BAB IV

DATA ANALYSIS AND RESEARCH FINDING

4.1 Data Analysis

Data analysis was intended to find out wether the application of Clustering Technique significantly affects the students' vocabulary mastery. The analysis was computed by applying the t test formula to discover the hypothesis of this research was accepted or rejected, before it the researcher should do requirement test at the first by using normality and homogencity test.

4.1.1 The Descriptive of Data

The data on this research were quantitative data, they were taken from experiment research design, and there were 60 students who were taken as sample of this research. They were divided into two class, namely experiment and contolclass.the students in experiment class were taught by applying clustering Technique and the students in control class were taught by using lecturing method. The population of this research was the students of the seventh grade at MTs.Islamiyah YPI BatangKuis. The experiment class was VII-1 and the control class was VII-2.

Table 4.1

Students' Score in Experiment Class

Num	Students' Initial Name	Pre-Test Score	Post-Test Score
1	ADF	75	85
2	A	55	75
3	AS	65	80
4	AL	60	75
5	AA	65	85
6	BS	55	75
7	CN	55	75
8	CA	55	85
9	DR	60	70
10	FI	55	55
11	FM	65	90
12	GS	60	75
13	HB	60	70
14	K	60	85
15	LMLS	75	90
16	MJH	75	90
17	MA	75	95
18	MFA	65	80
19	MIM	50	65
20	MQAM	70	80
21	MY	20	60

22	MZN	60	75
23	NR	50	80
24	NNS	65	80
25	ND	65	90
26	RAH	50	70
27	RM	65	90
28	RP	35	65
29	R	50	70
30	RAPB	60	75
31	RRL	75	90
32	RS	40	65
Total		**1890**	**2490**
Mean		**59.06**	**77.81**

Based on the table above, the students' vocabulary mastery that was taught by applying clustering technique showed the minimum score of pre-test was 20, the maximum score of pre-test was 75 and the mean of pre test was 59,06. Meanwhile the minimum score of post-test was 55, the maximim score of post-test was 95, and the mean of post-test was 77,81.

Table 4.2

Students' Score in Control Class

Num	Students' Initial Name	Pre-Test Score	Post-Test Score
1	ADPS	55	60
2	AAS	60	70
3	AP	25	30
4	A	60	65
5	ABP	60	65
6	CWL	50	55
7	ES	75	75
8	FH	70	75
9	FS	65	75
10	FC	60	60
11	FEW	70	75
12	FR	50	55
13	FNS	45	50
14	FAS	60	60
15	H	65	70

16	HRS	60	65
17	HS	60	60
18	HFL	50	55
19	IM	60	65
20	IMS	55	60
21	IP	55	60
22	LH	50	55
23	LS	60	65
24	MB	45	50
25	MKP	55	60
26	MAA	55	60
27	MP	55	65
28	NJN	65	70
Total		**1595**	**1730**
Mean		**56.96**	**61.78**

The table above explains that the students' vocabulary mastery that was taught by teachers' method showed the minimum score of pre-test was 25 and the maximum score was 75, the mean of pre-test was 56,96. Meanwhile the minimum score of post-test was 30 and the maximum score was 75, the mean of post-test was 61,78.

4.1.2. Normality Test

Table 4.3

Normality Test of Pre-Test in Experiment Class

Num	Score	F	Fcum	Zi	FZi	SZi	|FZi-SZi|
1	20	1	1	-3.234	0.0006	0.0312	0.03064
2	35	1	2	-1.992	0.0231	0.0625	0.039333
3	40	1	3	-1.578	0.0572	0.0937	0.036507
4	50	4	7	-0.750	0.2265	0.2187	0.007769
5	55	5	12	-0.336	0.3682	0.375	0.006703
6	60	7	19	0.077	0.5309	0.5937	0.062814
7	65	7	26	0.491	0.6885	0.812	0.123996
8	70	1	27	0.905	0.8174	0.8437	0.026322
9	75	5	32	1.319	0.9065	1	0.093485

