

CHAPTER 4

ANALYSIS AND DESIGN

4.1 Analysis

This program analyze the data of student's grade with statistics formulas. Statistics have many formulas, in this program the statistics formulas used are mean, median, mode, standard deviation, skewness and kurtosis. This is the formulas that used in this program and the example to resolve the count of statistics.³

1. Mean formula :

$$\bar{x} : \frac{\sum Xi}{n}$$

explanation:

$\sum Xi$ = total of frequencies
 n = frequencies

2. Median formula :

$$Me : \frac{n+1}{2}$$

explanation:

n = frequencies

*first sorting the data before calculat the formula
the calculation for the review findings indicate sequence data*

3. Mode formula :

Mo = data that often appear

³ Bambang Kustitunto and Rudy Badrudin, "Statistika 1 (Deskriptif)", Gunadarma, Jakarta, 1994

4. Standard deviation formula :

$$S = \frac{\sqrt{\sum (xi - \bar{x})^2}}{n - 1}$$

explanation:

xi = data

\bar{x} = mean

n = frequencies

5. Skewness formula :

$$Sk = \frac{3(Me - \bar{x})}{S}$$

explanation:

\bar{x} = mean

Me = median

S = standard deviation

$Sk = 3 \rightarrow$ normal data

$Sk > 3 \rightarrow$ positive data

$Sk < 3 \rightarrow$ negative data

6. Kurtosis formula :

$$\alpha = \frac{1}{n} \frac{\sum (xi - \bar{x})^4}{S^4}$$

explanation:

n = frequencies

xi = data

\bar{x} = mean

S = standard deviation

$\alpha = 0 \rightarrow$ mezo curtic

$\alpha < 0 \rightarrow$ platy curtic

$\alpha > 0 \rightarrow$ leptocurtic

Example to resolve the count of statistic :

Here the data value of 10 students

8, 9, 7, 9, 6, 5, 8, 8, 9, 7

Please count mean, median, mode, standard deviation, skewness and kurtosis with statistics formulas.

