

# Using Inductive approach to teach Statistics (Measures of Dispersion) with the aid of Microsoft Excel Spreadsheet

*F.H. Chen*

chen\_fang\_hao@moe.edu.sg  
Department of Mathematics  
Millennia Institute, Singapore

**Abstract:** This paper describes the usage of the inductive approach with the aid of Microsoft Excel in the teaching of statistical concepts such as Measures of Dispersion to engage students during mathematics lesson. In this approach, students are given tangible data in the form of real life data. With certain manipulation of the data, students are guided through proper scaffolding to observe certain patterns and trends. From here, hopefully, students can derive the appropriate mathematical formulae on their own. To manipulate the data with ease, Microsoft Excel spreadsheet is used. Excel spinners and macros are incorporated into the spreadsheet as means to control the data with ease. Data is plotted on animated bar charts and scatter plots to cater to students inclined to visual learning. Interaction between students and teacher is also increase as students can verify solutions with the Excel spreadsheet. The ease of data manipulation provides avenues for teachers to pose higher order thinking questions to the students. In addition, the flexibility and user friendliness of the Excel software also allow teachers to customize the spreadsheet according to the needs of the lesson e.g. by changing to the types of statistical graphs that they like or increasing the number of data sets. This method of teaching has shown to be effective in engaging the students both in the cognitive and the affective domains.

## 1. Introduction

The implementation of IT Masterplan 2 [1] was launched in July 2002 by the Singapore Ministry of Education. Its purpose is to integrate IT seamlessly into the curriculum to support a more student-centred learning environment, more autonomy in schools in terms of technology implementation, easy access to dynamic multimedia content and exploration of ubiquitous computing. Mathematics problem solving is central to the Singapore mathematics curriculum [2-3] and the appropriate use of IT can greatly aid the teaching of mathematical concepts and processes to the students in a more engaging manner [4-5].

Although the advantages of infusing IT into teaching are many, the problem of finding customized and affordable software is a prevalent issue in Singapore schools. Microsoft Excel software presents a good solution to this problem. With the implementation of IT Masterplan 1 [1] in Singapore, every teacher is equipped with a notebook installed with the latest Microsoft Excel version.

In this paper, a mathematics lesson using the inductive approach with the usage of Microsoft Excel will be described. Two Excel spreadsheets were designed to teach the statistical concept of measures of dispersion, namely range, standard deviation and variance according to the GCE 'A' levels Mathematics 8174 syllabus. The organization of the paper will begin in section 2 with a description of the reasons to use the inductive approach in the lesson. This is followed by a brief description of the design of the spreadsheets and the lesson structure in section 3. In section 4, the outcomes of administering the lesson to a class of 10 Pre-university 3 Arts students taking the GCE 'A' levels Maths 8174 subject and to another group of 40 teachers will be discussed. Section 5 will conclude this paper.

## **2. Motivation for using the Inductive approach using Microsoft Excel as a tool**

According to the constructivists' approach to learning [6-7], a student-centred environment must be created. To accomplish this, the students' prior knowledge must first be established. After this is done, students are guided by the teachers via questioning, IT tools and other pedagogical tools such as manipulatives to discover the mathematical concepts by themselves through linking the concepts to be taught to their prior knowledge as much as possible. As this is a student-centred teaching, the interaction among teacher and students will also increase. Therefore, this approach fulfills the affective and cognitive domains of learning.

The inductive approach to the teaching of mathematical concepts is part of the constructivists approach. According to Kline et al [8], "extracting the appropriate concept from a concrete situation, generalizing from observed cases, inductive arguments, arguments by analogy, and intuitive grounds for an emerging conjecture" are essential mathematical modes of thinking in the inductive approach. Therefore, in this lesson, students obtained many opportunities to manipulate raw data which they can associate in real life. Through controlled and guided manipulation of the data planned by the teacher, students will discover the patterns and trends behind the raw data and hopefully, deduce the mathematical formulae by themselves. This would result in greater retention by the students. Microsoft Excel spreadsheet offers user friendly features which allow easy manipulation of data. Moreover, the interactivity of the Excel spreadsheet and its colorful presentation adds color and zest to the lesson.

