
Supervised Learning Analysis (Prediction): Vote Margin per District – Sragen Local Election 2024

Pipin Tri Hastuti^{1*}, Nandita Sekar Sukma Dewi², Falih Walen Satria Sakti³, Dimas Akbar Pramudya⁴

^{1,2,3,4,5}Duta Bangsa University, Bhayangkara Street No. 55-57, Tipes, Serengan District, Surakarta City, Central Java 57154., Indonesia

Keywords

Information; system; technology; software

*Correspondence Email:

220103032@mhs.udb.ac.id

Abstract

This study examines whether the difference in votes between candidate pairs in the 2024 Sragen Regional Election can be predicted using basic administrative data at the sub-district level. The analysis was carried out quantitatively with analysis units of 20 sub-districts in Sragen Regency. For each sub-district, the composition of voters (number of male voters, number of female voters, total Permanent Voter List), voter participation (participation rate, calculated as the number of valid votes divided by the total registered voters), and electoral results (difference in votes, defined as the votes of Candidate 1 minus the votes of Candidate 2). The purpose of this study is to explain why some sub-districts produce a very large difference in votes for one candidate (strong base), while other sub-districts are very tight and only differ by hundreds of votes (swing areas). The results of the study show three main patterns. First, sub-districts with large numbers of voters and high participation tend to produce wide vote differences for dominant candidates. Second, sub-districts with a difference of only hundreds of votes can be identified as competitive areas that can easily change direction if there is a slight change in voter mobilization. Third, the multiple linear regression model with variables of total registered voters, participation rate, and proportion of male voters was able to explain about 38% of the variation in vote difference between sub-districts. These findings show that the difference in votes is not just the final result of the count, but is related to the ability to mobilize voters in densely populated areas as well as to the demographic structure of voters in those areas.

1. Introduction

In the election of regional heads (Pilkada), public discussion usually stops at who the winner is. In fact, in terms of political strategy, what is equally important is How big is the difference in votes between candidate pairs in each sub-district. A large difference in votes indicates a relatively safe support base, while a narrow difference indicates a swing area that is prone to change direction in the event of a change in voter mobilization (Collins et al., 2021). This means that the margin of victory itself can be read as an indicator of local political strength.

This research departs from the main question: Can the difference in votes between candidate pairs at the sub-district level be estimated using basic data that is already available administratively? Specifically, this study examined two factors: (1) the demographic structure of permanent voters (number of male voters, number of female voters, and total Permanent Voter Lists/DPT) and (2) voter participation (how many voters actually attended and exercised their right to vote). In the practice of holding regional elections in Indonesia, the voter participation rate is not only considered as an attendance figure, but also as a measure of citizen involvement and the legitimacy of election results (Asgar, 2023).

This participation issue is important because the trend is not always stable. In the 2020 Simultaneous Regional Elections, which were held in the midst of the COVID-19 pandemic, the national participation rate was reported to be around 76 percent and was hailed as a high achievement. However, regional political observers noted that in a number of regions ahead of the 2024 Regional Elections, participation tends to fall to the range of 60 percent and this is seen as potentially weakening the political legitimacy of the Regional Election results. In other words, turnout is not just a technical matter of "how many people come", but also related to how strongly the winner's mandate is perceived by the public.

In addition to participation, the demographic composition of voters also affects support patterns. A study of political behavior in Indonesia found that differences in the composition of male and female voters, women's political involvement, and local socio-economic structures also affect the intensity of vote competition between candidates (Bare'e, 2025). This finding indicates that the character of voters in a sub-district, not only the number of voters, can play a role in determining whether the difference in votes in the sub-district is large or very thin.

Based on this background, this study uses Sragen Regency in the 2024 Regional Elections as a case study. The unit of analysis is the sub-district. The dependent variable is the difference in votes between candidate pairs (the difference between Candidate A and Candidate B in each sub-district). Independent variables include the size of the voter base (total DPT and male/female voters) and voter participation rate (turnout). The analysis was carried out using a quantitative approach using multiple linear regression at the sub-district level. The objectives of this study are: (1) to map the pattern of vote difference between sub-districts in Sragen, (2) to analyze the relationship between demographic factors and participation with vote differences, and (3) to develop a simple prediction model to estimate vote difference based on these factors.

