

COMPUTER PROGRAM FOR STEM-AND-LEAF DIAGRAM

Nicolae EFTIMIE¹

Abstract: *The stem-and-leaf diagram represents a useful statistical process control tool. This graphical display provides information about the shape, variability and central tendency of the measurements' distribution. The paper presents a computer program that can be used for plotting three variants of the stem-and-leaf diagram. The first variant plots the regular stem-and-leaf diagram, the second one plots the diagram with the stems divided into two parts and the last one plots the diagram with the stems divided into five parts. The presented program can be used in scientific research and in plant practice for obtaining information about the measurements' distribution.*

Key words: *stem-and-leaf diagram, statistical process control, computer aided quality.*

1. Introduction

One important statistical process control tool is the stem-and-leaf diagram. This graphical representation is used to obtain information about the shape, variability and central tendency of the measurements' distribution [3].

An advantage of the use of the stem-and-leaf diagram is that the original data are not lost, as in the case of the histogram [1], [2].

Let's consider that a sample of size n is taken from a process, and the sample measurements are x_1, x_2, \dots, x_n .

In order to construct a stem-and-leaf diagram for the measurements, the following steps have to be followed [3]:

- Each measurement x_i will be divided into two parts: the stem, and the leaf. Let's consider that the measurement is made up of m digits. In this case, the stem will contain the first $m-1$ digits, and the leaf

will contain the last digit;

- The stem values will be recorded in a vertical column;

- The leaf that corresponds to each measurement will be recorded beside its stem.

Figure 1 presents a stem-and-leaf diagram.

Stem	Leaf
8	8786884997
9	552934163368545016261709978412869
10	01197261171027702748208110337520826482
11	4205382597228132
12	520

Fig. 1. *Stem-and-leaf diagram*

To get a better understanding of the measurements' distribution, each stem may be divided into two or five parts [3]. In the case when the stems are divided into two parts, the leaves 0-4 will be assigned to the lower part, and the leaves 5-9 will be assigned to the upper part.

¹ Centre "Advanced Technologies and Manufacturing Systems", *Transilvania* University of Braşov.

In the case when the stems are divided into five parts, the leaves 0-1 will be assigned to the first part, the leaves 2-3 will be assigned to the second part, and so on [3].

Figures 2 and 3 present the stem-and-leaf diagrams for the case when the stems are divided into two parts and into five parts respectively.

Stem	Leaf
8L	4
8U	878688997
9L	234133401210412
9U	559668556679978869
10L	011211102024201103320242
10U	97677778875868
11L	4203222132
11U	585978
12L	20
12U	5

Fig. 2. *Stem-and-leaf diagram with stems divided into two parts*

Stem	Leaf
8z	
8t	
8f	4
8s	767
8e	888899
9z	101101
9t	233322
9f	4445555
9s	6666776
9e	9899889
10z	0111110001100
10t	222233222
10f	445
10s	76777776
10e	98888
11z	01
11t	2322232
11f	455
11s	7
11e	898
12z	0
12t	2
12f	5
12s	
12e	

Fig. 3. *Stem-and-leaf diagram with stems divided into five parts*

The leaves that correspond to each stem may be ordered.

In this case, the plot is called an ordered stem-and-leaf diagram.

The diagram may also contain, in its left side, a column where it is recorded the number of measurements at and above each stem, in the upper half of the plot, and the number of measurements at and below each stem, in the lower half of the plot.

At the stem that corresponds to the median, the column indicates the number of observations at that stem [3].

Figure 4 presents an ordered stem-and-leaf diagram.

	Stem	Leaf
10	8	4677888899
43	9	0011112223334445556666778889999
(38)	10	00000011111112222223344566777788889
19	11	0122222334557889
3	12	025

Fig. 4. *Ordered Stem-and-leaf diagram*

2. Objectives

The application of the stem-and-leaf diagram is a time-consuming operation and requires special attention from the person who performs this activity.

The paper objective is to develop a computer program that makes easier the application of the ordered stem-and-leaf diagram.

Using the input data, which consist of a set of integer numbers, the program plots three variants of the stem-and-leaf diagram.

The first variant plots the regular stem-and-leaf diagram, the second one plots the diagram with the stems divided into two parts and the last one plots the diagram with the stems divided into five parts.

3. Material and Methods

The paper proposes a method of application for the stem-and-leaf diagram based on the use of a computer program.

The program was developed by means of MS Excel 2013 and Visual Basic for Applications for Excel 2013.

The application contains an input data area that consists of five columns, where any number of integer values can be recorded (Figure 5).