To find Z score by using the formula :

$$Z_i = \frac{xi - \bar{x}}{s}$$

1. $Z_i = \frac{20 - 59,06}{12,07} = -3,234$

2. $Z_i = \frac{35 - 59,06}{12,07} = -1,992$

3. Z_i $= \frac{40-59,06}{12,07} = -1.578$

4. Z_i $= \frac{50-59,06}{12,07} = -0,750$

5. Z_i $= \frac{55-59,06}{12,07} = -0,336$

To find out S(Zi) we use the formula :

$$S(Z_i) = \frac{Fcum}{n}$$

1. $S(Z_i) = \frac{1}{32} = 0,031$

2. $S(Z_i) = \frac{2}{32} = 0,062$

3. $S(Z_i) = \frac{3}{32} = 0,093$

4. $S(Z_i) = \frac{7}{32} = 0,218$

5. $S(Z_i) = \frac{12}{32} = 0,375$

From the explanation above, it can be seen that the Liliefors Observation or L_o = 0,124 with n = 32 and at real level α = 0,05 from the list critical coefficient or Liliefors table L_t = 0,156. It can be concluded that the data distribution was **normal**, because L_o (0,124) < L_t (0,156).

Table 4.4

Normality Test of Post-Test in Experiment Class

Num	Score	F	Fcum	Zi	FZi	SZi	|FZi-SZi|
1	55	1	1	-2.301	0.0106	0.0312	0.020556
2	60	1	2	-1.796	0.0361	0.0625	0.026309
3	65	3	5	-1.292	0.0981	0.1562	0.058136
4	70	4	9	-0.788	0.2153	0.2812	0.065911
5	75	7	16	-0.283	0.3883	0.5	0.111677
6	80	5	21	0.220	0.5873	0.6562	0.068933
7	85	4	25	0.724	0.7657	0.7812	0.015479
8	90	6	31	1.229	0.8905	0.9687	0.078224
9	95	1	32	1.733	0.9585	1	0.041488

To find Z score by using the formula :

Z_i $= \frac{xi - \bar{x}}{s}$

1. $Z_i = \frac{55-77,81}{9,91} = -2,301$

2. $Z_i = \frac{60-77,81}{9,91} = -1,796$

3. $Z_i = \frac{65-77,81}{9,91} = -1.292$

42

4. $Z_i = \frac{70-77,81}{9,91} = -0,788$

5. $Z_i = \frac{75-77,81}{9,91} = -0,283$

To find out S(Zi) we use the formula :

$$S(Z_i) = \frac{Fcum}{n}$$

1. $S(Z_i) = \frac{1}{32} = 0.031$
2. $S(Z_i) = \frac{2}{32} = 0.062$
3. $S(Z_i) = \frac{5}{32} = 0.156$
4. $S(Z_i) = \frac{9}{32} = 0.281$
5. $S(Z_i) = \frac{16}{32} = 0.5$

From the explanation above, it can be seen that the Liliefors Observation or L_o = 0,111 with n = 32 and at real level α = 0,05 from the list critical coefficient or Liliefors table L_t = 0,156. It can be concluded that the data distribution was **normal**, because L_o (0,111) < L_t (0,156).

Table 4.5

Normality Test of Pre-Test in Control Class

Num	Score	F	Fcum	Zi	FZi	SZi	|FZi-SZi|
1	25	1	1	-3.343	0.0004	0.03571	0.0353
2	45	2	3	-1.251	0.1053	0.10714	0.0017
3	50	4	7	-0.728	0.2331	0.25	0.0168
4	55	6	13	-0.205	0.4185	0.46429	0.0456
5	60	9	22	0.317	0.6246	0.78571	0.1611
6	65	3	25	0.840	0.7997	0.89286	0.0931
7	70	2	27	1.363	0.9136	0.96429	0.0506
8	75	1	28	1.886	0.9704	1	0.0295