1.1 Literature Review

1.1.1 Description (concept, research position, and gap)

The study of voter behavior at the local level basically does not only discuss candidate preferences, but also the dynamics of vote differences between candidates in one region. A large margin of victory in a sub-district is usually read as an indication that the region is a safe base for a particular candidate, while a thin margin indicates a competitive or swing region, which is an area that can still change its direction of support in the event of a change in campaign strategy or voter mobilization. Contemporary research in Indonesia links this variation in vote margins to two main factors: (1) the socio-demographic structure of voters in the region, and (2) voter turnout, specifically how many registered voters actually turn out at polling stations and cast their votes. The level of participation itself is currently treated not only as an administrative figure, but as an indicator of the legitimacy of the results of the Regional Elections and the depth of political involvement of citizens. Election organizers and local political observers in recent years have openly linked voter participation to the strength of the political mandate of elected regional heads, especially after concerns emerged that declining participation in a number of regions in the 2024 regional elections is said to be below 70% and is considered to threaten the quality of local democracy.

This literature is important because it shows that the vote margin cannot be separated from two things that are actually available as routine administrative data: the composition of the electorate (e.g. the proportion of men and women on the Permanent Voter List/DPT) and the participation rate. However, most studies still focus on voting behavior (e.g. who is elected) or on participation (whether people come to vote), not on how far a candidate wins in each sub-district. This means that there is a gap: there have not been many studies that

explicitly try to predict the difference in votes between candidate pairs at the sub-district level using only basic demographic data and voter participation that are already available to the Pilkada organizers.

1.2.1 Methodology (approaches to previous studies and their relevance to this study)

In general, research on electoral political behavior in Indonesia in the last 3–5 years has mostly used two quantitative approaches that are relevant to this study. The first approach is area-based analysis (ecological/aggregate analysis), which is to read the relationship between the characteristics of a region and the pattern of political participation and support in that region. For example, research on Indonesia's middle class shows that structural factors such as the level of modernization and socio-economic composition of regions affect voter participation, and that more economically "developed" regions do not necessarily have higher turnout. The analysis was carried out by connecting regional socio-economic indicators with voter participation data, then tested using a regression model to see the extent to which the social composition of the community explains the difference in voter turnout between regions.

The second approach is a demographic analysis of voters, especially those that highlight gender (Persepsi et al., 2024). Some recent studies emphasize that female voters in Indonesia are not only the majority numerical group (more than half of voters are registered nationally), but also tend to have very active participation rates, even in some national and local elections recorded slightly higher than men. This study generally uses large-scale survey data and/or election administrative data, then compares the political involvement of men and women, including willingness to attend polling stations and voting consistency. The two methodological currents of aggregate area analysis and demographic composition analysis are very relevant to this study. The study followed a similar pattern: the unit of analysis was the sub-district (not the individual), the independent variable was administratively measurable data (DPT size, male/female composition, voter participation rate), and the bound variable was the difference in votes between candidate pairs in the sub-district. Thus, this study positions itself as an adaptation of the well-established electoral quantitative method, but directed directly to the strategic question of "what is the margin of victory of a candidate in a sub-district," not just "who is elected" or "who comes to vote."

1.3.1 Key results from previous research (empirical findings) and implications

The results of previous studies provide several important patterns that form the basis of this research hypothesis. First, the level of voter participation is proven to be seen directly as a measure of the political legitimacy of regional candidates. In the 2020 Simultaneous Regional Elections, the Indonesian Election Commission reported an average participation of around 76.09%, and this figure was officially claimed as a high achievement even though the election took place in pandemic conditions in hundreds of regions (East Lampung Health Office, 2020). However, entering the 2024 Regional Elections, regional political experts began to mention that in a number of regions participation dropped below 70% (Dzulfian Syafrian, 2025) and called this trend a threat to the quality of local democracy because the political mandate of regional heads has become seen as "narrower," as if it only comes from the most loyal voters. Second, in terms of demographics, women's political participation in Indonesia remains high and consistent; women make up more than half of the national electorate and their attendance at polling stations in some major elections is reported to be slightly above that of men. This shows that the gender structure of voters has the potential to affect the aggregate results of the regions, including the difference in votes between candidates at the sub-district level. Third, there is evidence that regional socio-economic factors such as the size of the middle class and the level of modernization are related to participation patterns, but not always linearly ("more modern" regions do not automatically have higher participation. Taken together, the findings imply that the variation in vote difference between sub-districts is likely to be determined not purely by political preferences, but also by: (a) how large the number of voters in the region actually turned out to vote, and (b) what the demographic makeup of the voters looked like. A gap that has not been widely answered by previous studies is whether this combination of participation and demographic factors can be used directly to predict the width of the vote difference between candidate pairs at the sub-district level in the context of certain regional elections, in this case, Sragen Regency in 2024.