1. Mean

$$\bar{x} = \frac{76}{10} = 7,6$$

2. Median

5 6 7 7 8 8 8 9 9 9

$$Me = \frac{10 + 1}{2} = 5,5$$

$$\text{data } 5 = 8$$

$$\text{data } 6 = 8$$

$$\text{median} = \frac{8 + 8}{2} = 8$$

3. Mode

data does not have a mode value

4. Standard deviation

$$S = \sqrt{\frac{16,4}{9}} = 1,35$$

5. Skewness

$$Sk = \frac{3 (8 - 7,6)}{1,35} = 0,8$$

$$Sk < 3 \rightarrow \text{negative data}$$

6. Kurtosis

$$\alpha = \frac{1}{10} \cdot \frac{64,112}{3,32} = 1,9311$$

$$\alpha > 0 \rightarrow \text{leptocurtic}$$

4.2 Design

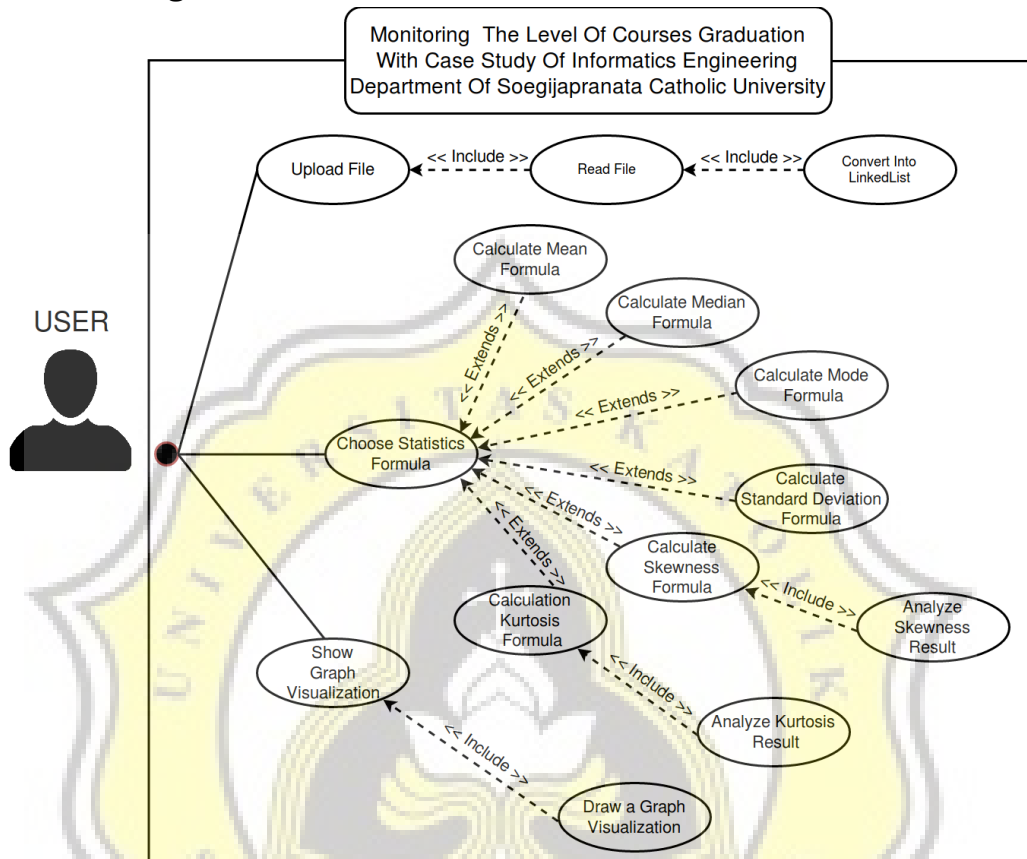
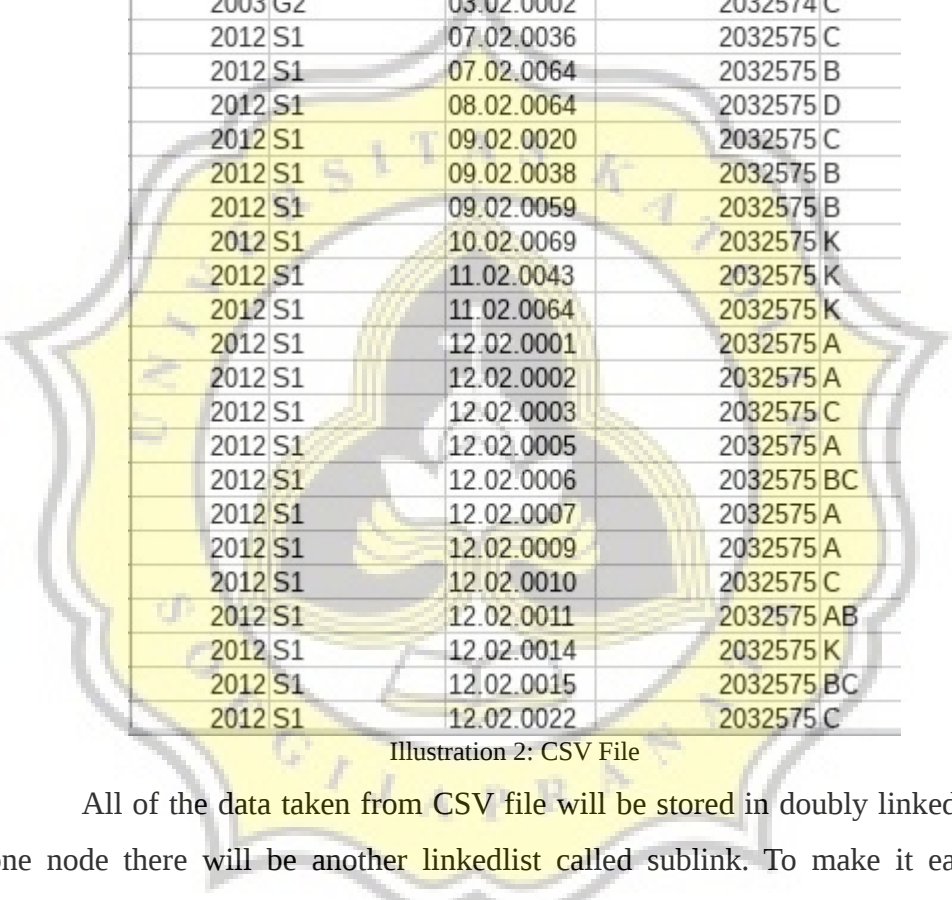


Illustration 1: Use Case Diagram

The use case diagram above show that, the “upload file” use case is a process of inputing csv file by a user. Then the system will be read and convert the csv file into doubly linkedlist. The csv file represent all student’s grade. The data of student grade consist of year, semester, student number (nim), subject code, and grade. In student's grade data, it is possible that one subject can be taken more than once by a student. This happen while student take remedy. Even the data will be redudant, it is still be counted. As the illustration, below will be shown the csv file of student’s grade.



