	campaign_id	state	race_type	election_date	total_votes	votes_received	percent_of_vote	election
1	1	Wisconsin	Senate	2024-11-05	4904806	2684511	54.73	Win
2	2	Georgia	Governor	2024-11-05	772532	274086	35.48	Loss
3	3	Ohio	Local	2024-11-05	4494758	1873743	41.69	Loss
4	4	Arizona	Senate	2024-11-05	3027883	1005046	33.19	Loss
5	5	Florida	House	2024-11-05	4165386	1999020	47.99	Loss
6	6	Georgia	Local	2024-11-05	290957	180996	62.21	Win
7	7	Georgia	Governor	2024-11-05	806734	300993	37.31	Loss
8	8	Pennsylvania	House	2024-11-05	4880903	2551545	52.28	Win
9	9	Ohio	Governor	2024-11-05	3845026	2214336	57.59	Win

Query:

SELECT COUNT(*)

FROM election_results

;

Result: 46

Table: event_attendance

attendance_id	event_id	attendee_name	attendee_type	attended_as	engagement_le
1	1150	James Thomas	Voter	Attendee	Low
2	1522	Isabella Cooper	Volunteer	Organizer	Low
3	613	Brandon Mitchell	Voter	Attendee	Low
4	419	Michelle King	Volunteer	Attendee	Low
5	1316	Amanda Turner	Voter	Attendee	Low
6	763	Kevin Thomas	Voter	Attendee	Low
7	1726	Jason Allen	Donor	Attendee	Low

Query:

SELECT COUNT(*)

FROM event_attendance

;

Result: 6500

Table: events

	event_id	campaign_id	event_date	event_location	event_type	expected_attendance	actual_attendance	funds_raised	doors_knocked	calls_made
▶	1	28	2023-09-06	Minneapolis, MN	Canvassing				1652.0	
	2	45	2024-09-23	Tampa, FL	Canvassing				1664.0	
	3	43	2024-09-08	Phoenix, AZ	Rally	162.0	173.0			
	4	45	2024-07-12	Seattle, WA	Canvassing				2868.0	
	5	20	2023-09-12	Portland, OR	Rally	4419.0	4642.0			
	6	11	2024-04-20	Nashville, TN	Phone Bank					9585.0
	7	24	2024-01-19	Columbus, OH	Phone Bank					9608.0
	8	46	2024-05-09	Chicago, IL	Fundraiser	196.0	182.0	140314.57		
	9	23	2024-09-21	Minneapolis, MN	Rally	2344.0	2811.0			
	10	39	2024-05-21	Chicago, IL	Canvassing				2999.0	

Table events columns:
- event_id — int
- campaign_id — int
- event_date — text
- event_location — text
- event_type — text
- expected_attendance — text
- actual_attendance — text
- funds_raised — text
- doors_knocked — text
- calls_made — text

Query:

SELECT COUNT(*)

FROM events

;

Result: 1800

Table: outreach

	outreach_id	campaign_id	voter_name	contact_method	contact_date	response	follow_up_needed	volunteer_name
▶	1	5	Ava Brown	Door Knock	2023-11-28	No Answer	No	Matthew Davis
	2	44	Ryan Cooper	Email	2024-09-13	No Answer	No	Jessica Martinez
	3	30	Elizabeth Clark	Phone Call	2024-07-01	No Answer	No	Samantha Anderson
	4	9	Jessica Phillips	Text	2022-10-01	Positive	Yes	Jennifer Reed
	5	13	Samantha Thomas	Mailer	2024-07-23	No Answer	No	Ava Moore
	6	20	Jason Gonzalez	Text	2023-10-19	Neutral	Yes	Ava Scott
	7	17	David Hall	Email	2023-07-23	No Answer	No	Elizabeth Morgan
	8	19	Lauren Green	Phone Call	2024-10-28	Neutral	Yes	Citizens for Liberty

Table outreach columns:
- outreach_id — int
- campaign_id — int
- voter_name — text
- contact_method — text
- contact_date — text
- response — text
- follow_up_needed — text
- volunteer_name — text

SELECT COUNT(*)

FROM outreach

;

Result: 25000

Table: voters

Query:

SELECT COUNT(*)

FROM voters

;

Result: 50000

-- Starting my analysis with campaign fundraising; we want to know how much money did each campaign raise?