2. Research Methods

2.1 Research Design

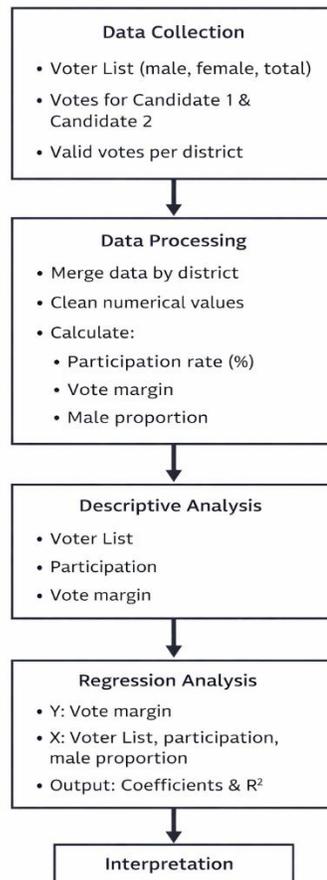


Figure 1. Research Flow

This study uses a quantitative approach based on regional aggregate data. This means that the analysis is not carried out at the individual voter level, but at the sub-district level. Each sub-district in Sragen Regency is treated as an analysis unit that has certain demographic characteristics, a certain level of voter participation, and a certain vote difference.

The purpose of this design is to answer the main question of the study: whether regional characteristics (voter composition and voter participation) can be used to estimate how large the vote difference between candidate pairs in the regional elections will be. This approach is explanatory, not just descriptive. So it's not just "sub-district A has a certain percentage of participation", but also "is the high participation related to a wider vote difference?" This explanatory approach is commonly used to test the relationship between local political variables in Indonesia, especially when researchers want to look at the contribution of several factors at once to electoral results

Broadly speaking, the flow of the research method can be visualized as a flowchart containing the following steps:

1. Collect administrative data for the Regional Elections per sub-district (DPT, number of votes, participation).
2. Clean and consolidate data into one master table per sub-district.
3. Calculates derivative variables such as participation rates and vote differences.

4. Conducting a descriptive analysis to see the basic pattern of each sub-district.
5. Run multiple linear regression analysis to test interfactor relationships.
6. Interpret the results to answer the research objectives.

2.2 Unit Analysis, Population, and Sampling Approach

a. unit analysis.

The unit of analysis in this study is the sub-district. One line of data = one sub-district.

b. Research population.

The population in the context of this study is all sub-districts in Sragen Regency that participated in the 2024 Regional Elections and has complete data regarding:

- the number of male voters in the Permanent Voter List (DPT),
- the number of female voters in the DPT,
- total DPT,
- the number of valid votes for each pair of candidates,
- total valid votes,
- voter turnout,
- difference in votes between two candidates.

There are 20 sub-districts that meet these criteria, so the number of observations = 20.

c. Sampling Technique.

This study did not randomly sample. All sub-districts were analyzed. In other words, the approach is not a "sample survey" but a **total enumeration** (regional census). This makes sense because the number of sub-districts is relatively small and all the data is available. So this research no longer needs probabilistic sampling techniques such as simple random sampling.

d. Territorial context.

The reason for choosing Sragen Regency as a case study is because the inter-districts in this area are not homogeneous. There are sub-districts with a very large number of voters and a difference of thousands (a strong base of one candidate), while other sub-districts have a smaller number of voters and a very thin vote difference (swing areas). Internal variations like this make Sragen suitable as a "small laboratory" to read the relationship between regional characters, participation rates, and victory margins.