## **3. Spreadsheet design and lesson structure**

As mentioned in section 1, two Excel spreadsheets were created for different segments in the lesson with different specific learning objectives. In this section, they will be described separately..

### **3.1 First spreadsheet: Why do we need measures of dispersion?**

To use the inductive in the lesson, the weights of five characters (Tom, Dick, Harry, Jason and Victor) are used as raw data. These weight data are shown in Fig.1. Students are comfortable in manipulating weight data as they are real life entities which the students can relate to. In Fig.1, it can be seen that the weights of the different characters are displayed in different colored cells in the spreadsheet for easy visualization. These data are plotted on a bar chart with the colors of the bars corresponding to the respective data. This caters to students who prefer charts and figures than numbers. Spinners were inserted for each character for easy manipulation of the weight data. Each click will result in the increment or decrement of the weight by one. Prolonged click on the spinner will result in the increase and decrease of the weight data continually. The spreadsheet is highly dynamic. Any changes in the weight data initiated by the user will trigger corresponding changes in the bar charts and the mean cell (the different measures of dispersion will change accordingly shown in Fig.3).

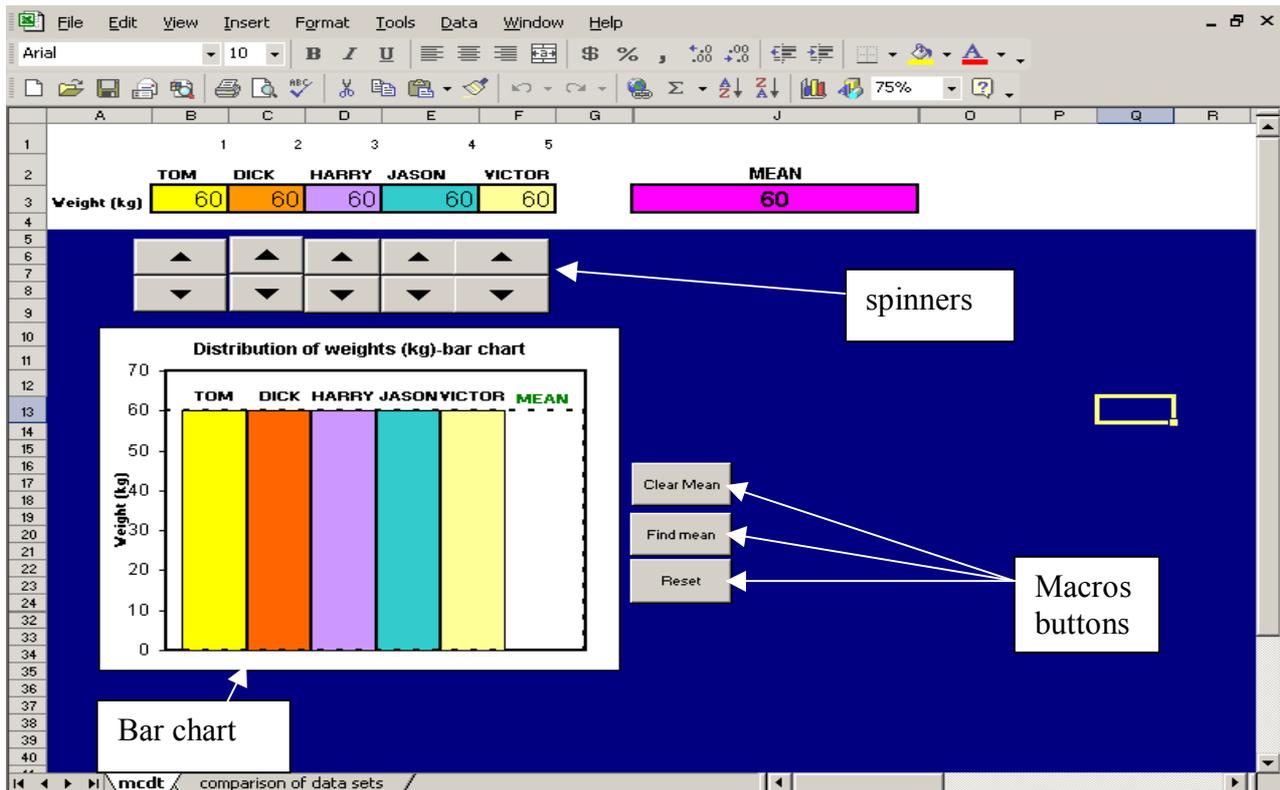


Fig.1: First spreadsheet

In addition, buttons were also included in the spreadsheet. Each button has a macros created for a specific function in the spreadsheet. The “Find mean” button will calculate the average of the weight data. The “Reset” button will reset all data to 60 kg as shown in Fig.1 and the “Clear mean” button will hide the value of the mean which is useful if the teacher wants the students to calculate the mean themselves.