-- May also notice differences in fundraising by political affiliation

SELECT c.campaign_id, c.campaign_name, c.party_affiliation,

    SUM(d.donation_amount) AS total_funds_raised

FROM campaigns AS c

JOIN campaign_donations AS d

        ON c.campaign_id = d.campaign_id

GROUP BY c.campaign_id, c.campaign_name, c.party_affiliation

ORDER BY total_funds_raised DESC

;

Result:

campaign_id	campaign_name	party_affiliation	total_funds_raised
27	The People's Voice	Republican	40571
15	Moving Forward	Republican	36116
31	Stronger Communities	Republican	32722
24	Unite for Change	Republican	32066
22	Bold Solutions	Democratic	31945
7	Stronger Together	Democratic	31664
30	Fighting for Justice	Republican	31428
34	Equality Now	Independent	29622

-- This shows how much money each campaign raised

WITH donation_cte(campaign_id, campaign_name, party_affiliation, total_funds_raised) AS

(

SELECT c.campaign_id, c.campaign_name, c.party_affiliation,

    SUM(d.donation_amount) AS total_funds_raised

FROM campaigns AS c

JOIN campaign_donations AS d

        ON c.campaign_id = d.campaign_id

GROUP BY c.campaign_id, c.campaign_name, c.party_affiliation

ORDER BY total_funds_raised DESC

)

SELECT party_affiliation, SUM(total_funds_raised) AS total

FROM donation_cte

GROUP BY party_affiliation

ORDER BY total DESC

;

Result:

party_affiliation	total
Republican	341470
Democratic	331147
Independent	293348

-- This CTE allows us to see that Republicans raised only $10,000 more than Democrats

-- We can also see that Independents raised a significant chunk of the total money raised during the cycle

WITH top_10_cte(campaign_id, campaign_name, party_affiliation, total_funds_raised) AS

(

SELECT c.campaign_id, c.campaign_name, c.party_affiliation,

    SUM(d.donation_amount) AS total_funds_raised

FROM campaigns AS c

JOIN campaign_donations AS d

        ON c.campaign_id = d.campaign_id

GROUP BY c.campaign_id, c.campaign_name, c.party_affiliation

ORDER BY total_funds_raised DESC

LIMIT 10)

SELECT party_affiliation, COUNT(party_affiliation)

FROM top_10_cte

GROUP BY party_affiliation

ORDER BY COUNT(party_affiliation) DESC

;

Result:

party_affiliation	COUNT(party_affiliation)
Republican	5
Democratic	4
Independent	1

-- This CTE shows that half of the top 10 campaigns, in terms of fundraising, were Republican campaigns

-- Only 4 Democrat campaigns made it to the top 10, with Independent campaigns holding the remaining spot

-- Does this information coincide with a winning campaign?

WITH donation_cte(campaign_id, campaign_name, party_affiliation, total_funds_raised) AS

(

SELECT c.campaign_id, c.campaign_name, c.party_affiliation,

    SUM(d.donation_amount) AS total_funds_raised

FROM campaigns AS c

JOIN campaign_donations AS d

    ON c.campaign_id = d.campaign_id

GROUP BY c.campaign_id, c.campaign_name, c.party_affiliation

)

SELECT d.campaign_name, d.party_affiliation, d.total_funds_raised, e.election_result

FROM donation_cte AS d

JOIN election_results AS e

    ON d.campaign_id = e.campaign_id

ORDER BY d.total_funds_raised DESC

;

Result:

campaign_name	party_affiliation	total_funds_raised	election_result
The People's Voice	Republican	40571	Win
Moving Forward	Republican	36116	Loss
Stronger Communities	Republican	32722	Loss
Unite for Change	Republican	32066	Loss
Bold Solutions	Democratic	31945	Loss
Stronger Together	Democratic	31664	Loss
Fighting for Justice	Republican	31428	Loss
Equality Now	Independent	29622	Win
New Vision for America	Democratic	29414	Loss
United for Progress	Democratic	29140	Loss