2.3 Data Collection Techniques

The data used is administrative data for the Regional Elections at the sub-district level. This data includes:

1. Number of registered male voters (DPT male).
2. Number of registered female voters (female DPT).
3. Total DPT (total number of registered voters).
4. Obtaining valid votes for candidate pairs 1.
5. Obtaining valid votes for candidate pairs 2.
6. Total valid votes (the sum of the valid votes of the two candidate pairs).
7. Voter participation rate (how many registered voters actually attended and voted).
8. The difference in votes between candidate pairs (candidate 1's vote minus candidate 2's vote, can be negative).

The data processing steps are carried out as follows:

- Merge the DPT table per sub-district with the table of vote results per sub-district based on the name of the sub-district.
- Cleanup column naming, converting the value that was originally text to numerical ones, and ensuring numerical consistency (e.g. total valid votes do not exceed the total DPT).
- Calculates new derivative variables that are indirectly present in the raw table, such as voter participation rate (%) and proportion of male voters.

2.4 Definition of Variable Operational (Measures)

This subsection explains how each variable is defined and measured, so that it is clear what "participation", "vote difference", and so on are.

(1) Total DPT (Total Permanent Voter List)

- Operational definition: the number of all voters who are officially registered in the sub-district in the 2024 Regional Elections. This is the number of male voters + female voters.
- Unit: An absolute number (integer).
- Analytical function: shows the "large voter base" of the sub-district. Districts with large DPT are considered to have greater political weight because they contribute a large number of votes.

(2) DPT Male and DPT Female

- Operational definition: the number of male voters and the number of female voters registered in the DPT in the sub-district.
- From these two values, the proportion of males ($\text{proporsi_lk} = \text{male DPT} / \text{total DPT}$) is also calculated.
- Important reasons: the gender structure of voters is considered relevant because some studies show that women's political behavior in Indonesia is not passive; Female voters often have high and consistent participation rates. Thus, the gender composition of the region may also affect the difference in votes.

(3) Voter Participation Rate (Turnout Rate)

- Operational definition: the number of valid votes cast in the sub-district divided by the total DPT of that sub-district, then expressed in percentages.
- This represents how many citizens are not only registered as voters, but actually present and vote.
- Why this matters: the level of voter participation in regional elections in Indonesia in recent years has been used politically as an indicator of the legitimacy and strength of the mandate of regional heads. In the 2020 Simultaneous Regional Elections, for example, the participation rate of around 76% was claimed to be a high achievement even though the election took place in the midst of a pandemic. However, ahead of the 2024 Regional Elections, many parties are starting to worry because there are regions with participation falling to the range of 60 percent, which is considered to weaken the legitimacy of the Regional Election results

(4) Difference of Votes Between Candidate Pairs (Margin of Victory)

- Operational definition: the number of valid votes of candidate pair 1 minus the number of valid votes of candidate pair 2, for each sub-district.
- The value can be negative. A negative score means that the candidate pair 2 is superior.
- The greater the negative value (e.g. -9,000, -10,000), the stronger the dominance of candidate pair 2 in the sub-district.
- A value close to zero (e.g. -163, -246) indicates a swing district: an area where the fighting is tight and can still change direction.
- This variable is treated as a dependent variable (variable Y) in the statistical model, because this is the "outcome to be described": how wide the victory is in the district.

2.5 Data Analysis Procedure

Data analysis is carried out in two major stages:

Stage 1 – Descriptive Analysis.

This stage is used to describe the basic pattern per sub-district:

- how large is the voter base (Total DPT),
- how high is voter participation (% turnout),
- how big or small the difference in votes between candidate pairs.

Stage 2 – Multiple Linear Regression Analysis.

This stage is used to test whether the factors of the region can predict the magnitude of the difference in votes. The model used is an OLS (Ordinary Least Squares) regression with:

- Dependent variable: Vote Difference (Candidate 1 – Candidate 2).
- Independent variables: Total DPT, Voter Turnout Rate, and Proportion of Male Voters.

2.6 Reliability, Validity, and Replication

This section explains why the method used is trustworthy.

1. Data reliability.

The data used is official administrative data at the sub-district level, not individual opinion survey data. Due to its administrative nature, the numbers can be double-checked and consistent between documents.

2. Measurement validity.

Each variable is operationally defined. For example, "participation rate" is described in an explicit formula, not an assumption. "Vote difference" is also clearly defined (Candidate 1 – Candidate 2). This is important so that readers understand what each term means and not misinterpret it.