The lesson begins by telling the students to key in three different weight data sets into the spreadsheet but with the same mean weight of 60 kg. Students could perform this by themselves and verify with the spreadsheet. These three data sets were then plotted on three bar charts shown in Fig.2 and students were asked this question: “If I say that the mean weight is 60 kg, can you tell me which data sets does it refer to?”

From this question, students will realize the insufficiency of the mean to describe data sets and the necessity of another measure which measure a different kind of characteristics of the data sets.

From observing the bar charts shown in Fig.2, students are asked to observe a trend in the data patterns. Putting the data patterns in bar charts helps students to see the patterns more easily. Students can then infer that as the data charts move in a clockwise fashion, the number of characters’ weight deviating from the mean weight increases. From here, students can then induce that this new measure could measure the deviation of data from the mean as the values are changing with the different data sets. Thus, the measures of dispersion (range, standard deviation and the variance) could be introduced formally here shown in Fig.3 with the relevant columns unhide in the spreadsheet. Students can then determine the values for the range, variance and standard deviation for the different data sets by manipulating the weight data using spinners or typing in the spreadsheet. By following this procedure, students are given concrete physical meaning to the statistical concept of dispersion. This will further motivate them to find the formulae to determine

the values. Note that up till this point, the formulae for the different measures of dispersion have not been revealed yet.

