Determine Sample Size for Precision Results on Quick Count

Y Ridwan[1,2], **R H S Aji**[3]

[1]Institute of Teacher Training Siliwangi, Bandung
[2]Swamedia Reasearch and Communication, Jakarta
[3]Faculty of Economics and Business, UIN Syarif Hidayatullah, Jakarta

*Corresponding author's e-mail: yusepridwan@gmail.com

Abstract. This research aims to answer the problem of the appropriate sample size in the case of the quick count of the election so that the results obtained are close to the actual results. Although there are practical procedures that are widely used to calculate the sample size in the quick count methodology, in reality, the results obtained often deviate from the actual results, so the issue of precision is always an interesting discussion. The formulation of the problem regarding the size of the sample and how the level of precision of the forecast results are important issue to be discussed. This research method is included in experimental research where the analysis used is the Kruskal-Wallis test. The data used is primary data from the real count results of the regency election Sumedang by consultants and teams. The results showed that there was a significant difference between the seven sample size groups in vote acquisition and the percentage of votes at the polling station (TPS), where the sample sizes n =408, n=500, n=875 and n=1674 were the most appropriate sample sizes in the implementation of the quick count.

1. Introduction

Quick Count (QC, quick count) is a way that is provided to find out the results of the vote in the implementation of presidential elections or regional head elections. *Quick Count* has the main function as a control tool for the results of manual calculations by election organizers when performing manual calculations on the final result. As a method of estimating early calculations, *quick count* is believed to be able to provide estimates and detect early in preventing manual miscalculations, calculation irregularities or disclosing vote calculation fraud, so it is believed that the QC will encourage the results of a calculation in an honest and fair general election (Kismiantini, 2007). From the function of good QC control, but along with it often the result of his estimate becomes a verdict of the claim of victory by the contestants before the announcement of the final results of the recapitulation of official votes issued by the general election. This results in conflict between contestants, and public distrust of the general election, if there is a difference in results, especially the difference that occurs is not significant enough between the results of the QC and the results of manual calculations by the organizers.

There are several cases around differences in the results of official calculations from the election organizers, namely the General Election Commission (KPU), such as the case in the 2014 presidential election. No less than 11 pollsters participated in enlivening the 2014 Presidential Election by conducting a quick count (Susanto, 2019). Some of the names of the pollsters suddenly came to the attention of the public and were widely discussed in the mass media because they released different QC results data and *claimed* the victory of one of the presidential candidates by certifying the results of the QC conducted by his institution is the most accurate (precision). Call it the difference in QC results from 4 QC organizers, announcing that the Prabowo Subianto - Hatta Rajasa got a vote advantage over their competitor, Joko Widodo-Yusuf Kala. But the QC results from other QC organizers also released different results, there were at least 7 other QC institutions that announced the superiority of presidential candidate Joko Widodo – Jusuf Kala in the appeal with presidential candidate Prabowo Subianto-Hatta Rajasa. The difference in the results of QC predictions by various institutions in the presidential elections made a stir in the community.

Another example of QC results that resulted in noise in the community included the election of the regional head (pilkada) of Tasikmalaya regency (CNN Indonesia, 2020), a well-known Survey Institute announced the results of the QC where candidate number 4 was superior to the candidate number 2 with a margin 2.7%. But in fact, the results of the recapitulation of the vote of the Regional Election Commission (KPUD) of Tasikmalaya Regency, candidate number 2 narrowly superior to candidate number 4 by a margin of 0.7 percent. This condition makes noise in the community, especially for supporters who have been celebrating victory, but ultimately have to accept defeat. Academics and democracy activists, thus impacting suspicion of the methodology on QC sampling due to differences in prediction results by QC machines with official results published by KPUD.

The more information about transparency about the background of QC implementation, the more questions arise from the public about scientific procedures in conducting QC activities. Especially regarding sampling methodology or known as *sampling technique.* From the aspect *of human error* is also questioned, who is the person behind the pollsters. Reflecting on the Presidential election in 2014 and the Tasikmalaya Regency Election in 2020 and occurred also in several other regions, the institution that held the QC has published the results of a quick count of election results so that when using the correct research rules (scientific methodology), there are many errors in results.

The QC method is based on direct observation at a randomly selected polling station (TPS), where the polling station becomes a unit of analysis, so sampling cannot be done before registering for availability from the polling station or village to be monitored (Kismianti, 2007). The true strength of the quick count data depends on how the sample is taken. The sample will determine the voter's vote to be used as a basis for predicting the outcome of the election. QC predictions will be accurate if they meet the requirements of precise and rigorous statistical and sampling methods so that they will describe or represent actual population characteristics.

Sampling techniques are often needed in research because it is not possible to collect all data from each population unit that we take (Kumar et al., 2013; Now, 2003). Therefore, determining the appropriate sample size is very important to draw valid conclusions from the research findings. However, it is often considered a difficult step in empirical research design (Dattalo, 2008). While there are several tables and rules of thumb for calculating sample sizes in social science research, many researchers are still unclear about which they should use to determine the appropriate sample size in their studies, especially when their studies use survey research collect data. . Previous literature has highlighted that sample size is one of the main limitations of empirical studies published in top journals (see Aguinis & Lawal, 2012; Green et al., 2016; Uttley, 2019).