-- It is surprising that the win percentage for the campaigns that raised the most is actually below 50%

-- In fact, of the top 10 fundraising campaigns, only 2 of them won, suggesting fundraising is not the best predictor of election wins

-- Now let's actually quantify the win/loss for top 10 fundraisers

WITH donation_cte(campaign_id, campaign_name, party_affiliation, total_funds_raised) AS

(

SELECT c.campaign_id, c.campaign_name, c.party_affiliation,

       SUM(d.donation_amount) AS total_funds_raised

FROM campaigns AS c

JOIN campaign_donations AS d

       ON c.campaign_id = d.campaign_id

GROUP BY c.campaign_id, c.campaign_name, c.party_affiliation

ORDER BY total_funds_raised DESC

LIMIT 10

)

SELECT e.election_result, COUNT(*) AS total_campaigns,

       ROUND(COUNT(*) * 100.0 / (SELECT COUNT(*) FROM donation_cte), 0) AS percentage

```sql
FROM donation_cte AS d

JOIN election_results AS e

        ON d.campaign_id = e.campaign_id

GROUP BY e.election_result

;
```

Result:

election_result	total_campaigns	percentage
Loss	8	80
Win	2	20

```sql
-- 80% of the top 10 fundraising campaigns lost their elections

-- This suggests there are other factors that drive victories in this election

-- Let's explore the effectiveness of voter outreach

WITH outreach_summary(campaign_id, total_contacts) AS

(

SELECT o.campaign_id,

                COUNT(o.outreach_id) AS total_contacts

FROM outreach AS o

GROUP BY o.campaign_id

)

SELECT c.campaign_name, c.party_affiliation, os.total_contacts, e.election_result

FROM campaigns AS c

JOIN outreach_summary AS os

        ON c.campaign_id = os.campaign_id

JOIN election_results AS e

        ON c.campaign_id = e.campaign_id

ORDER BY os.total_contacts DESC

;
```

Result:

campaign_name	party_affiliation	total_contacts	election_result
United for a Better Tomorrow	Republican	614	Win
Defending Democracy	Republican	587	Loss
For the People	Independent	577	Win
One Nation, One Future	Democratic	577	Win
New Vision for America	Democratic	569	Loss
The People's Voice	Republican	568	Win
Securing Our Future	Independent	558	Win
Stronger Communities	Republican	557	Loss
Standing for You	Democratic	556	Loss
Building Bridges	Independent	551	Loss

-- More outreach does not necessarily equate to more wins

-- Since the win rate for top fundraising campaigns was only 20%, does higher outreach success correlates with election wins?

WITH outreach_summary(campaign_id, total_contacts) AS

(

SELECT o.campaign_id,

COUNT(o.outreach_id) AS total_contacts

FROM outreach AS o

GROUP BY o.campaign_id

)

SELECT e.election_result,

COUNT(*) AS total_campaigns,

ROUND(COUNT(*) * 100.0 / (SELECT COUNT(*) FROM outreach_summary), 2) AS percentage

FROM outreach_summary AS os

JOIN election_results AS e

ON os.campaign_id = e.campaign_id

WHERE os.total_contacts > 500

GROUP BY e.election_result

;

Result:

election_result	total_campaigns	percentage
Win	21	45.65
Loss	24	52.17

-- 45 of the total 46 campaigns appear here, so I think changing the threshold for total contacts would provide a more clear distinction

-- If we change the threshold of very high outreach being 550 total contacts, how does our output change?

WITH outreach_summary AS (

    SELECT o.campaign_id,

        COUNT(o.outreach_id) AS total_contacts

    FROM outreach AS o

    GROUP BY o.campaign_id

)

SELECT e.election_result,

    COUNT(*) AS total_campaigns,

    ROUND(COUNT(*) * 100.0 / (SELECT COUNT(*) FROM outreach_summary WHERE total_contacts >= 550), 2) AS percentage

FROM outreach_summary AS os

JOIN election_results AS e ON os.campaign_id = e.campaign_id

WHERE os.total_contacts >= 550

GROUP BY e.election_result

;

Result:

election_result	total_campaigns	percentage
Loss	8	61.54
Win	5	38.46

-- Interestingly, even among "very high" outreach campaigns, more than 50% still lost their election

-- Outreach does have a slightly higher win rate than fundraising, with 38% for outreach and 30% for fundraising

-- Fundraising and outreach alone are not predictors of election success, I want to explore two other areas: event attendance and voter demographics

-- For event attendance; do campaigns with higher event participation perform better in their elections?