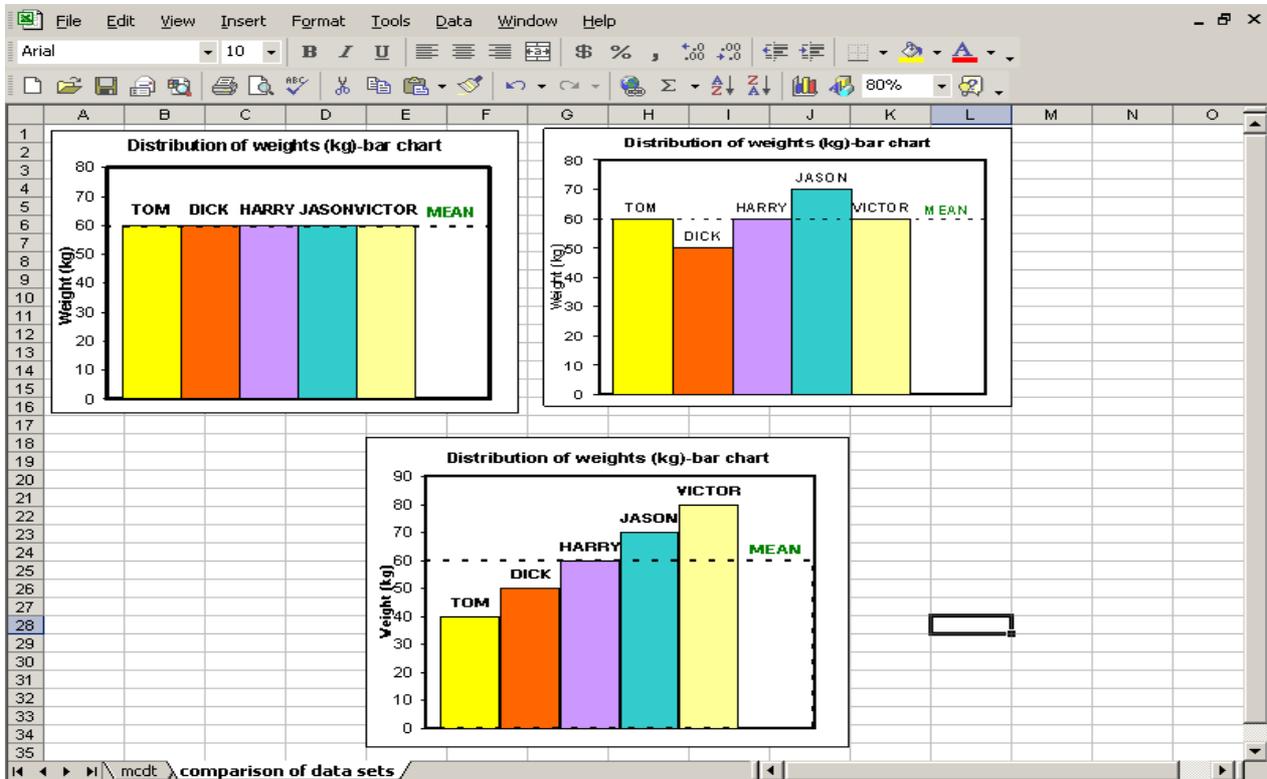


Fig.2: Bar charts representing the three different data sets with the mean of 60 kg indicated

To allow students to discover some salient properties of the range, standard deviation and variance by themselves, one of the character's weight (e.g. Jason's weight) is incremented slowly to a value (eg 100 kg) and then decreased slowly back to another value. This will cause the values of range, standard deviation and variance to change accordingly to a certain pattern. The values chosen must include zero deviation from the mean. Some properties that students can infer by themselves include:

- 1) Range, standard deviation and variance cannot be negative
- 2) Variance is the square of standard deviation
- 3) Range is calculated by taking the difference the highest value and the lowest value
- 4) Standard deviation and variance follows a parabola curve characteristics

Thus, from here, we can see that most of the salient properties of the measures of dispersion are already observed by the students through mere observation in the changes of the data. What's more, students have already found the formulae for range. This would not have been possible without Microsoft Excel. It is also noteworthy to mention that, in a sharing session with teachers, the non-Mathematics teachers can also deduce the properties by themselves.