The success of QC is also greatly influenced by the accuracy and precision used, this precision will affect the number of samples used then the number of samples taken will be greater, and vice versa, the greater the precision used will be the smaller the number of samples taken, but the precision and samples taken will likely affect the accuracy of QC results, by weighing it, so that the purpose of this paper is to find out how big the exact sample size in the QC at the city district level.

2. Research methodology

2.1. Data Sources and Research Variables

The data used in this study are primary data obtained from the real calculation process on the ground (real count) at the time of the Sumedang regency election in 2018. The variables used in this study are the number of votes of candidates for regent number 1 and the total valid votes at each polling station (TPS) in Sumedang regency and the sample size used, where the data collected reached 1925 polling stations from 2026 polling stations or 95%.

2.2. Method and Analysis

This research method is included in the experimental study, which consists of dependent variables and independent variables, dependent variable is sample sizes, while independent variables are the average of candidate votes in each polling station and the percentage of candidate votes. The group's analysis design used non-parametric Kruskal-Wallis test analysis. Kruskal-Wallis is an alternative test for the F test and a one-way ANOVA test for testing the difference in means or for testing the similarity of several mean values and analysis of variance that we can use if the assumption of normality is not met. The design tables include:

Table 1. Candidate 1 vote with sample size.

Candidate Vote	Sample Size (B)						
	n = 92 (B_1)	n = 158 (B_2)	n = 324 (B_3)	n = 408 (B_4)	n = 500 (B_5)	n = 875 (B_6)	n = 1674 (B_7)
Candidate vote (A)	(AB_1)	(AB_2)	(AB_3)	(AB_4)	(AB_5)	(AB_6)	(AB_7)

Table 2. Percentage of candidate vote by sample size.

Percentage of Candidate vote	Sample Size (B)						
	n = 92 (B_1)	n = 158 (B_2)	n = 324 (B_3)	n = 408 (B_4)	n = 500 (B_5)	n = 875 (B_6)	n = 1674 (B_7)
Percentage (P)	($P\ B_1$)	($P\ B_2$)	($P\ B_3$)	($P\ B_4$)	($P\ B_5$)	($P\ B_6$)	(PB_7)

2.3. Population and Sample

2.3.1. Population

The voter population in the implementation of *quick count* is all voters who have the right to vote and have been registered on the Permanent Voter List (DPT) by the Election Commission (KPU). The population of polling stations in the implementation of quick *count* is all polling stations (TPS) in all districts of Sumedang.

2.3.2. Sample

As for the sample of voters in the implementation of *quick count* is the people who will vote and have the right to vote in polling stations that have been randomly selected. Polling stations (TPS) sample is a TPS that is randomly selected will be a sample in the implementation of the quick count. This paper uses TPS as its unit of analysis whose sampling is done randomly. Therefore, to find out the size of the sample of polling stations that must be taken in order to represent or represent the population, it can be determined through the following equation from Parel, Cristina P:

$$n = \frac{NZ^2 p(1-p)}{Nd^2 + Z^2 p(1-p)} \tag{1}$$

n= size or number of polling place samples, Z=Reliability coefficient or standard normal variable value, *d* = Tolerable error rate (margin of *error),* p= 0.5 (proportion of those who voted in an election), q= proportion of those who did not vote in an election i.e. (1−p), N= rate or population number of polling stations.

2.4. Research Hypothesis

Based on the analysis design above, a provisional hypothesis is proposed which will be tested in this study in the form of a research hypothesis as follows:

- $H_0 : \mu_{AB1} = \mu_{AB2} = \mu_{AB3} = \mu_{AB4} = \mu_{AB5} = \mu_{AB6} = \mu_{sAB7}$
  The average candidate votes in the sample size group are the same, or there is no difference in the mean between groups
  $H_1 : \mu_{AB1} \neq \mu_{AB2} \neq \mu_{AB3} \neq \mu_{AB4} \neq \mu_{AB5} \neq \mu_{AB6} \neq \mu_{sAB7}$
  The average candidate votes in the sample size group are not the same, or there is a difference in the mean between the groups
- $H_0 : \mu_{PB1} = \mu_{PB2} = \mu_{PB3} = \mu_{PB4} = \mu_{PB5} = \mu_{PB6} = \mu_{PB7}$
  The average percentage of votes in the sample size group is the same, or there is no difference in the average percentage between the sample size groups
  $H_1 : \mu_{PB1} \neq \mu_{PB2} \neq \mu_{PB3} \neq \mu_{PB4} \neq \mu_{PB5} \neq \mu_{PB6} \neq \mu_{PB7}$
  The average percentage of votes in the sample size group is not the same, or there is a difference in the average percentage between the sample size groups.

3. Results of research and discussion

The Regional Election Commission (KPU) of Sumedang stated that in the recapitulation of permanent voter data in the 2018 Sumedang District Head Election there were 834276 voters spread in 2026 polling stations. Real data recapitulation (Real Count) conducted by the internal team can be as follows:

Table 3. Real count recapitulation results by district.