Data					
246	245	255	247	229	
253	251	251	246	258	Stem-and-Leaf Diagram
243	258	252	251	257	
235	263	247	241	246	Stem-and-Leaf Diagram: Stems divided into two parts
258	268	246	255	241	
252	261	251	257	255	
252	250	258	261	250	
243	242	254	260	245	Stem-and-Leaf Diagram: Stems divided into five parts
257	247	249	230	241	
243	272	232	254	261	
251	248	253	237	251	
249	264	250	246	253	
242	255	243	259	242	
227	259	255	247	249	
255	266	239	261	234	
253	249	249	232	245	
236	250	269	236	255	
259	262	259	249	267	
244	246	255	258	244	
250	243	243	250	246	
264	243	256	261	250	
240	243	238	242	234	
252	239	255	230	246	
260	250	264	245	244	
254	256	248	253	254	

Fig. 5. The input data area and the control buttons

To prevent the recording of values that are not integer numbers, a data validation rule was created for the cells from the input data area.

The interface of the program also contains three control buttons that are used for plotting the three types of diagrams: the regular stem-and-leaf diagram, the stem-and-leaf diagram with the stems divided into two parts and the stem-and-leaf diagram with the stems divided into five parts (Figure 5).

The paper explains the operation of the program by means of an example.

The input data for the presented example were obtained as follows:

- By means of the Excel's number generator it was generated a set of 125 normally distributed numbers, with the parameters $\mu = 250$ and $\sigma = 9$;
- Each number was rounded to the nearest integer value.

4. Results and Discussions

For each command button of the computer program, it was written a module of code, which was developed in Visual Basic for Applications for Excel 2013.

Each module consists of four procedures:

- A main procedure;
- A procedure for the stems;
- A procedure for the leaves;
- A procedure for the cumulative number of measurements.

Further, there are presented the three modules of the program.

4.1. The Module for Regular Stem-and-Leaf Diagram

The module *Diagr_V1* contains the procedures that are used by the program for plotting the regular stem-and-leaf diagram. These procedures are: *Stem_and_Leaf*, that represents the main procedure, *Stems_V1*, which calculates and plots the values of stems, *Leaves_V1*, which determines and plots the leaves that correspond to each stem and *Cumulative_Count_V1*, which computes and plots the cumulative number of measurements.

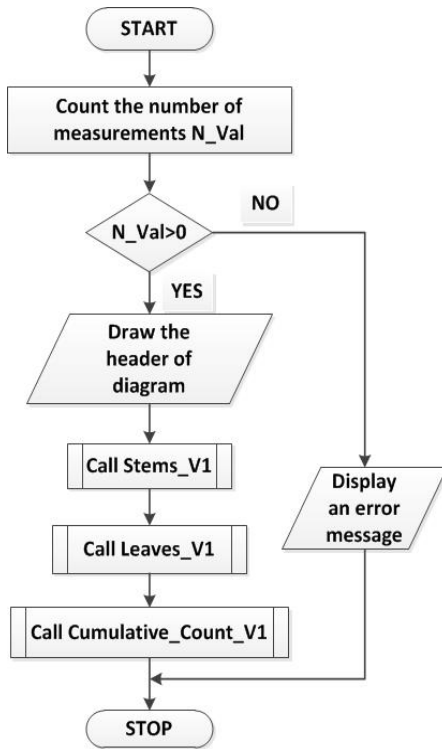
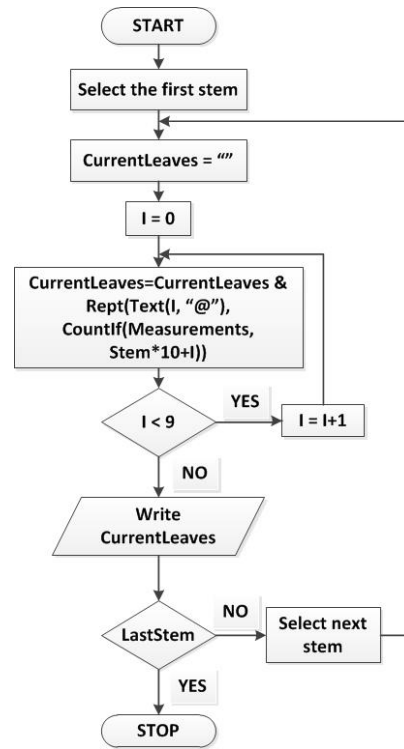
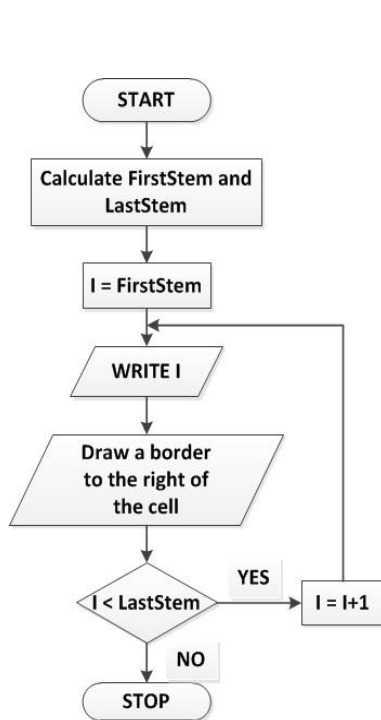
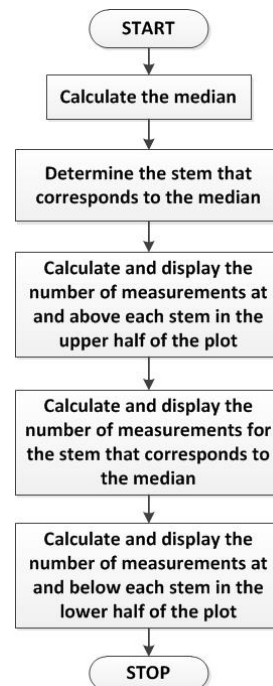
For plotting the regular stem-and-leaf diagram, presented in Figure 6, the user of the program must click the control button labelled "*Stem-and-Leaf Diagram*".