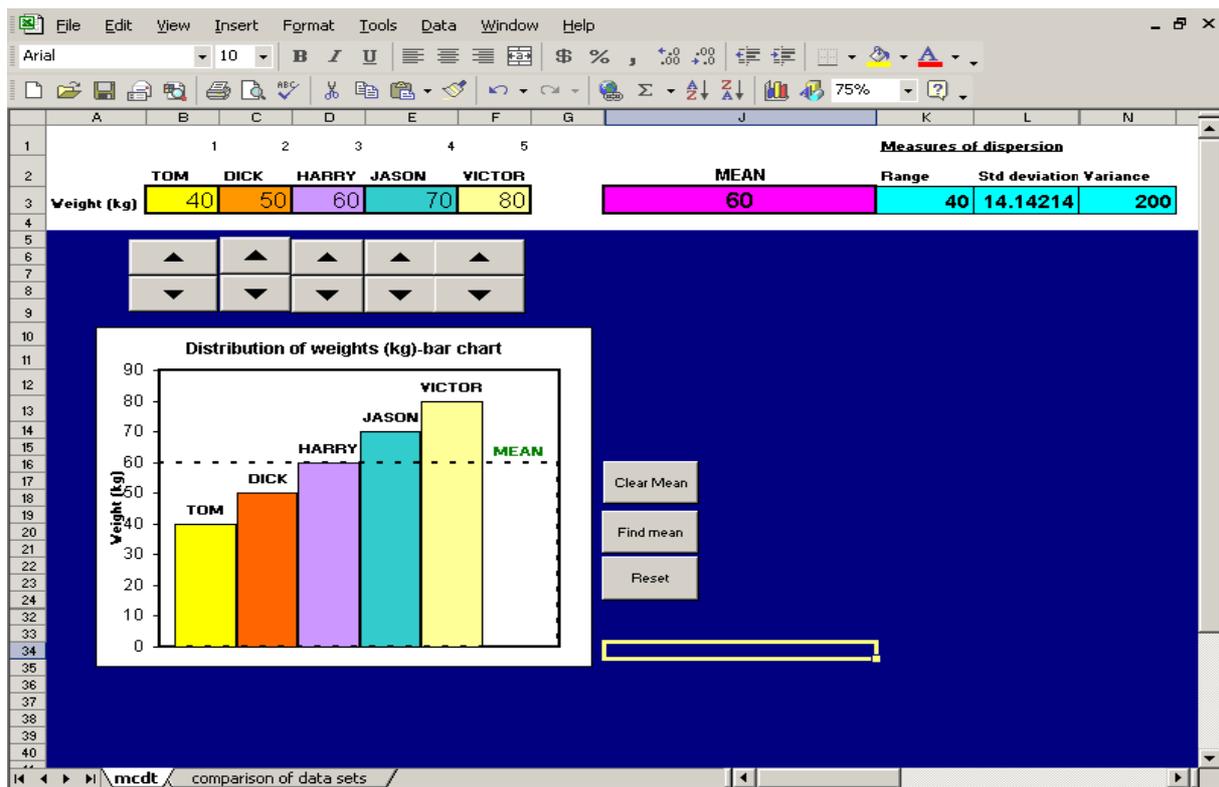


Fig.3: First spreadsheet with the measures of dispersion revealed with the relevant columns unhide

### 3.2 Second spreadsheet: Range, standard deviation and variance—which is the most commonly used? Why?

After the properties have been established, the next segment of the lesson is to determine which measure is the most commonly used. To have a good measure of dispersion, one criterion is the ability to differentiate different data sets appropriately. Students are tasked to find two data sets with the same mean (i.e. 60kg ) and the same range (e.g.20 kg). The spreadsheet can be used to aid students in discovering the two data sets. The two data sets can be found by tuning any other points beside the highest and lowest data values in such a way that the mean remains constant. Thus, students will know that range is not an adequate measure of dispersion. This leaves us with standard deviation and variance.

In the final segment of the lesson, students will be taught the formulae of standard deviation and variance using the inductive approach again. This will allow them to see that standard deviation is the most adequate and suitable measure of dispersion as compared to variance. To accomplish this, a second spreadsheet is designed with this purpose in mind as shown in Fig. 4.

In Fig.4, the spreadsheet consists of spinners and buttons. The buttons have the specific functions to build up to the formula of variance and standard deviation. In mathematical symbols, the functions of the different buttons and cells are described below:

- 1) “Sq(data-mean)”button and “(x-mean)^2” cell:  $(x - \bar{x})^2$
- 2) “x-mean”cell :  $(x - \bar{x})$
- 3) “clear sq(data-mean)” button: clears off data in the “(x-mean)^2” row

- 4) “sum(data-mean)” button:  $\sum(x - \bar{x})$
- 5) “sum sq(data-mean)” button:  $\sum(x - \bar{x})^2$
- 6) “ave sum sq(data-mean)” button:  $\frac{\sum(x - \bar{x})^2}{n}$
- 7) “Reset” button: clears off data and reset all weight data to 60 kg