A	B	C	D	E
THNAJAR	KD MSUJI	NIM	KDMK PUS	NILAI
2012	G2	10.02.0057	2053182	A
2009	G2	07.02.0007	2053182	A
2012	S1	09.02.0047	2032575	A
2003	G2	02.02.0008	2032574	A
2009	G2	06.02.0043	2053182	B
2012	G1	09.02.0005	2053182	AB
2003	G2	03.02.0002	2032574	C
2012	S1	07.02.0036	2032575	C
2012	S1	07.02.0064	2032575	B
2012	S1	08.02.0064	2032575	D
2012	S1	09.02.0020	2032575	C
2012	S1	09.02.0038	2032575	B
2012	S1	09.02.0059	2032575	B
2012	S1	10.02.0069	2032575	K
2012	S1	11.02.0043	2032575	K
2012	S1	11.02.0064	2032575	K
2012	S1	12.02.0001	2032575	A
2012	S1	12.02.0002	2032575	A
2012	S1	12.02.0003	2032575	C
2012	S1	12.02.0005	2032575	A
2012	S1	12.02.0006	2032575	BC
2012	S1	12.02.0007	2032575	A
2012	S1	12.02.0009	2032575	A
2012	S1	12.02.0010	2032575	C
2012	S1	12.02.0011	2032575	AB
2012	S1	12.02.0014	2032575	K
2012	S1	12.02.0015	2032575	BC
2012	S1	12.02.0022	2032575	C

Illustration 2: CSV File

All of the data taken from CSV file will be stored in doubly linkedlist. In one node there will be another linkedlist called sublink. To make it easier to understand, the illustration below will be shown the visualization of linkedlist that will be used in this system.

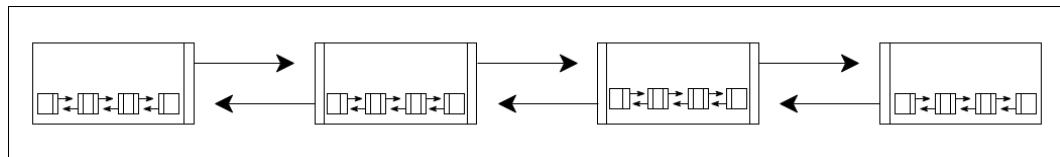


Illustration 3: Linkedlist

As shown in the illustration above, the doubly linkedlist have 2 pointer. There are next and prev pointer. The next pointer is used to connect the next node, and the prev pointer is to connect the previous node. It is used to make linkedlist's operation easier to do.

The data that will be stored in a nodelink are year, semester, and subject code. While the sublink store student's number (nim) and grade. The illustration below will be shown the visual of nodes.

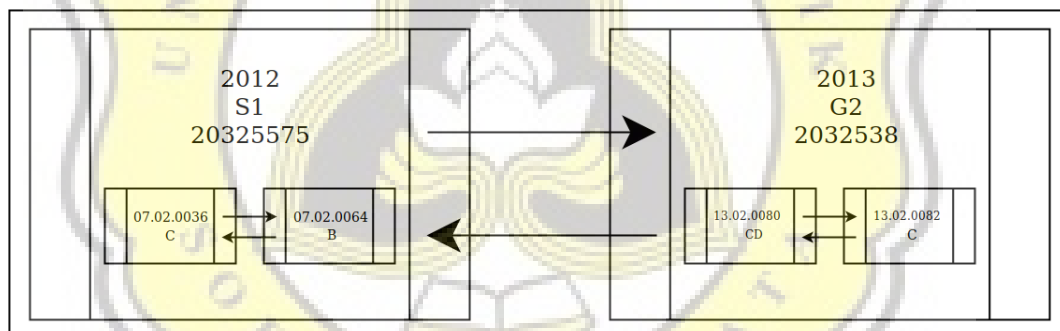


Illustration 4: Value of Nodes

The use case of “choose statistics formula” in the illustration 1, show that user can choose the statistic formulas from the system. The statistic formulas that provided in this system are mean, median, mode, standard deviation, skewness and kurtosis. Below is the completion of the statistic formulas :

