Descriptive statistics in ICT education: A practical and real-life perspective

William Phiri

Zambia Research and Development Centre (Zrdc) and ICU School of Education, P.O. Box 30226, Lusaka, Zambia

Abstract
My experience of teaching and lecturing Mathematics and Statistics for more than 15 years from secondary school up to university level, it has become evident that statistical courses have always been taught in isolation of real life and practical problems. In this regard, many pupils, students and the general Zambian populace have found very little relevance of statistics to their real life experiences as it is mainly taught in abstract terms. Additionally, the emphasis and emerging of ICTs in almost all sectors of life including education has worsened the negative perception people develop over the application of statistics in real life situations. This paper argues that a clear understanding of statistics remains critical in the acceleration of a problem–solving ICT education. Moreover, the Ministry of General Education in Zambia has clearly stated in its policies that ICTs and E-learning are mandatory in all the schools. This discourse therefore aims to present descriptive statistics from the practical perspective and show the role it plays in ICT education. The paper will further demonstrate how descriptive statistics could be used to solve practical situations from various fronts.

Keywords: Descriptive statistics, Problem – solving ICT education

1. Introduction
Basic definitions, descriptions and explanation of statistical terms
1.1 Statistics: Statistics is defined as a scientific study of handling quantitative information (processed data). It is a methodology of collection, classification, description and interpretation of data obtained through the conduct of surveys and experiments (Aggarwal, 2004) [2]. Therefore, the main purpose of statistics is to describe and draw inferences about the numerical properties of populations.

1.2 Descriptive Statistics: Descriptive Statistics is a procedure for organizing, summarising and describing quantitative data about samples or about the populations where complete population data are available. It does not involve the drawing of an inference from a sample to its population (Mumba, 2008) [7]. A very good example of descriptive statistics is measures of central tendency (mean, mode, median) and measures of variability (range, variance, standard deviation).

1.3 Inferential Statistics: Inferential statistics makes inferences (conclusions) from small (sample) to large (population) groups of subjects or events on the basis of probability theory (Chimedza, Chipoyera and Mupambireyi, 2014) [3] and (Kasonde – Ng'andu) [4]. It should be noted that inferences about population drawn from sample measure may involve some error or discrepancy, the magnitude of which can be estimated on the basis of the probability theory.

1.4 Variable: A property or characteristics on which the members of a group or set differ from one another (Mann, 1992) [6]. For example Sex, age, height, weight, intelligence, socio-economic status. It is important to note that constant is the opposite of a variable, signifying the condition that the members of a group do not differ among themselves on a given property.

1.5 Scale of Measurement: According to Stattreck (2016) [11], a measurement involves the assignment of a value and defines the nature of values.

1.5.1 Characteristics of scale of measurements
- Identity: Values have unique meanings
- Magnitude : Values can be ranked ordered
- Equal interval: Scales units are equal
- Absolute Zero Point: Smallest value is zero

1.5.2 Types of scale of measurements
- Nominal scale   Less Precision
- Ordinal scale
- Interval scale
- Ratio scale   More Precision

1.6 Nominal Scale: Nominal variables allow for only qualitative classification. That is, they can be measured only in terms of whether the individual items belong to some distinctively different categories, but we cannot quantify or even rank order those categories. For example, all we can say is that 2 individuals are different in terms of a certain variable (for example, they are of different race), but we cannot say which one has more of the quality represented by the variable (Aggarwal, 2004 [2]; Stattreck, 2016 [11]). It is clear from the above that a nominal scale:
- Allow the least precision
- Values are verbal labels
- Values are not numbers

A Good and typical examples of nominal variables are gender, race, colour, city, marital status, etc. A practical example is the

gender information we solicit from respondents when answering a questionnaire that can be as follows:

       1. Male   2. Female

It is clear, the numbers associated with the options above have no numerical significance so that, not only comparison between values is impossible, but also descriptive statistics like the mean and standard deviation would make no sense if calculated.

1.7 Ordinal Scale: Ordinal variables allow us to rank order the items we measure in terms of which has less and which has more of the quality represented by the variable, but still they do not allow us to say how much more(Neter, 1988) [8] and (Chimedza, *et al*, 2014) [3]. A typical example of an ordinal variable is the socioeconomic status of families. It can be deduced therefore that an ordinal scale mainly indicate rank order where identity and magnitude are taken into consideration but run short of equal interval and absolute zero are absent.

1.8 Interval Scale: Interval variables allow us not only to rank order the items that are measured, but also to quantify and compare the sizes of differences between them. However, interval scale variables do not have an absolute zero. It may be argued in summary that interval scale supports arithmetic operations.

1.9 Ratio Scale: Ratio variables are very similar to interval variables; in addition to all the properties of interval variables, they feature an identifiable absolute zero point. Typical examples of ratio scales are measures of time or space. Height is a ratio scale variable since, if a person is twice as tall as another, he/she will remain so, irrespective of the units used (centimetres, inches, etc…). In a nutshell, ratio scale can be said to be similar to interval scale and its lowest value is zero such as weight.