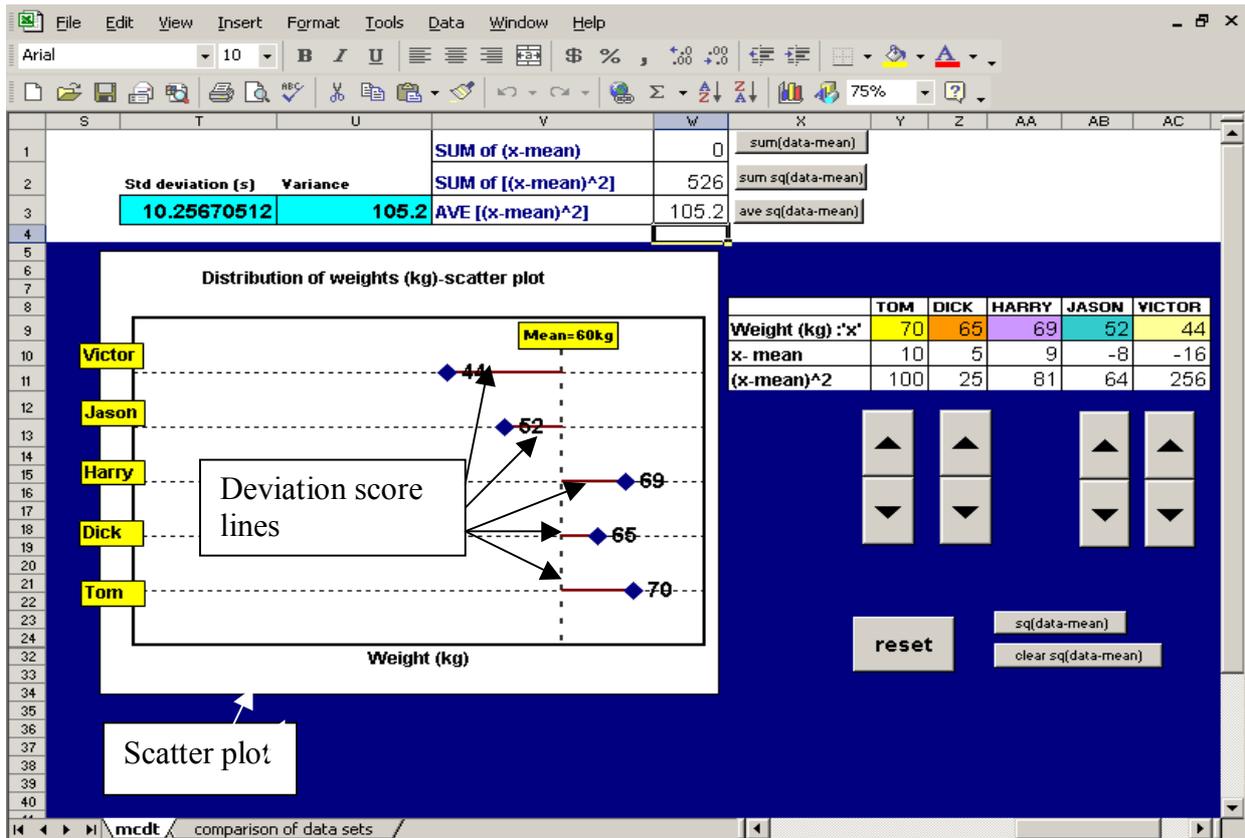


Fig.4: Second spreadsheet

The scatter plot is used to illustrate the concept of deviation score  $(x - \bar{x})$  using the red lines shown in Fig.4 with the mean of 60kg as the constant reference point.

Students are to key in one data set into the spreadsheet. The deviation scores  $(x - \bar{x})$  will be instantaneously displayed with the deviation scores displayed on the scatter plot. To find the representative value for the deviation scores, the average of the deviation score must be calculated.

However, students will discover that  $\sum(x - \bar{x}) = 0$  when the “sum(data-mean)” button is pressed. Students can press the different spinners in the spreadsheet to increase or decrease the data values (Harry’s weight is controlled by the computer to keep the mean at 60 kg). Higher-thinking questions could be posed to the students to explain the reason for  $\sum(x - \bar{x}) = 0$ . Students can infer

from the scatter plot that the sum of the deviation scores (indicated by the red lines) below the mean is equal to the sum of the deviation scores above the mean.