1. Mean formula

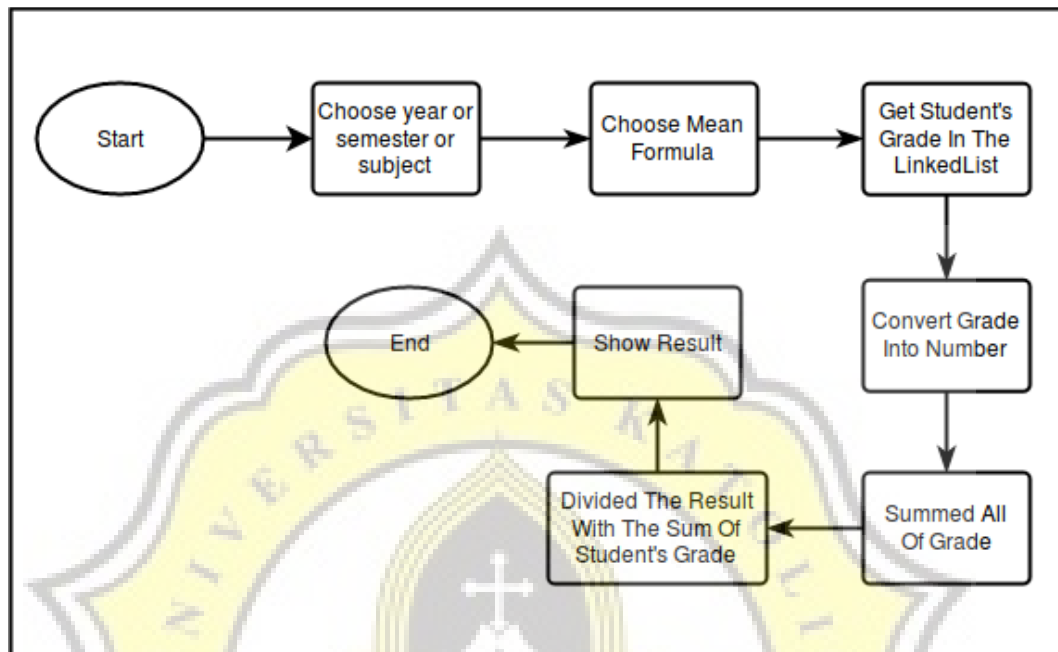


Illustration 5: Mean Flowchart

The mean formula is formula to calculate an average of student's grade. The illustration above show how the system get those average. The first step is user choose year or semester or subject that will be calculated, then user choose the mean formula. Based on those year or semester or subject, the system will search data in the linkedlist. After that the system will convert the grade into number and summed all of the grade. Those result will be divided by the total of the student's grade. And the system will show the result.