Replication.

The analysis steps (combining the DPT table and the vote table, calculating turnout, calculating the vote difference, running multiple linear regression) are described in sequence. Other researchers can apply the same procedure to the data of other district elections, or to the following year's regional elections. The methodological guidelines emphasize that methods must be described in sufficient detail so that other researchers can assess the feasibility of the method and be able to repeat it.

3. Result and Discussion

3.1 Statistics Descriptive

Descriptive statistics are used to understand the general picture of voter character and vote results at the sub-district level. A summary of the main variables is presented in Table 1 (Summary of District Statistics), which contains:

- GET MALE,
- DPT female,
- total DPT per sub-district,
- voter turnout rate (in percent),
- the difference in votes between candidate pairs (Candidate 1 – Candidate 2).

The summary results are as follows:

1. Size of the voter base (Total DPT).
 - The average number of registered voters (DPT) per sub-district is around 38,115 voters.
 - The sub-district with the most voters has 59,638 registered voters.
 - The sub-district with the least voters has 18,058 registered voters. This means that there are quite extreme differences in the scale of the sub-districts: the largest sub-districts are almost 3× larger than the smallest sub-districts. Politically, large sub-districts like this have a much greater weight of votes for the final result at the district level.
2. Voter participation (turnout). The voter participation rate is calculated as the number of valid votes divided by the total DPT, then expressed in percent.
 - The average voter participation rate between sub-districts is around 76.14%.
 - The lowest participation was recorded at around 72.15%.
 - The highest participation was close to 80.06%. In other words, in the sub-districts with the highest participation, about 4 out of 5 registered voters actually came and cast valid votes. In the national context, the participation rate of around 75-80% is seen as a "quite high" level of participation in the discourse of the Regional Elections in recent years, and has even been used as a benchmark for the success of the implementation of the 2020 Simultaneous Regional Elections by election organizers.

3. The difference in votes between the candidate pairs. The main variable of this study is the difference in votes per sub-district, calculated as the votes of Candidate 1 minus the votes of Candidate 2. A negative score means that Candidate 2 is superior.
- The average difference in votes is around -3,909 votes per sub-district. This means that in general, Candidate 2 is quite far ahead in many sub-districts.
 - The largest difference in votes was around -10,840 votes (Candidate 2 won very largely).
 - The smallest difference in votes was only -163 votes (almost a tie). This shows that not all sub-districts are "won" by the same distance. There are sub-districts with landslide victories, and there are sub-districts that are actually almost drawn.

In terms of reporting, these figures are in Table 3.1

Table 1. Total Population of Each Sub-district

NO	District	Man	Woman	Sum	Male Percentage	Percentage of Women
1	Kaliambe	20.675	20.465	41.140	50,26%	49,74%
2	Plupuh	19.667	19.851	39.518	49,77%	50,23%
3	Masaran	29.536	30.102	59.638	49,53%	50,47%
4	Kedawung	25.504	26.041	51.545	49,48%	50,52%
5	Sambirejo	15.782	16.085	31.867	49,52%	50,48%
6	Gondang	18.090	18.656	36.746	49,23%	50,77%
7	Sambungmacan	18.845	19.359	38.204	49,33%	50,67%
8	Ngrampal	16.220	16.978	33.198	48,86%	51,14%
9	Karangmalang	26.945	28.451	55.396	48,64%	51,36%
10	Sragen	26.448	28.188	54.636	48,41%	51,59%
11	Sidoharjo	22.184	22.954	45.138	49,15%	50,85%
12	Tanon	22.772	23.289	46.061	49,44%	50,56%
13	Squirt	19.696	20.342	40.038	49,19%	50,81%
14	Miri	14.148	14.412	28.560	49,54%	50,46%
15	Sumberlawang	19.019	19.638	38.657	49,20%	50,80%
16	Mondokan	15.450	15.410	30.860	50,06%	49,94%
17	Sukodono	13.084	13.478	26.562	49,26%	50,74%
18	Gesi	8.889	9.169	18.058	49,22%	50,78%
19	Seaweed	11.472	11.457	22.929	50,03%	49,97%
20	Jenar	11.681	11.878	23.559	49,58%	50,42%
	Kabupaten Sragen	376.107	386.203	762.310		