To resolve the problem of  $\sum (x - \bar{x}) = 0$ , the deviation scores are squared to remove any negative values. The steps are performed by pressing “Sq(data-mean)” button, “sum sq(data-mean)” button and then “ave sum sq(data-mean)” button. The final value obtained is equal to the variance (the blue cell beside it). Thus, students can infer that the variance is  $\frac{\sum (x - \bar{x})^2}{n}$  and from section 3.1,

standard deviation is  $\sqrt{\frac{\sum (x - \bar{x})^2}{n}}$  as they already knew that the variance is the square of standard deviation. From here, students can conclude that standard deviation is more often used than variance because the units for standard deviation is consistent with the data i.e. kg while the units for variance is kg<sup>2</sup>.

#### 4. Outcomes of implementation of lesson

As mentioned in section 1, this lesson was taught to a class of 10 Pre-university 3 students. In addition, this was also taught to a group of 40 teachers from junior colleges, secondary schools and curriculum planning unit under a simulated classroom lesson experience during the Ministry of Education Excel Festival [9].

##### 4.1 Outcomes of sharing to a group of 40 teachers

Evaluation forms were given at the end of the simulated lesson to the group of 40 teachers to gather feedback about the simulated classroom lesson. The results of the evaluation are shown in Fig 5 below.

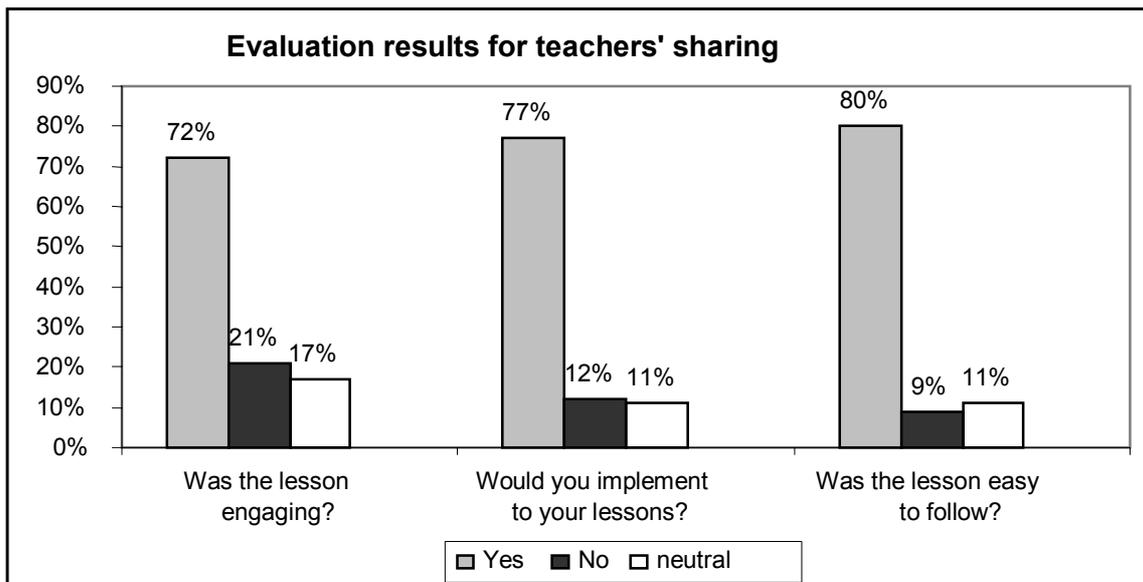


Fig.5: Evaluation results of teachers' sharing

From Fig.5, majority of the teachers have feedback that the lesson was engaging, easy to follow and that they will implement it to their lessons. Many commented that they found it easier to understand the concepts as compared to the conventional deductive approach. Specifically, they have commented that the interactivity between the students, spreadsheet and the teacher greatly engaged the students to the lesson. Teachers commented that they were impressed with the spreadsheets and sought to improvise the spreadsheets to meet the learning needs of their students.