	Stem	Leaf					
2	22	79					
15	23	0022445667899					
58	24	011122223333333344455555666666667777889999999					
(50)	25	000000001111112222333333444455555555666777888889999					
17	26	0011111234446789					
1	27	2					

Fig. 6. The regular stem-and-leaf diagram

When this button is clicked, the main procedure *Stem_and_Leaf* is called.

Figures 7-10 present the algorithms for the procedures *Stem_and_Leaf*, *Stems_V1*, *Leaves_V1* and *Cumulative_Count_V1*.

Fig. 7. *The Stem_and_Leaf algorithm*Fig. 9. *The Leaves_V1 algorithm*Fig. 8. *The Stems_V1 algorithm*Fig. 10. *The Cumulative_Count_V1 algorithm*

4.2. The Module for Stem-and-Leaf Diagram with Stems Divided into Two Parts

The main procedure of the module *DiagrV2*, *Stem_and_Leaf_Div_Two*, is similar to the procedure *Stem_and_Leaf*.

In order to plot the diagram, this procedure calls the procedures: *Stems_V2*, *Leaves_V2* and *Cumulative_Count_V2*.

The diagram is presented in Figure 11.

	Stem	Leaf
0	22L	
2	22U	79
8	23L	002244
15	23U	5667899
34	24L	0111222233333333444
58	24U	55556666666677778899999
(27)	25L	00000001111111222233334444
40	25U	5555555566677788889999
17	26L	001111123444
5	26U	6789
1	27L	2
0	27U	