No	Districts	Vote Of Candidate Number 1	Vote Of Other Candidate	Total Vote	Polling Stations
1	Buahdua	7.666	10.157	17.823	23
2	Cibugel	2.449	9.104	11.553	50
3	Cimalaka	16.155	17.532	33.687	33
4	Cimanggung	17.469	26.531	44.000	82
5	Cisarua	3.923	7.518	11.441	148
6	Cisitu	6.818	9.984	16.802	37
7	Conggeang	6.290	11.552	17.842	57
8	Darmaraja	5.931	12.007	17.938	56
9	Ganeas	7.913	6.999	14.912	64
10	Jatigede	6.481	7.728	14.209	51
11	Jatinangor	11.679	20.503	32.182	51
12	Jatinunggal	8.881	15.188	24.069	104
13	Pamulihan	15.233	15.237	30.470	90
14	Paseh	8.321	13.188	21.509	91
15	Rancakalong	11.996	11.159	23.155	74
16	Situraja	8.761	15.065	23.826	68
17	Sukasari	6.832	12.986	19.818	71
18	Sumedang Selatan	27.000	18.779	45.779	67
19	Sumedang Utara	26.102	22.870	48.972	135
20	Surian	1.670	5.311	6.981	128
21	Tanjungkerta	8.828	11.260	20.088	84
22	Tanjungmedar	7.020	7.936	14.956	43
23	Tanjungsari	23.273	24.028	47.301	154
24	Tomo	4.086	9.149	13.235	47
25	Ujung Jaya	6.549	10.684	17.233	54
26	Wado	6.968	13.214	20.182	63
TOTAL		**264.294**	**345.669**	**609.963**	**1.925**

The data collected only reached 1925 TPS from 2026 or 95% conducted for 3 days, this data will be used as a reference in data processing.

3.1. Sample Size

The size of the voter population is 834,276 voters spread across 2026 polling stations. The calculation of sample rate taken relies heavily on *d* = precision (*margin of error)*, in this study tried to use the precision of 1%, 2.5%, 5%, 7.5%, and 10%, with a confidence level of 95% and a default value of *Z* = 1.96. The reason for choosing the precision amount is based on quick count comparisons conducted by several leading institutions in Indonesia, for comparison, the 2020 Jambi Provincial Governor Election, LSI with a sample of 325 TPS from 8236 TPS or 2415862. The 2020 Riau Islands Governor Election, by LSI with the number of samples, is 250 TPS from 8416 TPS or 2415862. Election of the 2020 Karawang Regency Regent, by Indicator with a sample of 200 TPS from 4436 TPS. The election of the mayor of Medan City 2020, by Charta Politica with a total sample of 300 polling stations from 4303 polling stations. Meanwhile, the quick count conducted by Swamedia Research and Communication (SRC) on the selection of regents in Sumedang district with a total sample of 408 polling stations from 2026 polling stations. From the comparison of the use of the number of samples, in this study we want to compare the number of samples as in the table below based on the following formula:

$$n = \frac{2026\,(1.96^2)(*\,0.5(1-0.5))}{(2026 * 0.001^2) + (1.96^2 * 0.5(1-0.5))} \tag{2}$$

$$n = 1673.1 \;\approx 1674$$

The number of samples with different precision is obtained as follows:

Table 4. Sample size based on precision used.

Precision (%)	Sampel Size (n)
1	1674
2.5	875
3.80	500
4.34	408
5	324
7.5	158
10	92

Samples of 408 and 500 were taken because the number of samples in that range is often used in various District/City level QCs, with the size of the number of samples above to be tested at the average difference level.

3.2. Descriptive Statistics

3.2.1. Votes from candidate at each polling station

The sample selection process was carried out by random sampling with 100 repetitions from each sample size group. For example, if n = 92, then the data is taken as many as 92 TPS and then carried out 100 repetitions, as well as the other sample size groups so that the following descriptive statistics can be obtained:

Table 5. Description of the vote distribution of candidate vote at polling station based on sample size.

Descriptive statistics	Sample Size						
	n = 92	n = 158	n = 324	n = 408	n = 500	n = 875	n = 1674
Mean	135,08	136,63	137,09	137,20	137,00	137,12	137,30

Descriptive statistics	Sample Size						
	n = 92	n = 158	n = 324	n = 408	n = 500	n = 875	n = 1674
Median	135,08	136,97	137,62	137,11	137,16	137,03	137,28
Variance	33,48	18,92	10,49	8,32	6,07	2,73	0,30
Std. Deviation	5,79	4,35	3,24	2,88	2,46	1,65	0,55
Minimum	119,37	124,82	129,10	130,49	131,11	133,65	136,16
Maximum	147,57	145,02	144,84	143,46	142,20	140,55	138,67
Skewness	-0,065	-0,427	-0,361	-0,153	-0,277	0,013	0,241

The above empirical results showed that the average votes of the candidate ranged from 135.08 votes per polling station for a sample of 92 and the highest average was 137.3 votes in the sample of 1674. The highest votes were 147.57 in the sample of 92 and by 138 votes for the sample of 1674, while the lowest votes were 119.37 for the sample of 92 and 136 in the sample of 1674, the other distribution can be seen in the table above.