2. Median formula

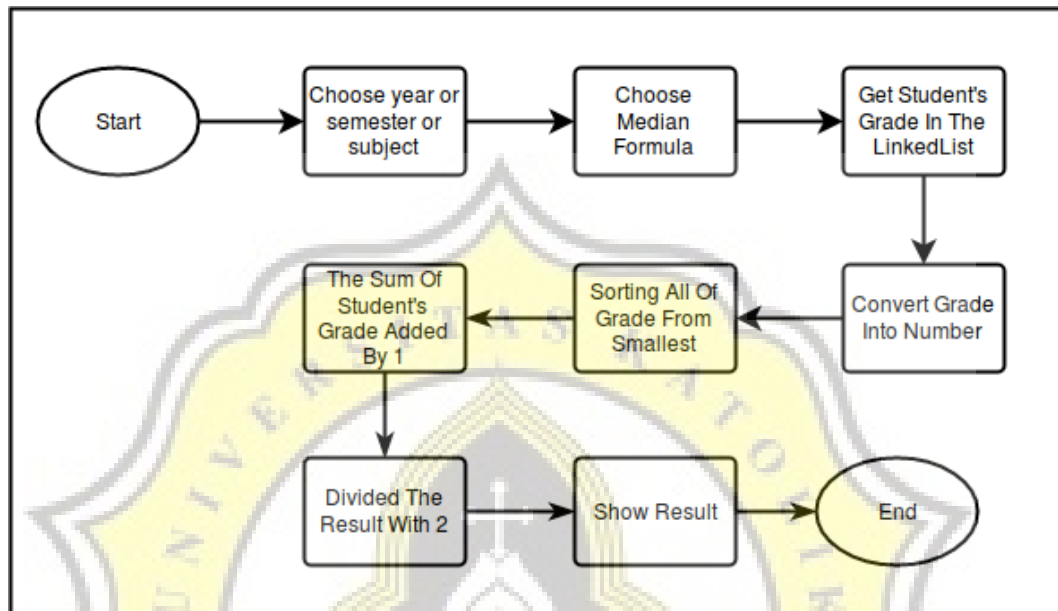


Illustration 6: Median Flowchart

The median formula is formula to calculate the middle value of student's grade. The illustration above show how the system get those middle value. The first step is user choose year or semester or subject that will be calculated, then user choose the median formula. Based on those year or semester or subject, the system will search data in the linkedlist. After that the system will convert the grade into number and sorting all of the grade from the smallest. After all the data is sorted, then the sum of student's grade is added by 1. Those result will be divided by 2. And the system will show the result.

3. Mode formula

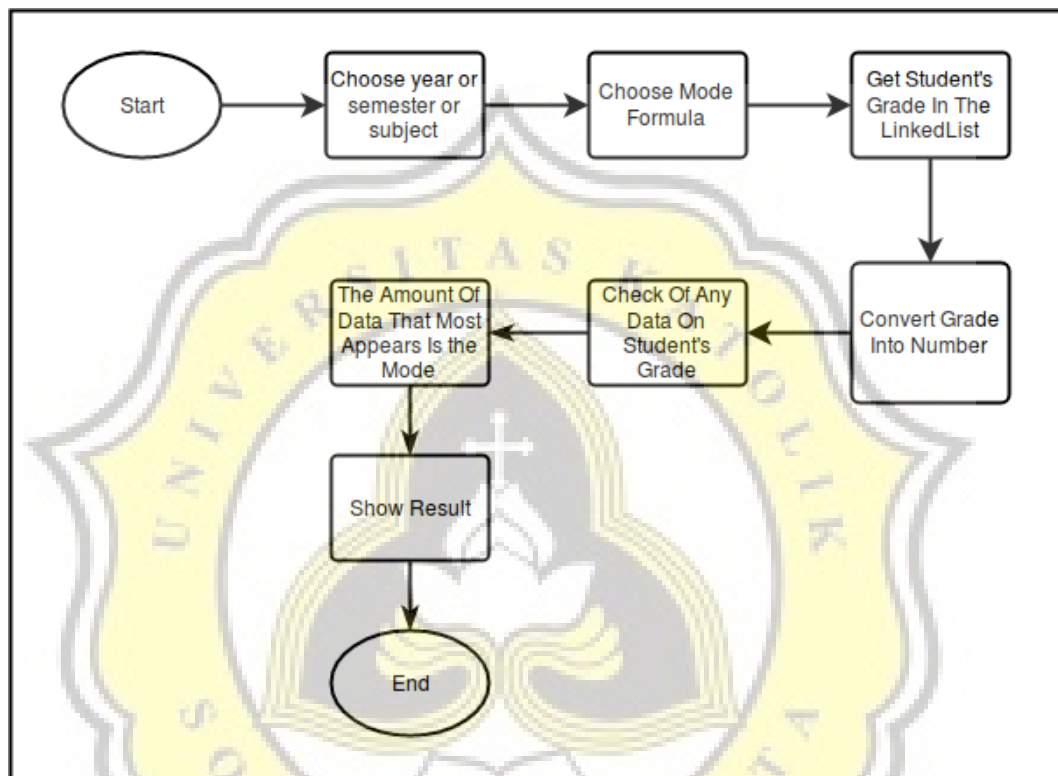


Illustration 7: Mode Flowchart

The mode formula is formula to know the value that often arise in student's grade. If the data often appear more than one, then the data does not have mode value. The illustration above show how the system get those value that often arise. The first step is user choose year or semester or subject that will be calculated, then user choose the mode formula. Based on those year or semester or subject, the system will search data in the linkedlist. After that the system will convert the grade into number and check of any data on student's grade. And the amount of data that most appears is the mode. And the system will show the result.