3.2 Base Pattern vs Swing Region (Sound Difference Visualization)

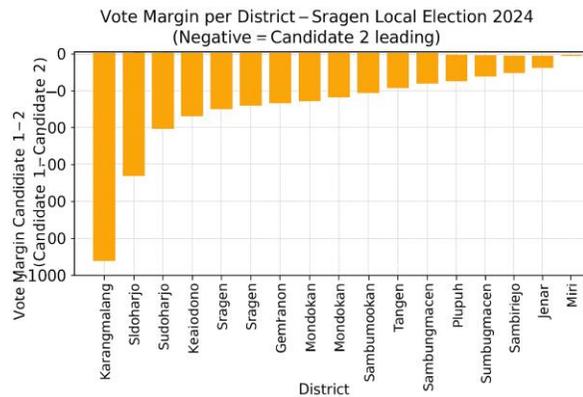


Figure 2. Difference in Votes for Each District

To read the pattern of political competition between sub-districts, the difference in votes between candidate pairs is plotted in the form of a bar chart (**Figure 3.1**). The X-axis is the name of the sub-district, the Y-axis is the difference in votes (Candidate 1 – Candidate 2). The sub-districts are then sorted from the largest difference (absolute widest) to the smallest difference (thinnest).

From this graph, two politically distinct sub-districts emerge:

1. The base sub-district is strong. This is a sub-district that shows a difference of thousands to more than ten thousand votes, all leading to the advantage of Candidate 2 (because the difference is a big negative). Examples:
 - o Karangmalang ($\approx -10,840$ votes),
 - o Sidoarjo ($\approx -9,678$ votes),
 - o Sukodono ($\approx -6,643$ votes),
 - o Kedawung ($\approx -6,632$ votes),
 - o Masaran ($\approx -4,993$ votes).

These sub-districts can be referred to as "voting barns" for Candidate 2. In such a region, electoral dominance is relatively stable: the margin of victory is too large to be reversed with just a slight change in preference.

2. Swing / competitive sub-district. This is a sub-district with a very small difference in votes, only hundreds or thousands of votes. For example:
 - o Miri (≈ -163 votes),
 - o Jenar (≈ -246 votes),
 - o Sambirejo ($\approx -1,129$ votes),
 - o Sumberlawang ($\approx -1,465$ votes).

A sub-district like this is a contested area. A change in voter participation or a shift in support for a small group can immediately overturn the local winner.

The political interpretation: the district-level victory is not "an equal victory in all districts." The pattern is more like this: there are several large pockets that give landslide victories (bases), plus tight sub-districts that have been successfully "secured". Visualizations like this also help identify future risk points: if one day the base sub-district starts to decline in participation, the district margin as a whole could be eroded.

3.3 Multiple Linear Regression Analysis

This section reports on inferential statistical testing, i.e. multiple linear regression (OLS). This analysis answers the core question of the study: can voter demographic factors and voter participation rates predict how big the vote difference between candidate pairs at the sub-district level is?

Table 2. Classification Results

District	Untung Wibowo Sukowati, B.Comm. – Drs. Suwandi, M.M. (Candidate 1)	Sigit Pamungkas, S.IP., M.A. – Suroto (Candidate 2)	Sum	Participation Ratio	Difference	Classification Labels
Kaliambe	14.907	16.964	31.871	1,290829	-2.057	0
Plupuh	14.388	16.827	31.215	1,265994	-2.439	0
Masaran	21.377	26.370	47.747	1,249042	-4.993	0
Kedawung	16.628	23.260	39.888	1,292243	-6.632	0
Sambirejo	11.235	12.364	23.599	1,350354	-1.129	0
Gondang	11.681	15.585	27.266	1,347686	-3.904	0
Sambungmacan	13.128	15.775	28.903	1,321801	-2.647	0
Ngrampal	10.711	14.709	25.420	1,30598	-3.998	0
Karangmalang	16.647	27.487	44.134	1,255177	-10.840	0
Sragen	18.659	23.027	41.686	1,310656	-4.368	0
Sidoharjo	12.599	22.277	34.876	1,294242	-9.678	0
Tanon	15.439	19.234	34.673	1,32844	-3.795	0
Squirt	13.503	17.618	31.121	1,286527	-4.115	0
Miri	11.011	11.174	22.185	1,287356	-163	0
Sumberlawang	13.726	15.191	28.917	1,336826	-1.465	0
Mondokan	9.518	12.747	22.265	1,386032	-3.229	0
Sukodono	6.397	13.040	19.437	1,366569	-6.643	0
Gesi	5.360	8.062	13.422	1,345403	-2.702	0
Seaweed	7.109	10.253	17.362	1,320643	-3.144	0
Jenar	8.620	8.866	17.486	1,347306	-246	0
Kabupaten Sragen	252.643	330.830	583.473			