-- Specifically, does event attendance correlate more with winning an election than fundraising or outreach? Also, any outliers?

```
WITH event_summary_cte(campaign_id, total_event_attendance) AS

(

SELECT e.campaign_id,

            SUM(e.actual_attendance) AS total_event_attendance

FROM events AS e

GROUP BY e.campaign_id

)

SELECT c.campaign_name,

    c.party_affiliation,

    es.total_event_attendance,

    e.election_result

FROM campaigns AS c

JOIN event_summary_cte AS es
```

```
    ON c.campaign_id = es.campaign_id

JOIN election_results AS e

    ON c.campaign_id = e.campaign_id

ORDER BY es.total_event_attendance DESC

;
```

Result:

campaign_name	party_affiliation	total_event_attendance	election_result
For the People	Independent	50053	Win
Bold Solutions	Democratic	45518	Loss
Fairness and Freedom	Democratic	41073	Win
Forward Together	Independent	38958	Loss
Standing for You	Democratic	38925	Loss
Together We Rise	Democratic	38658	Win
Unite for Change	Republican	36318	Loss
Stronger Communities	Republican	36000	Loss
A Future We Believe In	Democratic	35742	Loss
One Nation, One Future	Democratic	35388	Win

-- This does seem to indicate that winning campaigns drew more attendance overall

-- Some of the highest attended events still turned out losing campaigns

-- I want to apply a win percentage for high attendance (30,000) events


```
WITH event_summary_cte AS

(

SELECT e.campaign_id,

            SUM(e.actual_attendance) AS total_event_attendance

FROM events AS e

GROUP BY e.campaign_id

)

SELECT e.election_result,

    COUNT(*) AS total_campaigns,
```

ROUND(COUNT(*) * 100.0 / (SELECT COUNT(*) FROM event_summary_cte WHERE total_event_attendance >= 30000), 2) AS percentage

FROM event_summary_cte AS es

JOIN election_results AS e

ON es.campaign_id = e.campaign_id

WHERE es.total_event_attendance >= 30000

GROUP BY e.election_result

;

Result:

election_result	total_campaigns	percentage
Loss	10	52.63
Win	9	47.37

-- Event attendance is the strongest predictor so far, but it still does not surpass the 50% threshold for election success

-- Next, I will attempt to see if voter demographics is the strongest predictor of election success

-- These demographics may include: younger voters, high amounts of registered voters, or party-affiliation

WITH voter_age_cte AS

(

SELECT v.campaign_id,

            AVG(v.age) AS avg_voter_age

FROM voters AS v

GROUP BY v.campaign_id

)

SELECT c.campaign_name,

```sql
        c.party_affiliation,

        vac.avg_voter_age,

        e.election_result

FROM campaigns AS c

JOIN voter_age_cte AS vac

    ON c.campaign_id = vac.campaign_id

JOIN election_results AS e

    ON c.campaign_id = e.campaign_id

ORDER BY vac.avg_voter_age ASC

;
```

Result:

campaign_name	party_affiliation	avg_voter_age	election_result
Standing for You	Democratic	23.3042	Loss
A Brighter Future	Independent	23.3791	Win
A New Path	Independent	23.4044	Loss
Path to Prosperity	Republican	23.4108	Win
Common Ground	Republican	23.4200	Win
Progress for All	Republican	23.4539	Loss
Unite for Change	Republican	23.4646	Loss
New Vision for America	Democratic	23.4741	Loss
Honest Leadership	Democratic	23.4769	Loss
Empower the Future	Democratic	23.4835	Win
People First	Democratic	23.4839	Loss
Together We Rise	Democratic	23.4883	Win
United for a Better To...	Republican	23.5024	Win
Forward Together	Independent	23.5108	Loss

-- Younger voter bases had mixed results, with winning and losing elections across different parties