To find Z score by using the formula :

$Z_i = \frac{xi - \bar{x}}{S}$

1. $Z_i = \frac{25 - 59,96}{9,55} = -3,343$

2. $Z_i = \frac{45 - 59,96}{9,55} = -1,251$

3. $Z_i = \frac{50 - 59,96}{9,55} = -0,728$

4. $Z_i = \frac{55 - 59,96}{9,55} = -0,205$

5. $Z_i = \frac{60 - 59,96}{9,55} = -0,317$

To find out S(Zi) we use the formula :

$S(Z_i) = \frac{Fcum}{n}$

1. $S(Z_i) = \frac{1}{28} = 0,035$
2. $S(Z_i) = \frac{3}{28} = 0,017$
3. $S(Z_i) = \frac{7}{28} = 0,25$
4. $S(Z_i) = \frac{13}{28} = 0.464$
5. $S(Z_i) = \frac{22}{28} = 0,785$

From the explanation above, it can be seen that the Liliefors Observation or $L_o = 0,161$ with n = 28 and at real level $\alpha = 0,05$ from the list critical coefficient or Liliefors table $L_t = 0,167$. It can be concluded that the data distribution was **normal**, because $L_o (0,161) < L_t (0,167)$.

Table 4.6

Normality Test of Post-Test in Control Class

Num	Score	F	Fcum	Zi	FZi	SZi	|FZi-SZi|
1	30	1	1	-3.329	0.0004	0.0357	0.0352

2	50	2	3	-1.234	0.1085	0.1071	0.0013
3	55	4	7	-0.710	0.2386	0.25	0.0113
4	60	8	15	-0.187	0.4258	0.5357	0.1099
5	65	6	21	0.336	0.6318	0.75	0.1181
6	70	3	24	0.860	0.8052	0.8571	0.0519
7	75	4	28	1.384	0.9168	1	0.0831

To find Z score by using the formula :

$Z_i = \frac{xi - \bar{x}}{S}$

1. $Z_i = \frac{30 - 61{,}78}{9{,}54} = -3{,}329$

2. $Z_i = \frac{50 - 61{,}78}{9{,}54} = -1{,}234$

3. $Z_i = \frac{55 - 61{,}78}{9{,}54} = -0{,}710$

4. $Z_i = \frac{60 - 61{,}78}{9{,}54} = -0{,}187$

5. $Z_i = \frac{65 - 61{,}78}{9{,}54} = 0{,}336$

To find out S(Zi) we use the formula :

$S(Z_i) = \frac{Fcum}{n}$

1. $S(Z_i) = \frac{1}{28} = 0{,}035$

2. $S(Z_i) = \frac{3}{28} = 0{,}107$

3. $S(Z_i) = \frac{7}{28} = 0{,}25$

4. $S(Z_i) = \frac{15}{28} = 0{,}535$

5. $S(Z_i) = \frac{21}{28} = 0{,}75$

From the explanation above, it can be seen that the Liliefors Observation or $L_o = 0{,}118$ with n = 28 and at real level $\alpha = 0{,}05$ from the list critical coefficient or Liliefors table $L_t = 0{,}167$. It can be concluded that the data distribution was **normal**, because $L_o (0{,}118) < L_t (0{,}167)$.

4.1.3. Homogeneity Test

Homogeneity test used F-test to know what the samples come from the population that homogeneous or not.