3.3.1 Model specifications

The models used are:

$$\begin{aligned} \text{Vote Difference} = & \beta_0 \\ & + \beta_1 \cdot (\text{Total DPT}) \\ & + \beta_2 \cdot (\text{Voter Participation Rate}) \\ & + \beta_3 \cdot (\text{Proportion of Male Voters}) \\ & + \varepsilon \end{aligned}$$

Description of variables:

- Dependent variable (Y):
The difference in votes between candidate pairs in the sub-district (Candidate 1 – Candidate 2). A negative score means that Candidate 2 is superior.
- Independent variable (X):
 - Total DPT: the size of the sub-district voter base.

- Voter Turnout Rate: what percentage of registered voters actually turned out to vote.
- Proportion of Male Voters: the percentage of male voters in the sub-district DPT.

3.3.2 Regression results

The main results of the regression (reported in Table 2) can be summarized as follows:

- a. The Total DPT coefficient is negative (around -0.098). This means that the greater the number of registered voters in one sub-district, the deeper the vote difference in favor of the dominant candidate. Substantively: large sub-districts tend to be polling booths that widen the margin of victory.
- b. The Voter Participation Rate Coefficient is also negative (around -5302). This shows that when participation increases (e.g. from 74% to 75%), the vote gap tends to get "deeper" (Candidate 2 is further ahead). This means: mobilization to the polling station not only increases the number of voters present in general, but tends to strengthen candidates who are already dominant in the region. This is consistent with the view that voter turnout is often used as a symbol of the strength of local political mandates
- c. The coefficient of the Proportion of Male Voters is positive (around +191,599). This means that the higher the proportion of male voters in the DPT of a sub-district, the difference in votes moves towards a more balanced (not too deep) direction. In other words, sub-districts with a relatively "more masculine" composition of voters tend to be more competitive. This is in line with the finding that female voters in Indonesia generally have high and consistent levels of political involvement, so areas with strong female voter bases can provide a wider margin of victory for certain candidates
- d. The R-squared value of the model is around 0.378. This means that about 37.8% of the variation in vote difference between sub-districts can be explained by the three quantitative factors above. The rest (about 62.2%) are likely to come from other factors not included in the model, such as the candidate's personal image, the strength of the local network, village patronage, very specific public service issues in a particular sub-district, and the factor of the emotional proximity of the voter to the candidate's figure.

3.4 Discussion

1) **Voter participation widens the margin of victory.**

The regression results showed that sub-districts with higher participation rates tended to produce larger vote differences for dominant candidates. This means that turnout is not neutral. When the political machine of a pair of candidates is able to encourage its voters to come to the polling station, the margin of victory not only increases, but increases disproportionately. This supports the view that voter participation is now understood as a source of political legitimacy of regional heads, not just technical data on attendance. In practical terms, this also means that the strategy of winning districts is not only "convincing people to vote for us," but also "making sure they are actually present to vote."

2) **Demographic structure has an electoral effect.**

The proportion of male voters was associated with a smaller (tighter) vote margin, while sub-districts with a relatively larger proportion of female voters tended to show a wider difference. This is in line with the findings that female voters in Indonesia are not passive; in many contests they are consistently present and often part of the candidate's support base (**Setyonaluri & Prihatini, 2024**). In other words, the gender composition of voters is not only administrative data, but also a signal of the political character of the sub-district.

3) **Large sub-districts are margin machines.**

Districts with large DPT not only contribute a large absolute number of votes, but also tend to become "barns of difference". Therefore, victory at the district level as a whole is determined not only by "who excels in how many districts," but especially by "how deeply superior he is in large districts." This is clearly seen in sub-districts such as Karangmalang and Sidoharjo, which contribute to a difference of thousands to close to five digits. This situation fits with the local political logic: secure a large base, then survive in competitive districts.