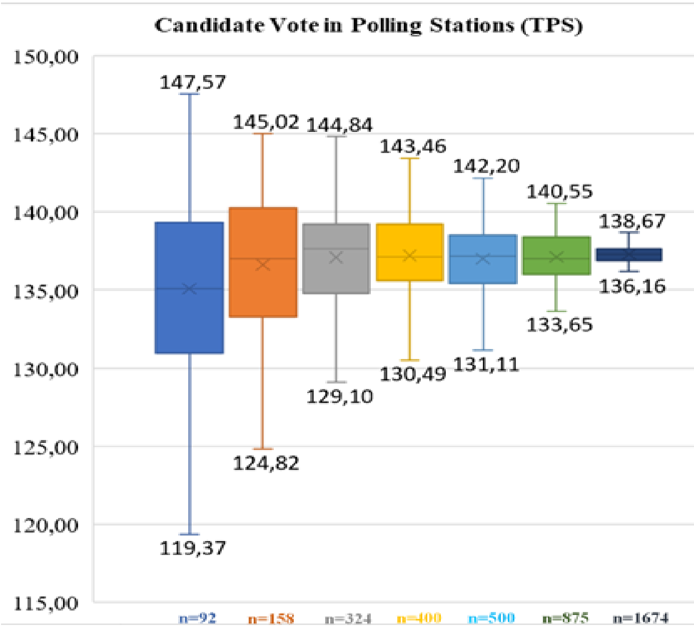

Figure 1. *Boxplot* distribution of Candidate vote at Each Polling Station Based on Sample Size

Table 5 data can also be explained by *boxplot*, where there are 7 data groups, namely sample size data groups 92, 158, 324, 408, 500, 875, and 1674. The average position of *boxplot* on sample size 92 is below that of *boxplots* of other sample sizes, this indicates that the percentage value of sample size 92 is lower than the average percentage of other sample size votes, as stated in table 6 where the sample mean size 92 is 135.08 at the lowest than the mean value on other sample sizes.

The sample size of 1674 is more homogeneous than the smaller sample size, this is indicated by the length of the 1674 box is shorter than all the boxes or in other words, The percentage variant of candidate vote at n = 1674 is 0.30 smaller than the *percentage* variance in other sample sizes, and the percentage variant of the sample n = 92 has the largest value of 33.48, as well as the sample n = 158 which has a variant of 18.92, from descriptive statistics and boxplot, it can be concluded that small size samples have an average difference with the size of the larger sample, it can be seen also that the larger the sample taken then the smaller the standard deviation, this shows that, the small size samples tend to be close to average.

3.2.2. Percentage of candidate votes at each polling station

Table 6. Description of the percentage distribution of candidate votes in each polling station based on sample size.

Descriptive statistics	Sample Size						
	n = 92	n = 158	n = 324	n = 408	n = 500	n = 875	n = 1674
Mean	42,72	43,09	43,21	43,28	43,25	43,27	43,32
Median	42,78	43,05	43,25	43,27	43,31	43,29	43,32
Variance	2,47	1,33	0,68	0,56	0,42	0,17	0,02
Std. Deviation	1,57	1,15	0,83	0,75	0,65	0,41	0,14
Minimum	38,85	40,18	41,31	41,50	41,79	42,48	43,01
Maximum	46,06	45,97	45,58	45,00	44,77	44,30	43,70
Skewness	-0,084	-0,275	-0,028	0,111	-0,092	0,247	0,219

The above empirical results show that the average percentage of candidate votes is 42.72% for a sample of 92 and the other average is about 43%. The highest percentage gain was 46.06% in the sample of 92 and by 43.70% of the votes for the sample of 1674, while the lowest percentage was 38.85% for the sample of 92 and 43.01% in the sample of 1674, the distribution of other percentages can be seen in the table above.

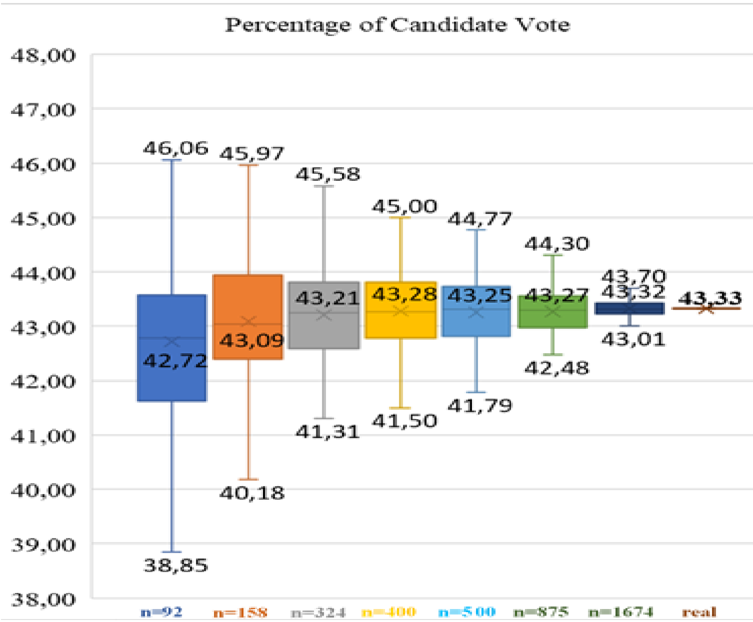

Figure 2. *Boxplot* Distribution percentage of Candidate vote in each polling station based on the sample size

Table 6 data can also be explained by boxplot, where there are 7 groups of data, namely sample size data groups 92, 158, 324, 408, 500, 875, and 1674. The average position of *boxplot* on sample size 92 is below that of *boxplots* of other sample sizes, this indicates that the percentage value of sample size 92 is lower than the average percentage of other sample size sounds, as stated in table 6 where the mean of sample size 92 is 42.72% lower than the mean value in other sample sizes of about 43%.