Fig. 11. The stem-and-leaf diagram with stems divided into two parts

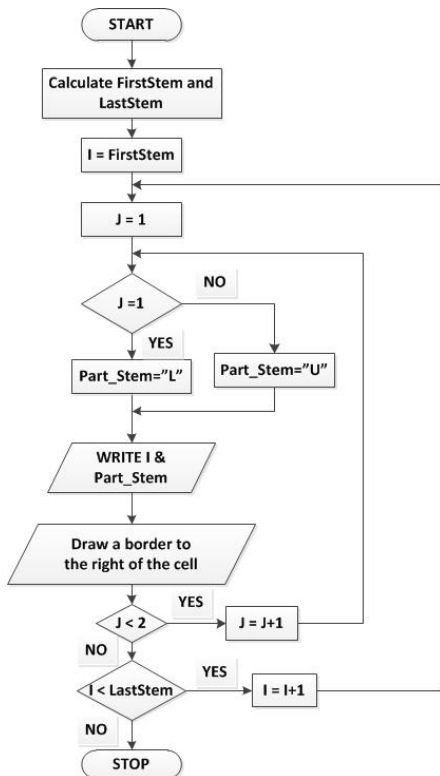


Fig. 12. The Stems_V2 algorithm

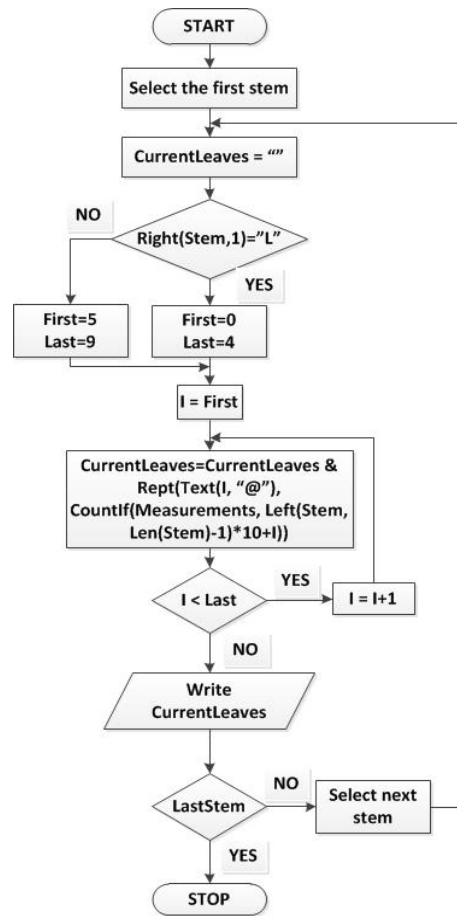


Fig. 13. The Leaves_V2 algorithm

Figures 12 and 13 present the algorithms *Stems_V2* and *Leaves_V2*.

The algorithm for the procedure *Cumulative_Count_V2* is similar to the algorithm for the procedure *Cumulative_Count_V1*.

4.3. The Module for Stem-and-Leaf Diagram with Stems Divided into Five Parts

The module *DiagrV3* contains the procedures used by the program for plotting the stem-and-leaf diagram with the stems divided into five parts.

The diagram (Figure 14) is plotted by means of the button “*Stem-and-Leaf Diagram: Stems divided into five parts*”.

	Stem	Leaf
0	22z	
0	22t	
0	22f	
1	22s	7
2	22e	9
4	23z	00
6	23t	22
9	23f	445
12	23s	667
15	23e	899
19	24z	0111
31	24t	222233333333
38	24f	4445555
50	24s	666666667777
58	24e	88999999
(14)	25z	00000000111111
53	25t	2222333333
44	25f	444455555555
31	25s	66777
26	25e	888889999
17	26z	0011111
10	26t	23
8	26f	444
5	26s	67
3	26e	89
1	27z	
1	27t	2
0	27f	
0	27s	
0	27e	

Fig. 14. The stem-and-leaf diagram with stems divided into five parts

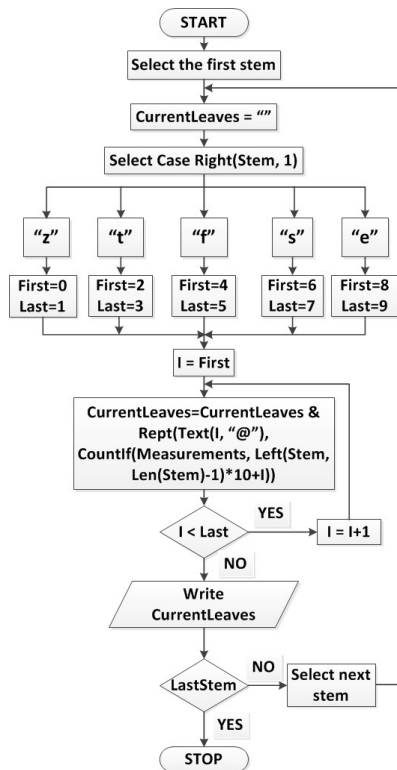


Fig. 15. The Leaves_V3 algorithm

In Figure 15 it is illustrated the use of the “Select Case” programming structure in the case of the *Leaves_V3* procedure.

The use of the structure was achieved in a similar manner for the other procedures of the module *DiagrV3*.

5. Conclusions

The presented program can be used for plotting three variants of the stem-and-leaf diagram. The user of the proposed program can chose the type of the diagram that provides the best information about the measurements’ distribution.

Thus, from the three variants, the user can select the regular stem-and-leaf diagram, the diagram with the stems divided into two parts and the diagram with the stems divided into five parts.

The program can be used in scientific research and in plant practice for obtaining information about the shape, variability and central tendency of the measurements’ distribution.

Acknowledgements

The author is thankful to the reviewers of this paper who analysed with competence the material.

References

1. Lepskyi, A.: *On the Stability of Comparing Histograms with Help of Probabilistic Methods*. In: *Procedia Computer Science* **31** (2014), p. 597-605.
2. Mitra, A.: *Fundamentals of Quality Control and Improvement, Third Edition*. New Jersey. John Wiley & Sons, Inc., 2008.
3. Montgomery, D. C.: *Applied Statistics and Probability for Engineers, Sixth Edition*. New Jersey. John Wiley & Sons, Inc., 2014.