Table 4.7
Homogeneity Test of Pre-test

Num	Data	Variants	$F_{observation}$	F_{table}	Conclusion
1	Pre-Test of Experiment Class	145,87	1,59	1,87	Homogenous
2	Pre-Test of Control Class	91,36			

$F_{observation} = \frac{S_1^2}{S_2^2}$

Where : S_1^2 = The biggest variant

S_2^2 = The smallest variant

Based on the variants of both samples of pre-test found that :

$S^2_{experiment}$ = 145,87 N = 32

$S^2_{control}$ = 91,36 N = 28

So :

$F_{observation} = \frac{S^2_{Kexs}}{S^2_{Kcont}}$

$F_{observation} = \frac{145,87}{91,36}$

$F_{observation}$=1,59

Then the coefficient of $F_{observation}$= 1,59 was compared with F_{table}, where F_{table}wasdetermined at real α = 0,05 and the same numerator dk = n-1 = (32-1 = 31) that was exist between dk numerator 30 and 40, the denominator dk = n-1 = (28-1 = 27). Then F_{table}can be calculated by linear interpolation. So that :

$F_{(0,05)(30;27)}$ = 1,88

$F_{(0,05)(40;27)}$ = 1,84

So :

$F_{table} = F_{(0,05)(30;27)} + \frac{31-30}{40-30} (F_{(0,05)(40;27)} - F_{(0,05)(30;27)})$

$= 1,88 + (\frac{1}{10}) (1,84 - 1,88)$

= 1,88 + 0,1 (-0,04)

= 1,88 – 0,004

= 1,87

Because of $F_{observation} < F_t$ atau (1,59 < 1,87) so it can be concluded that the variant was **homogenous.**

Table 4.8
Homogeneity Test of Post-test

Num	Data	Variants	$F_{observation}$	F_{table}	Conclusion
1	Pre-Test of Experiment Class	98,28	1,07	1,87	Homogenous
2	Pre-Test of Control Class	91,13			

$F_{observation} = \frac{S_1^2}{S_2^2}$

Where : S_1^2 = The biggest variant

S_2^2 = The smallest variant

Based on the variants of both samples of pre-test found that :

$S_{experiment}^2$ = 98,28 N = 32

$S_{control}^2$ = 91,13 N = 28

So :

$F_{observation} = \frac{S_{Kexs}^2}{S_{Kcont}^2}$

$F_{observation} = \frac{98,28}{91,13}$

$F_{observation}$=1,07

Then the coefficient of $F_{observation}$= 1,07 was compared with F_{table}, where $F_{table}$was determined at real α = 0,05 and the same numerator dk = n-1 = (32-1 = 31) that was exist between dk numerator 30 and 40, the denominator dk = n-1 = (28-1 = 27). Then F_{table}can be calculated by linear interpolation. So that :

$F_{(0,05)(30;27)}$ = 1,88

$F_{(0,05)(40;27)}$ = 1,84

So :

$F_{table} = F_{(0,05)(30;27)} + \frac{31-30}{40-30} (F_{(0,05)(40;27)} - F_{(0,05)(30;27)})$

$= 1,88 + (\frac{1}{10}) (1,84 - 1,88)$

$= 1,88 + 0,1 (-0,04)$

$= 1,88 - 0,004$

$= 1,87$

Because of $F_{observation} < F_t$ atau (1,07< 1,87) so it can be concluded that the variant was **homogenous.**

4.1.4. Hypothesis Test

The hypothesis was aimed to show the result of the observation sample quantitavely and also to know whether the application of clustering technique significantly affects the students' vocabulary mastery, so the hypoyhesis were:

$H_0 = \mu x \leq \mu y$

$H\alpha = \mu x > \mu y$

From the criteria of the hypothesis, Hα is accepted if $t_{observation} > t_{table}$

To find out whether the application of clustering technique significantly affects the students' vocabulary mastery. The analysis was computed by applying the t-test formula to discover the hypothesis of this research was accpeted or rejected. The formula was stated as the following:

$$t = \frac{Mx - My}{\sqrt{\left(\frac{dx^2 + dy^2}{nx + ny - 2}\right)\left(\frac{1}{nx} + \frac{1}{ny}\right)}}$$

Mx= the mean score of experiment group

My = the mean score of control group

dx = the deviation standard of experimental group

dy = the deviation standard of control group

nx = the total sample of experimental group

ny = the total of control group

Before calculating t-test data, it used the formula bellow to find the deviation standard of both class:

$Mx = \frac{\sum d}{n}$

Table 4.9

The Tabulation of Students' Score at Experiment Class

Num	Students' Initial Name	Pre-Test	Post-Test	$d = (t_2 - t_1)$	dx = d - Mx	(dx)2
1	ADF	75	85	10	-8.75	76.56
2	A	55	75	20	1.25	1.562
3	AS	65	80	15	-3.75	14.06
4	AL	60	75	15	-3.75	14.06
5	AA	65	85	20	1.25	1.562
6	BS	55	75	20	1.25	1.562
7	CN	55	75	20	1.25	1.562
8	CA	55	85	30	11.25	126.56
9	DR	60	70	10	-8.75	76.56
10	FI	55	55	0	-18.75	351.56
11	FM	65	90	25	6.25	39.06
12	GS	60	75	15	-3.75	14.06
13	HB	60	70	10	-8.75	76.56
14	K	60	85	25	6.25	39.06
15	LMLS	75	90	15	-3.75	14.06
16	MJH	75	90	15	-3.75	14.06
17	MA	75	95	20	1.25	1.562
18	MFA	65	80	15	-3.75	14.06
19	MIM	50	65	15	-3.75	14.06
20	MQAM	70	80	10	-8.75	76.56
21	MY	20	60	40	21.25	451.56
22	MZN	60	75	15	-3.75	14.06
23	NR	50	80	30	11.25	126.56
24	NNS	65	80	15	-3.75	14.06
25	ND	65	90	25	6.25	39.06
26	RAH	50	70	20	1.25	1.562
27	RM	65	90	25	6.25	39.06
28	RP	35	65	30	11.25	126.56
29	R	50	70	20	1.25	1.56
30	RAPB	60	75	15	-3.75	14.06
31	RRL	75	90	15	-3.75	14.06
32	RS	40	65	25	6.25	39.06
Total		**1890**	**2490**	**600**	**0**	**1850**
Mean		**59.06**	**77.81**			

$$Mx = \frac{\sum d}{n}$$

$= \frac{600}{32}$

$= 18{,}75$

Table 4.10

The Tabulation of Students' Score at Control Class

Num	Students' Initial Name	Pre-Test	Post-Test	d = (t_2-t_1)	dy = d - My	(dy)2
1	ADPS	55	60	5	0.18	0.0324
2	AAS	60	70	10	5.18	26.8324
3	AP	25	30	5	0.18	0.0324
4	A	60	65	5	0.18	0.0324
5	ABP	60	65	5	0.18	0.0324
6	CWL	50	55	5	0.18	0.0324
7	ES	75	75	0	-4.82	23.2324
8	FH	70	75	5	0.18	0.0324
9	FS	65	75	10	5.18	26.8324
10	FC	60	60	0	-4.82	23.2324
11	FEW	70	75	5	0.18	0.0324
12	FR	50	55	5	0.18	0.0324
13	FNS	45	50	5	0.18	0.0324
14	FAS	60	60	0	-4.82	23.2324
15	H	65	70	5	0.18	0.0324
16	HRS	60	65	5	0.18	0.0324
17	HS	60	60	0	-4.82	23.2324
18	HFL	50	55	5	0.18	0.0324
19	IM	60	65	5	0.18	0.0324
20	IMS	55	60	5	0.18	0.0324
21	IP	55	60	5	0.18	0.0324
22	LH	50	55	5	0.18	0.0324
23	LS	60	65	5	0.18	0.0324
24	MB	45	50	5	0.18	0.0324
25	MKP	55	60	5	0.18	0.0324
26	MAA	55	60	5	0.18	0.0324
27	MP	55	65	10	5.18	26.8324
28	NJN	65	70	5	0.18	0.0324
Total		**1595**	**1730**	**135**	**0**	**174.1072**
Mean		**55.64**	**61.78**			