The sample size of 1674 is more homogeneous than the smaller sample size, this is indicated by the length of the box shorter than all the boxes or in other words, the *percentage* variance at n = 1674 is 0.02 smaller than the *percentage* variance in other sample sizes, and the percentage variant of the sample n = 92 has the largest value of 2.47, as well as the sample size n = 158 which has a variant of 1.15, from descriptive statistics and boxplot, it can be concluded that the sample size n = 92 and n = 157 has an average difference with other sample sizes. From the table and boxplot above, it can be concluded that a small sample size has an average difference with a larger sample size, it can also be seen that the larger the sample is taken, the smaller the standard deviation, this indicates that the values tend to be close to average. The sample sizes n=324, n=408, n=500, n=875 and n=1674 are the most appropriate sample sizes in quick count execution.

3.3. Hypothesis test on sample size differences

This study proposed two hypotheses that must be tested empirically. Both hypotheses are about the alleged average difference between the vote and the percentage of the vote to the sample size.

3.3.1. Test the first hypothesis

The first hypothesis in the study was to compare the average votes of candidate number 1 between the sample size groups n=92, n=158, n=324, n=408, n=500, n=875, and n=1674, which are as follows:

Table 7. Mean ranks of Kruskal-Wallis test.

Sample Size		N	Mean Rank
Average Votes	n = 92	100	281.40
	n = 158	100	351.78
	n = 324	100	373.68
	n = 408	100	366.46
	n = 500	100	355.64
	n = 875	100	353.13
	n = 1674	100	371.40
	Total	700	

The Mean Rank value shows the average rank of each treatment. In the above case, the mean rank of candidate votes with a sample of 324 is higher than all groups, while the sample n of 92 is the smallest. Are these differences all statistically significant overall, then this is where the role of the Kruskal-Wallis Test, which is to measure statistically whether the difference in the average ratings will be significant or not?

The hypothesis formulated in the first test is as follows:

$$H_0 : \mu_{AB1} = \mu_{AB2} = \mu_{AB3} = \mu_{AB4} = \mu_{AB5} = \mu_{AB6} = \mu_{sAB7}$$

The average votes in the sample size group were the same, or there was **no difference** in averages between sample size groups of AB_1 (n=92), AB_2 (n=158), AB_3 (n=324), AB_4 (n=408), AB_5 (n=500), AB_6 (n=875), and AB_7 (n=1674).

$$H_1 : \mu_{AB1} \neq \mu_{AB2} \neq \mu_{AB3} \neq \mu_{AB4} \neq \mu_{AB5} \neq \mu_{AB6} \neq \mu_{sAB7}$$

The average votes in the sample size group are not the same, or **there is an** average difference between the sample size groups of 92, 158, 324, 408, 500, 875, and 1674.

The basis for the decision-making in the Kruskal-Wallis test is to compare the value of significance (Asymp. Sig) with a probability of 0.05, with the following conditions:

- If the value of Asymp.Sig > 0.05, then accept H_0 or no difference **in** average
- If the value of Asymp.Sig < 0.05, then reject H_0 or accept H_1 in other words there is a difference **in** average

Table 8. Test Statistics[a,b].

	Average Votes 1
Chi-Square	14.770
df	6
Asymp. Itself.	.022

a. Kruskal Wallis Test

b. Grouping Variable: SAMPLE SIZE

Based on the output of the test statistic above, it is known that the value of Asymp. Sig is 0.022 < 0.05, thus, it can be concluded that H_0 is rejected or H_1 is accepted which means that there is a real (significant) **difference** between the votes of the seven sample size groups, so it can be decided that the votes of the candidate on the sample 92, 158, 324, 408, 500, 875 and 1674 are not the same or different.

Because the test results *showed* H_0 was rejected (there was a difference), it was tried a follow-up test (post hoc test) with the Bonferroni test and the Games-Howell test to see which sample size groups experienced differences, as stated in the following table:

Table 9. Post Hoc Advanced Tests.