4. Standard deviation formula

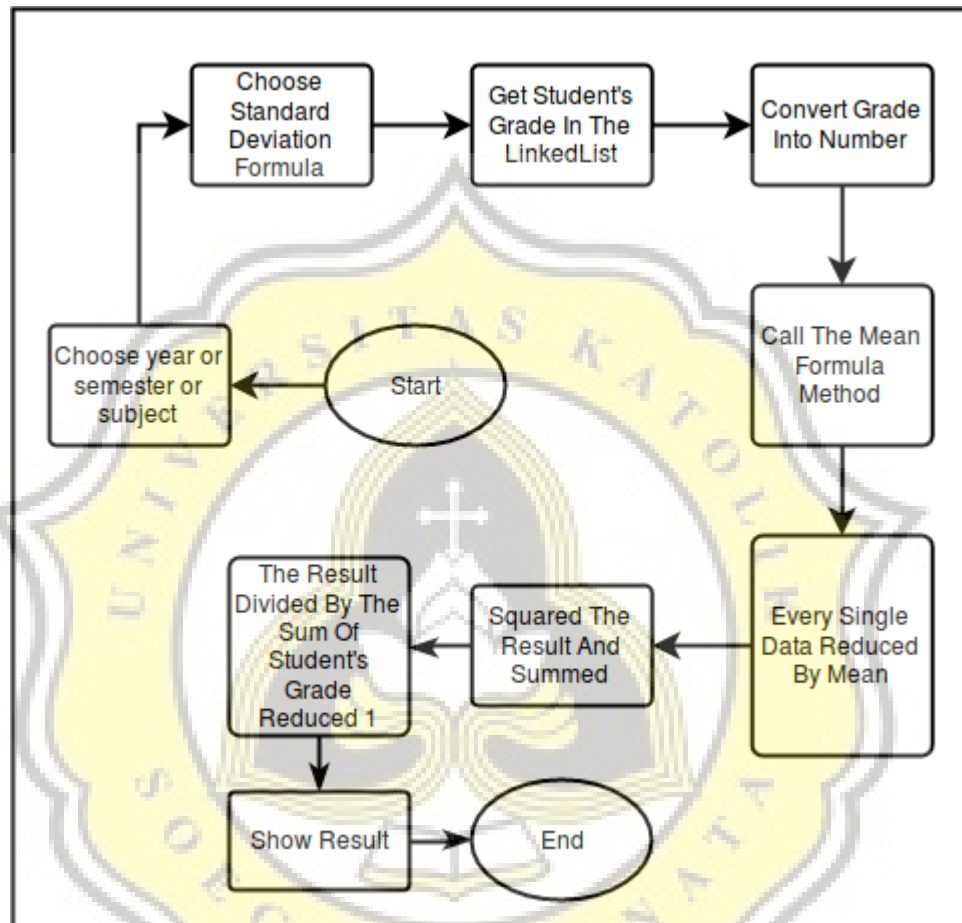


Illustration 8: Standard Deviation Flowchart

The standard deviation formula is formula to calculate the diversity of student's grade. The illustration above show how the system get those diversity. The first step is user choose year or semester or subject that will be calculated, then user choose the standard deviation formula. Based on those year or semester or subject, the system will search data in the linkedlist. After that the system will convert the grade into number and call mean formula method. Every single data will be reduced by mean. After that, those result will be squared and summed. Then it divided by the sum of the student's grade reduced by 1, and those result are square root. And the system will show the result.