4) **The swing region is fragile but crucial.**

Districts such as Miri and Jenar have a difference of only hundreds of votes. Regions like this are very volatile just because of a slight shift in issues or changes in the turnout of a particular group. From the point of view of electoral strategy, swing districts are locations where face-to-face campaigns, informal networks of RT/RW, and local figures (religious leaders, traditional leaders, community heads) can really determine the final outcome.

5) **There are still other factors beyond the model.**

An R-squared value of about 0.38 indicates that demographics and participation explain some of the variation in vote differences, but not all. This is natural. In Indonesian local politics, the personal factors of the candidates, the image of the incumbent, patronage relations, promises of certain infrastructure development, or very specific public service issues on a sub-district-by-district basis often play a strong role (Utami et al., 2024). That is, this model is useful quantitatively, but it should not be read as the only explanation.

4. Conclusions

This study shows that the difference in votes between candidate pairs at the sub-district level in the 2024 Sragen Regional Election does not occur randomly. Large differences tend to appear in sub-districts with a large number of voters (high DPT) and high voter participation. This means that the densely populated areas that have been successfully mobilized are the main source of victory margins for the dominant candidate. This is in line with the view that voter participation is not only a matter of attendance at polling stations, but also becomes the political legitimacy capital of elected regional heads.

The demographic structure of voters also has an effect. Districts with a relatively large composition of female voters tend to produce a wider vote gap, while sub-districts with a higher proportion of male voters tend to be more competitive (small vote difference). This is consistent with the finding that women voters in Indonesia are actively participating politically and can be a stable base of support for certain candidates.

Statistically, the multiple linear regression model used in this study was able to explain about 38% of the variation in vote difference between sub-districts with only three simple variables: the size of the DPT, the voter participation rate, and the proportion of male voters. This means that basic administrative data is sufficient to read a map of local political power, although other factors such as the figure of the candidate, local networks, public service issues, and patronage still play a role.

A brief suggestion for the next study: test this model in other areas, add local political context factors (e.g. the strength of the incumbent candidate, the village campaign machinery), and combine quantitative analysis with field interviews. That way, we not only know "who won," but also "why they won to such an extent in a certain sub-district," and what it means for the legitimacy of the next local government.

5. References

- Asgar, S. (2023). Public Political Participation in Elections and Regional Elections as the Basis for the Legitimacy of Government Power. *Innovative: Journal Of Social Science Research*, 3, 5982–5994. <http://j-innovative.org/index.php/Innovative/article/view/2802%0Ahttps://j-innovative.org/index.php/Innovative/article/download/2802/1996>
- Bare'e, Y. J. (2025). WOMEN'S POLITICAL PARTICIPATION IN VILLAGE HEAD ELECTIONS:(Analysis of Involvement in Voting, Campaigns, Political Discussions, and Cultural Constraints and. *Journal of Politics*, 3(1), 45–58. <https://jp.iregway.com/index.php/polahi/article/view/32>
- Collins, S. P., Storrow, A., Liu, D., Jenkins, C. A., Miller, K. F., Kampe, C., & Butler, J. (2021). *No Title 済無No Title No Title No Title*. 2024, 167–186.
- East Lampung Health Office. (2020). *Health Profile of East Lampung in 2020*.
- Dzulfian Syafrian, dkk. (2025). No 主観的健康感を中心とした在宅高齢者における健康関連指標に関する共

分散構造分析Title. In *Sustainability (Switzerland)* (Vol. 11, Issue 1).
http://scioteca.caf.com/bitstream/handle/123456789/1091/RED2017-Eng-8ene.pdf?sequence=12&isAllowed=y%0Ahttp://dx.doi.org/10.1016/j.regsciurbeco.2008.06.005%0Ahttps://www.researchgate.net/publication/305320484_SISTEM_PEMBETUNGAN_TERPUSAT_STRATEGI_MELESTARI

Perception, A., Against, P., Calon, P., Sari, P. E., Mariyanti, E., & Putri, S. L. (2024). White 2024. *Journal of Economics and Business of Dharma Andalas*, 26(1), 216–234.