Multiple Comparisons				
Dependent Variable: AVERAGE CANDIDATE VOTE 1				
(I) SAMPLE SIZE	(J) SAMPLE SIZE	Mean Difference (I-J)	POST HOC FURTHER TEST	
			Bonferroni	Games-Howell
n = 92	n = 158	-1,5531	*	
	n = 324	-2,0167	*	*
	n = 408	-2,1248	*	*
	n = 500	-1,9202	*	*
	n = 875	-2,0448	*	*
	n = 1674	-2,2279	*	*
n = 158	n = 92	1,5531	*	
	n = 324	-0,4636		
	n = 408	-0,5717		
	n = 500	-0,3671		
	n = 875	-0,4917		
	n = 1674	-0,6748		
n = 324	n = 92	2,0167	*	*
	n = 158	0,4636		
	n = 408	-0,1081		
	n = 500	0,0965		
	n = 875	-0,0281		
	n = 1674	-0,2112		
n = 408	n = 92	2,1248	*	*
	n = 158	0,5717		
	n = 324	0,1081		
	n = 500	0,2046		
	n = 875	0,08		
	n = 1674	-0,1031		
n = 500	n = 92	1,9202	*	*

Multiple Comparisons				
Dependent Variable: AVERAGE CANDIDATE VOTE 1				
(I) SAMPLE SIZE	(J) SAMPLE SIZE	Mean Difference (I-J)	POST HOC FURTHER TEST	
			Bonferroni	Games-Howell
	n = 158	0,3671		
	n = 324	-0,0965		
	n = 408	-0,2046		
	n = 875	-0,1246		
	n = 1674	-0,3077		
n = 875	n = 92	2,0448	*	*
	n = 158	0,4917		
	n = 324	0,0281		
	n = 408	-0,08		
	n = 500	0,1246		
	n = 1674	-0,1831		
n = 1674	n = 92	2,2279	*	*
	n = 158	0,6748		
	n = 324	0,2112		
	n = 408	0,1031		
	n = 500	0,3077		
	n = 875	0,1831		
*. The mean difference is significant at the 0.05 level.				

From the further test with the post hoc test above it is seen that the sample size group that showed the difference in the average vote of the candidate, namely the sampling of the sample size n = 92 is different from the sampling of n =158, n = 324, n = 408, n = 500, n = 875 and n = 1674. The Games-Howell test showed that taking samples n=92 and n=158 was different from sampling n=324, n=408, n=500, n=875, and n=1674.

3.3.2. Second hypothesis test

Table 10. Mean ranks of Kruskal-Wallis test

Sample Size		N	Mean Rank
percentage of candidate vote	n = 92	100	278.77
	n = 158	100	333.49
	n = 324	100	355.60
	n = 408	100	363.88
	n = 500	100	365.83
	n = 875	100	364.40
	n = 1674	100	391.54
	Total	700	

The Mean Rank value shows the average rank of each treatment. In the above case, the mean rank of candidate votes with a sample of n 1674 is higher than all groups, while the sample n of 92 is the smallest. Are these differences all statistically significant overall, then this is where the role of the Kruskal-Wallis Test, which is to measure statistically whether the difference in the average ratings will be significant or not?

The formulation of the hypothesis proposed in the second hypothesis test is as follows:

$$H_0 : \mu_{PB1} = \mu_{PB2} = \mu_{PB3} = \mu_{PB4} = \mu_{PB5} = \mu_{PB6} = \mu_{PB7}$$

The average percentage of votes in the sample size group is the same, or **there is no difference** in the average percentage between sample size groups of PB_1 (n=92), PB_2 (n=158), PB_3 (n=324), PB_4 (n=408), PB_5 (n=500), PB_6 (n=875), and PB_7 (n=1674).

$$H_1 : \mu_{PB1} \neq \mu_{PB2} \neq \mu_{PB3} \neq \mu_{PB4} \neq \mu_{PB5} \neq \mu_{PB6} \neq \mu_{PB7}$$

The average percentage of votes in the sample size group is not the same, or **there is an** average percentage difference between sample size groups of 92, 158, 324, 408, 500, 875, and 1674.

The basis for the decision-making in the Kruskal-Wallis test is to compare the value of significance (Asymp. Sig) with a probability of 0.05, with the following conditions:

- If the value of Asymp. Sig > 0.05, then accept H_0 or no difference **in** average
- If the value of Asymp. Sig < 0.05, then reject H_0 or accept H_1 in other words there is a difference **in** average

Table 11. Test Statistics[a,b].

	Percentage of Candidate Vote Number 1
Chi-Square	18.958
df	6
Asymp. Itself.	.004

a. Kruskal Wallis Test

b. Grouping Variable: SAMPLE SIZE

Based on the output of the test statistic above, it is known that the value of Asymp. Sig is 0.004 < 0.05, thus, it can be concluded that H_0 is rejected or H_1 is accepted which means that there is a real (significant) **difference** between the percentage of votes from the seven sample size groups, so it can be decided that the percentage of votes of the candidate on the sample n = 92, 158, 324, 408, 500, 875 and 1674 is not the same or different.

Because the test results *showed* H_0 was rejected (there was a difference), it was tried a follow-up test (post hoc test) with the Bonferroni test and the Games-Howell test to see which sample size groups experienced differences, as stated in the following table:

Table 12. Post hoc advanced tests.