My $= \frac{\sum d}{n}$

$= \frac{135}{28}$

= 4,82

Based on the calculation data, the result was as follow:

Mx = 18,75

My = 4,82

dx^2 = 1850

dy^2 = 147,10

n_x = 32

n_y = 28

So t-test can be counted as follows:

$$t = \frac{Mx - My}{\sqrt{\left(\frac{dx^2 + dy^2}{nx + ny - 2}\right)\left(\frac{1}{nx} + \frac{1}{ny}\right)}}$$

$$t = \frac{18{,}75 - 4{,}28}{\sqrt{\left(\frac{1850 + 147{,}10}{32 + 28 - 2}\right)\left(\frac{1}{32} + \frac{1}{28}\right)}}$$

$$t = \frac{14{,}47}{\sqrt{\left(\frac{1997{,}1}{60}\right)(0{,}06)}}$$

$$t = \frac{14{,}47}{\sqrt{(33{,}28)(0{,}06)}}$$

$$t = \frac{14{,}47}{\sqrt{1{,}99}}$$

$$t = \frac{14{,}47}{1{,}41}$$

t = 10,26

From the calculating of the data, it can be seen there was significant effect of clustering technique on the students' vocabulary mastery. In order to find out the significant effect of clustering technique, the researcher analyzed the data by applying t-test formula to prove the hypothesis of this research. It was obtained that the coefficient of $t_{observation}$ was 10,26.

In this research, the coefficient of t_{table} for the degree freedom (df) 58 at level of significance 0,05 is between df = 50 and df = 60. Because df = 58 there was not in t distribution, so the researcher used interpolation.

$t_{(50)} = 2{,}01$

$t_{(60)} = 2{,}00$

So :

$t_{(58)} = 2{,}01 + \frac{58-50}{60-50} (2{,}00 - 2{,}01)$

$t_{(58)} = 2{,}01 + \frac{8}{10} (-0{,}01)$

$t_{(58)} = 2{,}01 + 0{,}8 (-0{,}01)$

$t_{(58)} = 2{,}01 - 0{,}008$

$t_{(58)} = 2{,}00$

From the calculation above, it was found that the coefficient of $t_{observation}$(11,13) was higher than the coefficient of t_{table}(2,00). It can be seen as follows:

$$10{,}26 > 2{,}00$$

This result showed that null hypothesis was rejected, the hypothesis formulated as "there was significant effect of applying clustering technique on students' vocabulary mastery. It means that clustering technique significantly affected students' vocabulary mastery.

4.2. Research Finding

1. Based on the result of the calculation above, it was found that the students' vocabulary mastery when the researcher taught by applying clustering Technique got mean 59,06 in pre-test with the maximum score 75 and the minimum score was 20. While in post-test the students got mean 77,81 with the maximum score 95 and the minimum score was 55.

2. The students' achievement at writing descriptive text when the researcher taught by using lecturing method got mean 59,96 in pre-test with the maximum score 75 and the minimum score was 25. While in post-test the students got mean 61,78 with the maimum score 75 and the minimum score was 30.
3. Based on the statistical compulation t-test was found that the coefficient of $t_{observation}$ = 10,26where the coefficient of t_{table}= 2,00. It means that there was significant effect of applying clustering technique on the students' vocabulary mastery. It indicate that H_α was accepted and $H_0$was rejected.

4.2.1. Discussion

There was significant effect on students' vocabulary mastery by usingClustering technique. The students that was taught by clustering technique have higher score than were taught by lecturing method.

It had been explained in chapter 2 that clustering technique would be an effective way to improve students' vocabulary mastery.Through the application of clustering technique, it is hoped that the student can easily visualize and express their thoughts and ideas by giving a grammatical order way and increase their vocabulary mastery. Clustering technique is used to see a visual map of our ideas and able to make us think more creatively in making new association.

Based on the explanation above, the researcher concluded that the applying of clustering technique has significant effect to students' vocabulary mastery.