-- The Oldest voter base had a clearer correlation with winning elections


```sql
WITH voter_registration_cte AS

(
```

```sql
SELECT v.campaign_id,

    COUNT(*) AS total_voters,

    SUM(CASE WHEN v.registered_to_vote = 'Yes' THEN 1 ELSE 0 END) AS
registered_voters,

    ROUND(SUM(CASE WHEN v.registered_to_vote = 'Yes' THEN 1 ELSE 0 END) *
100.0 / COUNT(*), 2) AS registration_rate

  FROM voters AS v

  GROUP BY v.campaign_id

)

SELECT c.campaign_name,

    c.party_affiliation,

    vrc.registration_rate,

    e.election_result

FROM campaigns AS c

JOIN voter_registration_cte AS vrc

  ON c.campaign_id = vrc.campaign_id

JOIN election_results AS e

  ON c.campaign_id = e.campaign_id

ORDER BY vrc.registration_rate DESC

;
```

Result:

campaign_name	party_affiliation	registration_rate	election_result
For the People	Independent	91.44	Win
Opportunity for All	Independent	91.32	Win
Standing for You	Democratic	91.05	Loss
Progress for All	Republican	90.90	Loss
Stronger Communities	Republican	90.73	Loss
Change We Need	Independent	90.40	Loss
Moving Forward	Republican	90.36	Loss
A Future We Believe In	Democratic	90.35	Loss
Fairness and Freedom	Democratic	90.31	Win
Your Voice Matters	Democratic	90.30	Loss
Building Bridges	Independent	90.30	Loss

-- Registration rate doesn't necessarily predict success; losing campaigns also had high registration rates

-- I should analyze the turnout rates as well, since a campaign could have high registration but low turnout

WITH voter_party_cte AS

(

SELECT v.campaign_id,

COUNT(*) AS total_voters,

SUM(CASE WHEN v.party_affiliation = 'Democratic' THEN 1 ELSE 0 END) AS democratic_voters,

ROUND(SUM(CASE WHEN v.party_affiliation = 'Democratic' THEN 1 ELSE 0 END) * 100.0 / COUNT(*), 2) AS democratic_percentage

FROM voters AS v

GROUP BY v.campaign_id

)

SELECT c.campaign_name,

c.party_affiliation,

vpc.democratic_percentage,

e.election_result

FROM campaigns AS c

JOIN voter_party_cte AS vpc

    ON c.campaign_id = vpc.campaign_id

JOIN election_results AS e

    ON c.campaign_id = e.campaign_id

ORDER BY vpc.democratic_percentage DESC

;

Result:

campaign_name	party_affiliation	democratic_percentage	election_result
Democracy in Action	Independent	44.97	Loss
A New Path	Independent	43.65	Loss
The People's Voice	Republican	42.91	Win
Fighting for Justice	Republican	42.89	Loss
New Vision for America	Democratic	42.47	Loss
United for a Better Tomorrow	Republican	42.46	Win
Securing Our Future	Independent	42.13	Win
Unite for Change	Republican	41.93	Loss
Empower the Future	Democratic	41.75	Win
Fairness and Freedom	Democratic	41.65	Win
Path to Prosperity	Republican	41.43	Win
Defending Democracy	Republican	41.42	Loss
For the People	Independent	40.94	Win

-- Several winning Republican and Independent campaigns had roughly 40% Democratic voters, showing Democrat crossover may have helped these campaigns win

-- Most interesting: not one Democratic candidate had a voter base of 50%+ Democrats, so non-Dem voters helped these campaigns as well

-- Searching for outliers below

WITH ordered_voters AS (

    SELECT v.campaign_id, v.age,

        ROW_NUMBER() OVER (PARTITION BY v.campaign_id ORDER BY v.age) AS row_num,

        COUNT(*) OVER (PARTITION BY v.campaign_id) AS total_count

```sql
    FROM voters AS v
),
median_voter_age_cte AS (
  SELECT campaign_id,
      age AS median_voter_age
  FROM ordered_voters
  WHERE row_num = CEIL(total_count / 2) -- Get the middle row
)
SELECT c.campaign_name,
    c.party_affiliation,
    mva.median_voter_age,
    e.election_result
FROM campaigns AS c
JOIN median_voter_age_cte AS mva
  ON c.campaign_id = mva.campaign_id
JOIN election_results AS e
  ON c.campaign_id = e.campaign_id
ORDER BY mva.median_voter_age ASC
;
```