5. Skewness formula

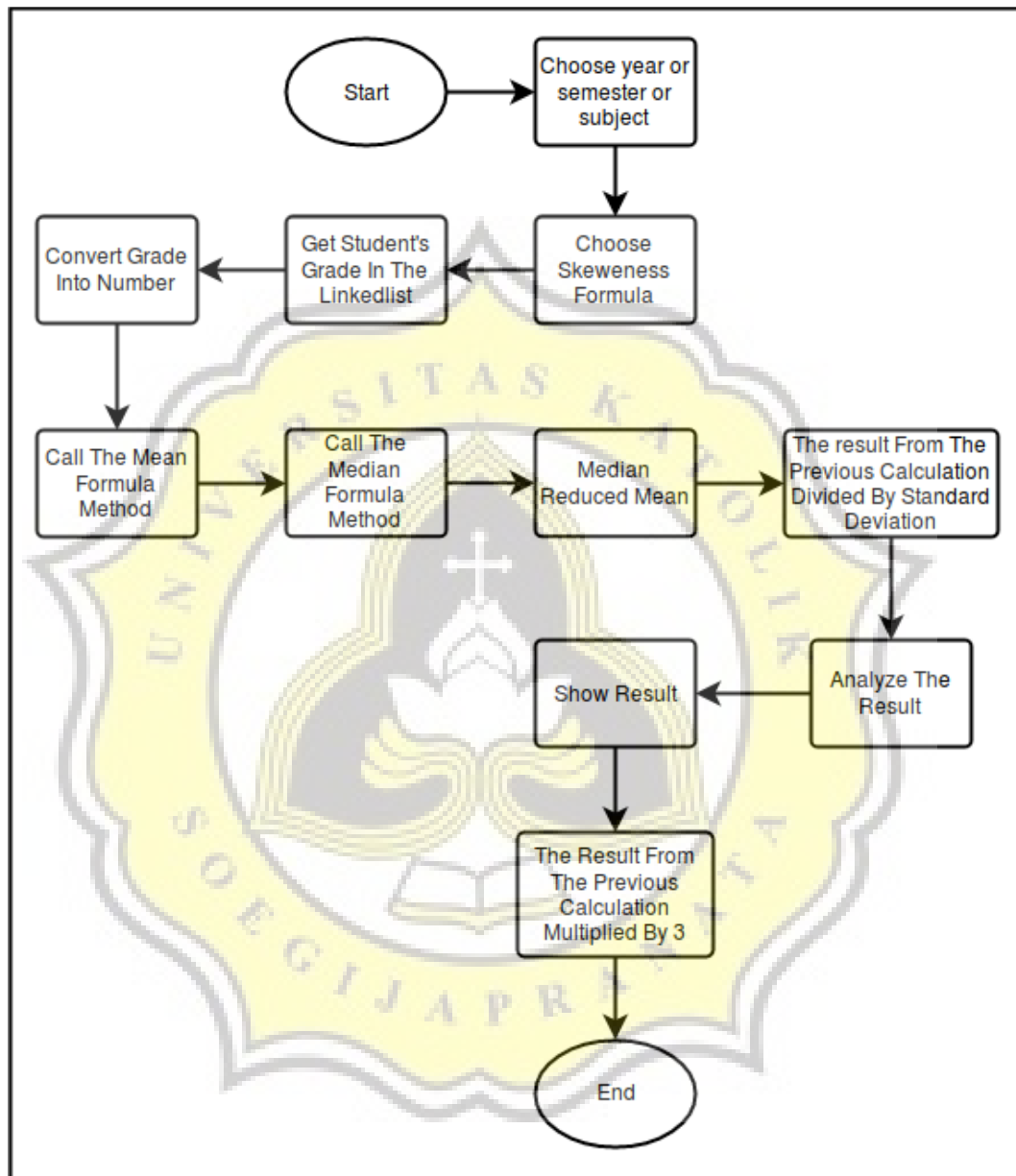


Illustration 9: Skewness Flowchart

The skewness formula is formula to calculate the tilt of student's grade. The illustration above show how the system get those tilt. The first step is user choose year or semester or subject that will be calculated, then user choose the skewness formula. Based on those year or semester or subject, the system will search data in the linkedlist. After that the system

will convert the grade into number and call median formula method. The median result will be multiplied by 3. After that, those result will be divided by standard deviation. Then the result will be analyze, if those result is 3 then the data is normal, and if the result is more then 3 then the data is positive, and if the result is less than 3 then the data is negative. And the system will show the result.