Also, the manipulation of data to allow students to deduce the formulae also gave a deeper meaning to the mathematical formulae. However, they commented that the process might be time consuming and for teachers who are not proficient in Microsoft Excel, they may encounter problems improvising the spreadsheet to tailor to the needs of the students.

## **4.2 Outcomes of implementation to students**

After the lesson was taught to the students, verbal feedback was gathered. Most students commented that the lesson were fun, interactive and engaging. Students commented that the use of weight data helps them to understand the concepts much better as they can give physical meaning to the formulae. The interactivity between the students and the spreadsheet engaged them to the lesson. Thus, students found it more meaningful and motivating to engage themselves in more drill and practice to gain mastery of the formulae later after this lesson. The feedback was found to be coherent with the teachers' feedback mentioned in section 4.1.

Personally, I found it a very enriching experience. What was most satisfying was the enthusiasm shown by teachers and students alike when the lesson was taught. This has encouraged me to explore more possibilities in using more IT tools in the teaching of Mathematics in the future.

## **5. Conclusion**

This paper describes the teaching of range, standard deviation and variance using the inductive approach with the aid of Microsoft Excel. Two spreadsheets incorporating spinners and macros were created. Students were given physical meaning to statistical concepts and formulae by controlled manipulation of the weight data. Feedback given by students and teachers has ascertain the fact that the inductive approach fulfills the cognitive and affective learning domains of the students.

### **Acknowledgements:**

The author would like to thank his colleagues of Millennia Institute for their valuable comments and input. The author would also like to thank his students of 04A3 for giving him the opportunity to teach this lesson to them.

## References

- [1] Koh T.S., Y.C. Koh and Wu W.T. (2004). Integration of Information Technology in Singapore School Mathematics Curriculum. *Proceedings of the Ninth Asian Technology Conference in Mathematics*, National Institute of Education, Singapore.
- [2] Lim, S.K. (2002). Mathematics Education Within the Formal Singapore Education System: Where Do We Go From Here? *EARCOME 2002 Proceedings*, Vol.1, Plenary and Regular Lectures, Singapore, 29-37.
- [3] Curriculum Planning and Development Division (CPDD) (2001). *Lower Secondary Mathematics syllabus: Special/Express, Normal Academic and Normal Technical*. Singapore: Ministry of Education [Online]. Available: [http://www1.moe.edu.sg/syllabuses/doc/Maths\\_LowSec.pdf](http://www1.moe.edu.sg/syllabuses/doc/Maths_LowSec.pdf) [2004, 4 Oct].
- [4] Alagic, M (2003). Technology in the Mathematics Classroom: Conceptual Orientation. *Journal of Computers in Mathematics and Science Teaching*, 22(4), 381-399.
- [5] Hung, D (2002). Situated Cognition and Problem-Based Learning: Implications for Learning and Instruction With Technology. *Journal of Interactive Learning Research*, 13(4),393-414.
- [6] Davis,A., *Constructivism*. In Davis, A. and Pettitt, D. (Ed) *Developing Understanding in Primary Mathematics*, The Falmer Press (1994).
- [7] Glasersfeld E.V., *A Radical Constructivists View of Basic Mathematical Concepts*. In Ernest, P. (Ed), *Constructing Mathematical Knowledge: Epistemology and Mathematics Education*, The Falmer Press (1994).
- [8] Kline, M. (1962). *On The Mathematics Curriculum Of The High School*. The Mathematics Teacher of March 1962. American Mathematical Monthly of March 1962 [Online]. Available: <http://michel.delord.free.fr/kline62.html> [2005].
- [9] Ministry of Education (MOE) Festival 2006 : “*Touching Hearts, Engaging Minds*”. Details can be obtained online at: <http://www.moe.gov.sg/excelfest/2006/index.htm>