Result:

campaign_name	party_affiliation	median_voter_age	election_result
Empower the Future	Democratic	24	Win
Stronger Communities	Republican	24	Loss
Lifting Every Voice	Republican	24	Win
Moving Forward	Republican	24	Loss
Honest Leadership	Democratic	24	Loss
United for Progress	Democratic	38	Loss
Change We Need	Independent	38	Loss
Stronger Together	Democratic	38	Loss
Building Bridges	Independent	38	Loss
Championing Change	Republican	38	Win
Bold Solutions	Democratic	38	Loss
Hope and Progress	Republican	38	Win
Defending Democracy	Republican	38	Loss
A New Day	Independent	38	Loss
For the People	Independent	38	Win
One Nation, One Future	Democratic	39	Win
Opportunity for All	Independent	39	Win
Voices for the People	Independent	61	Win
Securing Our Future	Independent	61	Win
Fairness and Freedom	Democratic	61	Win
Equality Now	Independent	61	Win

-- This shows that young voter bases do not guarantee wins

-- Older voter bases have the highest election success rate

-- Middle-aged campaigns struggled the most

-- Now let's move on to voter turnout analysis


```
WITH voter_turnout_cte AS (

  SELECT v.campaign_id,

      COUNT(*) AS total_voters,

      SUM(CASE WHEN v.voted_in_general_election = 'Yes' THEN 1 ELSE 0 END) AS
actual_voters,

      ROUND(SUM(CASE WHEN v.voted_in_general_election = 'Yes' THEN 1 ELSE 0 END) *
100.0 / COUNT(*), 2) AS turnout_rate

  FROM voters AS v
```

```sql
    GROUP BY v.campaign_id
)
SELECT c.campaign_name,
    c.party_affiliation,
    vtc.turnout_rate,
    e.election_result
FROM campaigns AS c
JOIN voter_turnout_cte AS vtc
    ON c.campaign_id = vtc.campaign_id
JOIN election_results AS e
    ON c.campaign_id = e.campaign_id
ORDER BY vtc.turnout_rate DESC
;
```

Result:

campaign_name	party_affiliation	turnout_rate	election_result
Forward Together	Independent	65.69	Loss
Progress for All	Republican	65.19	Loss
Change We Need	Independent	64.81	Loss
For the People	Independent	64.75	Win
Common Ground	Republican	64.62	Win
A Brighter Future	Independent	64.45	Win
Opportunity for All	Independent	64.21	Win
A New Path	Independent	64.20	Loss

-- This shows higher turnout does correlate with more wins, but it does not always

-- This also shows that winning campaigns rarely had below 62% turnout, which means mobilizing voters mattered

-- Below I will breakout republican and independent percentages among campaigns

WITH voter_party_cte AS (

```sql
    SELECT v.campaign_id,

        COUNT(*) AS total_voters,

        SUM(CASE WHEN v.party_affiliation = 'Republican' THEN 1 ELSE 0 END) AS
republican_voters,

        SUM(CASE WHEN v.party_affiliation = 'Independent' THEN 1 ELSE 0 END) AS
independent_voters,

        ROUND(SUM(CASE WHEN v.party_affiliation = 'Republican' THEN 1 ELSE 0 END) *
100.0 / COUNT(*), 2) AS republican_percentage,

        ROUND(SUM(CASE WHEN v.party_affiliation = 'Independent' THEN 1 ELSE 0 END) *
100.0 / COUNT(*), 2) AS independent_percentage

    FROM voters AS v

    GROUP BY v.campaign_id

)
SELECT c.campaign_name,

    c.party_affiliation,

    vpc.republican_percentage,

    vpc.independent_percentage,

    e.election_result

FROM campaigns AS c

JOIN voter_party_cte AS vpc

    ON c.campaign_id = vpc.campaign_id

JOIN election_results AS e

    ON c.campaign_id = e.campaign_id

ORDER BY vpc.republican_percentage DESC, vpc.independent_percentage DESC

;
```