Multiple Comparisons				
Dependent Variable:PERCENTAGE OF CANDIDATE VOTES NUMBER 1				
(I) Sample Size	(J) Sample Size	Mean Difference (I-J)	Post Hoc Further Test	
			Bonferroni	Games-Howell
n = 92	n = 158	-0,3754		
	n = 324	-0,4931	*	
	n = 408	-0,5621	*	*
	n = 500	-0,5375	*	*
	n = 875	-0,5541	*	*
	n = 1674	-0,6089	*	*
n = 158	n = 92	0,3754		
	n = 324	-0,1177		
	n = 408	-0,1867		
	n = 500	-0,1621		
	n = 875	-0,1787		
	n = 1674	-0,2335		

Multiple Comparisons				
Dependent Variable:PERCENTAGE OF CANDIDATE VOTES NUMBER 1				
(I) Sample Size	(J) Sample Size	Mean Difference (I-J)	Post Hoc Further Test	
			Bonferroni	Games-Howell
n = 324	n = 92	0,4931	*	
	n = 158	0,1177		
	n = 408	-0,069		
	n = 500	-0,0444		
	n = 875	-0,061		
	n = 1674	-0,1158		
n = 408	n = 92	0,5621	*	*
	n = 158	0,1867		
	n = 324	0,069		
	n = 500	0,0246		
	n = 875	0,008		
	n = 1674	-0,0468		
n = 500	n = 92	0,5375	*	*
	n = 158	0,1621		
	n = 324	0,0444		
	n = 408	-0,0246		
	n = 875	-0,0166		
	n = 1674	-0,0714		
n = 875	n = 92	0,5541	*	*
	n = 158	0,1787		
	n = 324	0,061		
	n = 408	-0,008		
	n = 500	0,0166		
	n = 1674	-0,0548		
n = 1674	n = 92	0,6089	*	*
	n = 158	0,2335		
	n = 324	0,1158		
	n = 408	0,0468		
	n = 500	0,0714		
	n = 875	0,0548		
*. The mean difference is significant at the 0.05 level.				

From the further test with the post hoc test above it is seen that the sample size group that shows the difference in the average percentage of candidate vote, namely sampling sizes n = 92 and n =158 is different from sampling n = 324, n = 408, n = 500, n = 875 and n = 1674. The Games-Howell test showed that the sample sizes n=92, n=158, and n=324 differed from the sampling of n=408, n=500, n=875, and n=1674. From these two further tests, it can be concluded that the sample rate n = 408, n = 500, n = 875, and n = 1674 is the most appropriate sample size in the implementation of the quick count.

In the process of this research, there are several limitations experienced for more attention to improving the research in the future. Some of the limitations of this study include 1. The sampling method used is simple random sampling, while quick counts always use stratified or multi-stage random sampling. 2. The number of TPS samples used was seven groups. 3. The analysis used is Kruskal-Walls, which should be able to use ANOVA where the number of variables studied can also be more than one, for example comparing the number of several samples and several sampling methods.

4. Conclusion

What is the correct number of samples in the quick count? According to Cohen (2007, p. 101) the larger the sample from the size of the existing population, the better, but there is a minimum number that must be taken by researchers, which is as many as 30 samples. From the theory, these rules, the sample of 30 can be tolerated but the more samples used, the better. Although in this paper, only "seven sample size groups" are tested in the case of quick count data, it can be concluded that:

- The number of samples n=92 and n=158 (precision 10 - 7.5%) is very vulnerable to producing accurate quick count data, if this number of samples is used in quick counts it will be possible to produce a sound range with great precision as shown in Figure 1 and Figure 2.
- The number of samples n=324 (5% precision) if used in a quick count it will produce data that can still be tolerated, although the results of the Games-Howel test show that the number of samples n=92 and n=158 is not different from n=324.
- The right number of samples to perform a quick count is n=408 or the larger it is because the larger the number of TPS samples taken, the smaller the precision range, as shown in Figure 1 and Figure 2.

5. Acknowledgment

The author thanked and congratulated the Candidate Couple number 1, DR. H. Dony Ahmad Munir, ST., M.M (Regent of Sumedang) - H. Erwan Setiawan, ST. (Vice Regent of Sumedang) and to all the contributing teams, especially the Survey, Quick Count and Real Count data collection teams in the Sumedang district elections so that the data obtained can be useful and useful, not to forget to The late Entjeng Shobirin Nadj as CEO / Founder of Swamedia Research and Communication and all staff.

References

[1] Aji, R HS, 2016 Statistik, polling, dan pemahaman metodologi pada teknik penelitian survai ekonomi Islam, Al-Iqtishad: Jurnal Ilmu Ekonomi Syariah 3 (2), 235-254
[2] Anselin, L 1993 Exprolatory Spatial Data Analysis and Geographic Information System California Santa Barabara: National Center for Geographic Information and Analysis
[3] Bais, F, Schouten, J G, Lugtig, P, Toepoel, V, Arends-Toth, J, Douhou, S, Kieruj, N, Morren, M, dan Vis, C 2019 Can survey item characteristics relevant to mode-specific measurement error be coded reliably Sociological Methods & Research, 48 (2), 263–295
[4] Bima 2013 Pembentukan Sampel Baru yang Masih Memenuhi Syarat Valid dan Reliable dengan Teknik Resampling pada Data Kuisioner Tipe Yes/No Prosiding Seminar Nasional Matematika UNY
[5] Cahyadi, K R 2013 Quick Count dalam Perspektif Pemilukada Journal Administrasi Negara Vol 2, Number 2
[6] CNN Indonesia 2020 Rapat Pleno KPU Tasikmalaya Ricuh, 3 Polisi Terluka https://www.cnnindonesia.com/nasional/20201216140240-32-582904/rapat-pleno-kpu-tasikmalaya-ricuh-3-polisi-terluka
[7] Cochran, W G 1977 Sampling Techniques Third Edition. America: A john Wiley & Sons, Inc
[8] Cohen, Louis., Lawrence Manion and Keith Morrison 2007 Research Methods In Education, Routledge Taylor & Francis E-Library, USA and Canada, P 101
[9] Desvira, D 2014 Penerapan Algoritma Greedy dalam Menentukan Sampel TPS pada Quick Count Makalah IF2211 Strategi Algoritma. Bandung: Institut Teknologi Bandung
[10] Eriyanto 1990 Metodelogi Polling Memberdayakan Suara Rakyat Bandung PT Remaja Rosdakarya
[11] Estok 2002 The Quick Count and Election Observation: Handbook for Civic Organizations and Political Parties Penerbit Nasional Democratic Institute for International Affairs
[12] Gilliland, D dan Melfi, V 2010 A Note on Confidence Interval Estimation and Margin of Error Journal of Statistics Education. Vol 18, Number 1
[13] Israel 1992 Determining Sampling Size University of Florida