6. Kurtosis formula

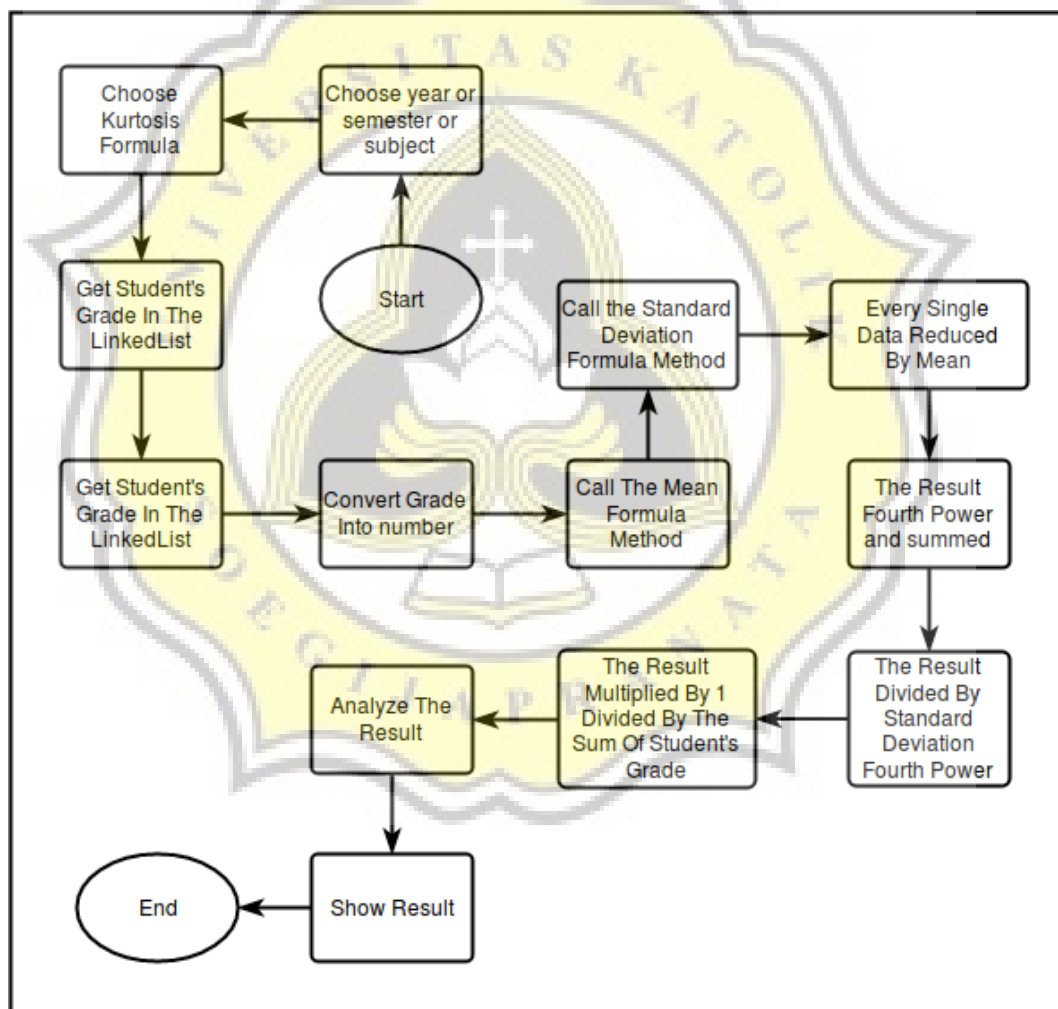


Illustration 10: Kurtosis Flowchart

The kurtosis formula is formula to calculate the kurtosis of student's grade. The illustration above show how the system get those kurtosis. The first step is user choose year or semester or subject that will

be calculated, then user choose the kurtosis formula. Based on those year or semester or subject, the system will search data in the linkedlist. After that the system will convert the grade into number and call formula method of mean and standard deviation. Every single data will be reduced by mean. After that, those result fourth power and summed. Then it divided by standard deviation fourth power. And those result will be multiplied by 1 and divided by the sum of student's grade.

Then the result will be analyze, if those result is 0 then the kurtosis is mezo curtic, and if the result is more then 0 then the kurtosis is leptocurtic, and if the result is less than 0 then the kurtosis is platycurtic. And the system will show the result.

The use case of “show graph visualization” in the illustration 4.2.1 show that user can choose show graphics menu, when the user choose that menu then the system will draw a graph with Javascript in CGI with C++. The graph will be represent in 2D. The graph will be represent the raw data.