Result:

campaign_name	party_affiliation	republican_percentage	independent_percentage	election_result
A Future We Believe In	Democratic	37.93	18.74	Loss
Stronger Together	Democratic	37.41	20.82	Loss
Change We Need	Independent	36.72	18.62	Loss
Honest Leadership	Democratic	36.69	21.88	Loss
Hope and Progress	Republican	36.55	20.06	Win
Moving Forward	Republican	36.39	20.32	Loss
Common Ground	Republican	36.22	21.18	Win

-- The biggest takeaway so far is that winning campaigns had strong crossover appeal

-- I will now attempt to combine all factors to determine which is the strongest predictor of election success

-- To do this, I will combine my previous CTEs into one query to analyze which factor has the heaviest weight on success

WITH voter_age_cte AS (

   SELECT campaign_id, age AS median_voter_age

   FROM (

     SELECT v.campaign_id, v.age,

        ROW_NUMBER() OVER (PARTITION BY v.campaign_id ORDER BY v.age) AS row_num,

        COUNT(*) OVER (PARTITION BY v.campaign_id) AS total_count

     FROM voters v -- ✅ Explicitly reference "voters" here

) AS ranked_voters

   WHERE row_num = CEIL(total_count / 2) -- Middle row for median age

),

voter_turnout_cte AS (

   SELECT v.campaign_id,

      COUNT(*) AS total_voters,

```sql
        SUM(CASE WHEN v.voted_in_general_election = 'Yes' THEN 1 ELSE 0 END) AS
actual_voters,

        ROUND(SUM(CASE WHEN v.voted_in_general_election = 'Yes' THEN 1 ELSE 0 END) *
100.0 / COUNT(*), 2) AS turnout_rate

    FROM voters v

    GROUP BY v.campaign_id

),

voter_party_cte AS (

    SELECT v.campaign_id,

        COUNT(*) AS total_voters,

        SUM(CASE WHEN v.party_affiliation = 'Democratic' THEN 1 ELSE 0 END) AS
democratic_voters,

        SUM(CASE WHEN v.party_affiliation = 'Republican' THEN 1 ELSE 0 END) AS
republican_voters,

        SUM(CASE WHEN v.party_affiliation = 'Independent' THEN 1 ELSE 0 END) AS
independent_voters,

        ROUND(SUM(CASE WHEN v.party_affiliation = 'Democratic' THEN 1 ELSE 0 END) *
100.0 / COUNT(*), 2) AS democratic_percentage,

        ROUND(SUM(CASE WHEN v.party_affiliation = 'Republican' THEN 1 ELSE 0 END) *
100.0 / COUNT(*), 2) AS republican_percentage,

        ROUND(SUM(CASE WHEN v.party_affiliation = 'Independent' THEN 1 ELSE 0 END) *
100.0 / COUNT(*), 2) AS independent_percentage

    FROM voters v

    GROUP BY v.campaign_id

)

SELECT c.campaign_name,

    c.party_affiliation,

    va.median_voter_age,
```

```sql
        vt.turnout_rate,

        vp.democratic_percentage,

        vp.republican_percentage,

        vp.independent_percentage,

        e.election_result

FROM campaigns c

JOIN voter_age_cte va

    ON c.campaign_id = va.campaign_id

JOIN voter_turnout_cte vt

    ON c.campaign_id = vt.campaign_id

JOIN voter_party_cte vp

    ON c.campaign_id = vp.campaign_id

JOIN election_results e

    ON c.campaign_id = e.campaign_id

ORDER BY vt.turnout_rate DESC, va.median_voter_age ASC

;
```

Result:

campaign_name	party_affiliation	median_voter_age	turnout_rate	democratic_percentage	republican_percentage	independent_percentage	election_res
Forward Together	Independent	24	65.69	39.07	33.52	21.06	Loss
Progress for All	Republican	23	65.19	39.25	34.93	21.05	Loss
Change We Need	Independent	38	64.81	38.71	36.72	18.62	Loss
For the People	Independent	38	64.75	40.94	35.25	19.50	Win
Common Ground	Republican	24	64.62	38.03	36.22	21.18	Win
A Brighter Future	Independent	23	64.45	39.15	33.86	21.71	Win
Opportunity for All	Independent	39	64.21	39.21	35.61	20.66	Win
A New Path	Independent	23	64.20	43.65	34.48	17.46	Loss
Lifting Every Voice	Republican	24	64.07	40.55	32.20	20.99	Win
Building Bridges	Independent	38	63.79	39.86	34.08	19.06	Loss