[14] Juwairiah 2009 Aplikasi Quickcount Pemilihan Presiden RI Menggunakan Teknologi Mobile Prosiding Seminar Nasional Informatika 2009 semnasIF 2009 UPN Veteran Yogyakarta, 1979, 23-28
[15] Kismiantini 2007 Pengumpulan Data dengan Quick Count dan Exit Poll Prosiding Seminar Nasional Penelitian, Pendidikan dan Penerapan MIPA, FMIPA Universitas Negeri Yogyakarta
[16] Kompas 2014 Bagaimana cara kerja 'Quick Count'? https://nasional.kompas.com/read/2014/07/11/06330421
[17] Kurnia, Ahmad 2015 Managemen Penelitian: Teknik Sampling Jakarta: Reconiascript Publishing
[18] LSI 2006 Panduan Menyelenggarakan Quick Count. [Terhubung berkala] http://www.20julbooklsi.pdf. [27 Desember 2015]
[19] Lugtig, P and Luiten A 2021 Do shorter stated survey length and inclusion of a QR code in an invitation letter lead to better response rates? Survey Methods: insights from the field
[20] Memon, M A, Ting, H, Ramayah, T, Chuah, F, & Cheah, J H 2020 Sample Size For Survey Research: Review And Recommendations, Journal of Applied Structural Equation Modeling, Vol 4 No 2, pp i-xx
[21] Muhtadi 2019 Sejarah Quick Count sebagai Alat Kontrol Kecurangan, Business Lounge Journal
[22] Nurdin 2018 Aplikasi Quick Count Pilkada dengan Menggunakan Metode Random Sampling Berbasis Android, Techsi, 10, 141-153
[23] Nurhayati 2008 Studi Perbandingan Metode Sampling antarA Simple Random dengan Stratified Journal Basis Data Vol 3, Number 1
[24] Omair 2014 Sample size estimation and sampling techniques for selecting a representative sample, Journal of Health Specialties, 2, 142-147
[25] Prahutama, A 2014 Analisis Kemenangan Pemilihan Gubernur (PILGUB) Jawa Tengah 2013 dengan Autokorelasi Spasial Jurnal Statistika, Vol 2, No 1, Halaman 6-12
[26] Parel, C.P, G.C Caldito, P.L Ferre, G.G De Guzman, C.C Sinsioco, dan R.H Tan 1973 Sampling Design and Procedure, PSSC (The Philippine Social Science Council), p 53
[27] Setiawan (2014), "(Monte Carlo) Resampling Technique in Validity Testing and Reliability Testing," International Journal of Computer Application, 91, 6-11.
[28] Syafitri, Utami, D., Bagus S. dan Salamatuttanzil. (2008). Pengujian Autokorelasi terhadap Sisaan Model Spatial Logistik. Seminar Nasional Matematika dan Pendidikan Matematika, Halaman: 2. Bogor: FMIPA IPB.
[29] Susanto 2019 Menghindari konflik akibat penyiaran hasil hitung cepat (quick count) pemilihan presiden 2019 (kajian penerapan nilai humanis dan religius), Journal of Humanities, Vol 1 (1), 62-88
[30] Taherdoost 2016 Sampling Method in Research Methodology; How to Choose a Sampling Technique for Research, International Journal of Advance Reseacrh in Management, 5, 18-27
[31] Taherdoost 2016 How to Design and Create an Effective Survey/Questionnaire; A step by step guide, international journal of advance reseacrh in management, 5, 37-41
[32] Ujiyati T P 2004 Quick Count http://www.lp3es.or.id/ program/pemilu2004/QCount.htm [19 November 2015]
[33] Ulya, S F, Sukestiyarno Y L, dan Hendikawati P 2018 Analisis Prediksi Quick Count dengan Metode Stratified Random Sampling dan Estimasi Confidence Interval Menggunakan Metode Maksimum Likelihood UNNES Journal of Mathematics 7(1): 108-119
[34] Yamane, T 1967 Elementery Sampling Theory U S America: Prentice Hall