-- Number one, we now know that voter turnout is the strongest predictor of election success; nearly all winning campaigns had turnout with 62%+

-- To further illustrate this, not one campaign below 60.7% turnout won their election

-- Median age does matter, but really only for Older voting blocs; Older voting bases are more likely to win

-- Another key point is that Republican/Independent support can flip races, but it is still secondary to turnout

-- The one last thing I really want to explore is what led to higher turnout; fundraising and outreach

```sql
WITH fundraising_summary AS (

    SELECT c.campaign_id,

        c.campaign_name,

        c.party_affiliation,

        c.final_fundraising_amount

    FROM campaigns c

),

outreach_summary AS (

    SELECT o.campaign_id,

        COUNT(*) AS total_outreach -- Count total outreach attempts per campaign

    FROM outreach o

    GROUP BY o.campaign_id

),

event_outreach_summary AS (

    SELECT e.campaign_id,

        COALESCE(SUM(e.doors_knocked), 0) AS doors_knocked,

        COALESCE(SUM(e.calls_made), 0) AS calls_made

    FROM events e

    GROUP BY e.campaign_id

),

voter_turnout_cte AS (

    SELECT v.campaign_id,
```

```sql
        COUNT(*) AS total_voters,

        SUM(CASE WHEN v.voted_in_general_election = 'Yes' THEN 1 ELSE 0 END) AS
actual_voters,

        ROUND(SUM(CASE WHEN v.voted_in_general_election = 'Yes' THEN 1 ELSE 0 END) *
100.0 / COUNT(*), 2) AS turnout_rate

    FROM voters v

    GROUP BY v.campaign_id

)
SELECT f.campaign_name,

    f.party_affiliation,

    f.final_fundraising_amount,

    COALESCE(o.total_outreach, 0) AS total_outreach,

    COALESCE(eo.doors_knocked, 0) AS doors_knocked,

    COALESCE(eo.calls_made, 0) AS calls_made,

    (COALESCE(o.total_outreach, 0) + COALESCE(eo.doors_knocked, 0) +
COALESCE(eo.calls_made, 0)) AS total_voter_contacts,

    vt.turnout_rate,

    e.election_result
FROM fundraising_summary f
JOIN voter_turnout_cte vt

    ON f.campaign_id = vt.campaign_id
LEFT JOIN outreach_summary o

    ON f.campaign_id = o.campaign_id
LEFT JOIN event_outreach_summary eo

    ON f.campaign_id = eo.campaign_id
JOIN election_results e

    ON f.campaign_id = e.campaign_id
```

ORDER BY vt.turnout_rate DESC

;

Result:

campaign_name	party_affiliation	final_fundraising_amount	total_outreach	doors_knocked	calls_made	total_voter_contacts	turnout_rate	election_resu
Forward Together	Independent	1534555.76	532	14305	43785	58622	65.69	Loss
Progress for All	Republican	1668298.48	539	9670	13344	23553	65.19	Loss
Change We Need	Independent	1646712.41	547	16748	16933	34228	64.81	Loss
For the People	Independent	1818652.14	577	4683	72774	78034	64.75	Win
Common Ground	Republican	972246.14	542	9685	20334	30561	64.62	Win
A Brighter Future	Independent	1646712.41	539	11238	48882	60659	64.45	Win
Opportunity for All	Independent	4047458.96	532	16562	30520	47614	64.21	Win
A New Path	Independent	1818652.14	539	7442	44544	52525	64.20	Loss

-- Outreach drives turnout more than fundraising

-- Campaigns that made more direct voter contact had higher turnout, regardless of how much money was raised

-- Importantly, fundraising alone was not enough - not spending enough money on outreach may have affected turnout

-- The strongest performing campaigns combined fundraising and outreach to drive higher voter turnout and election success