1.10 Population and Sample

Population	Sample
− Collection of all individuals, items or data under consideration in a statistical study (Kasonde – Ng'andu, 2013)[4] − A target group from which the researcher wishes to draw responses from as postulated by Leave (2005) [5] in Phiri (2015) [10].	− Small subset of the population − Selected representation of a larger population according to Tackman (2006) [12] in Phiri (2015) [10] − We are usually interested in populations − But because we cannot collect data from every human being or whatever, we collect data from a sample and use these data to infer things about the population as a whole

1.11 Mean: A hypothetical estimate of the 'typical' score. It is a single statistical model of the centre of a distribution of scores**.** We use the variance or standard deviation to tell us whether the mean is representative of our data**.** Standard deviation is a measure of how much error there is associated with the mean**.** A small standard deviation indicates that the mean is a good representation of our data.

1.12 Standard Deviation and Standard Error
1.12.1 Standard Deviation: Tells us how much observations in our sample differ from the mean value within our sample

1.12.2 Standard Error: Tells us how much the sample mean represents the population mean. It is the standard deviation of the sampling distribution of a statistic. For a given statistics (e.g. Mean) it tells us how much variability there is in this statistic across samples from the same population. Large values, therefore, indicate that a statistic from a given sample may not be an accurate reflection of the population from which the sample came from.

1.13 Degrees of Freedom (df): Almost all tests of significance require the calculation of degrees of freedom. It is a mathematical concept**.** The geometric interpretation of the concept relates to the movement of a point in relation to the number of dimensions it is attached with. A point on a line is free to move in one dimension only and thus has 1 degree of freedom. A point on a plane has 2df. Similarly, a point in space has 3df. When a sample statistic is used to estimate a parameter, the number of degree of freedom depends upon the number of restrictions placed upon the scores, each restriction reducing one df (Aggarwal, 2004)[2].

Degrees of freedom can be demonstrated as follows:

No	Set A Given Scores	Set B Altered Scores	
1	8	10	
2	7	6	Can vary in any way
3	6	12	
4	10	13	
5	19	9	
∑X	50	50	value gets fixed

Source: Ibid [2]

In set A, original scores have been presented. We can vary the first four scores, 1-4, in any way we like. But the value of the 5th score gets fixed due to the restriction that ∑X in each case must be 50. Hence in this situation the four scores out of five are free to vary and hence the df=5-1=4.Thus in this case, the number of df depended upon the number of scores minus the number of restrictions.

1.14 Significance Level and Confidence Intervals
Experimenters and researchers have selected some arbitrary standards called levels of significance (α) to serve as the cut-off points or critical points along the probability scale, to separate the significant difference from the non-significant difference between the two statistics, like means or SDs. Generally, the 0.05 and the 0.01 levels of significance are the most popular in social sciences research. The confidence Interval with which an experimenter rejects or retains a null hypothesis depends upon the level of significance adopted. The table below illustrates more on levels of significance (α) and Confidence Intervals (CI).

Levels of Confidence	Confidence Intervals	Interpretation
0.05	95%	If the experiment is repeated a 100 times, only on five occasions the obtained mean will fall outside the limits μ±1.96
0.01	99%	If the experiment is repeated a 100 times, only on 1 occasion, the obtained mean will fall outside the limits μ±2.58

The values 1.96 and 2.58 have been taken from the t- tables keeping large samples in view. The 0.01 level is more rigorous and higher a standard as compared to the 0.05 level and would require a larger value of the critical ratio for the rejection of the H_0. Hence if an obtained value of t is significant at 0.01 levels, it is automatically significant at 0.05 levels but the reverse is not always true.

1.15 Hypothesis Testing: A statistical hypothesis is an assumption about a population parameter. This assumption may or may not be true. Hypothesis testing therefore refers to the formal procedures used by statisticians to accept or reject statistical hypotheses. The best way to determine whether a statistical hypothesis is true would be to examine the entire population. Since that is often impractical, researchers typically examine a random sample from the population. If sample data are not consistent with the statistical hypothesis, the hypothesis is rejected. Basically two types of statistical hypothesis can be discussed at this level namely null and alternative. The null hypothesis, denoted by H_0, is usually the hypothesis that sample observations result purely from chance (Wegner, 1993) [13]. The alternative hypothesis, denoted by H_1 or H_a, is the hypothesis that sample observations are influenced by some non-random cause. For example, suppose we wanted to determine whether a coin was fair and balanced. A null hypothesis might be that half the flips would result in Heads and half, in Tails. The alternative hypothesis might be that the number of Heads and Tails would be very different. Symbolically, these hypotheses would be expressed as follows:

H_o: P = 0.5
H_1: P ≠ 0.5

1.16 Steps in Testing Hypothesis
1.16.1 State the hypotheses: This involves stating the null and alternative hypotheses. The hypotheses are stated in such a way that they are mutually exclusive. That is, if one is true, the other must be false.

1.16.2 Formulate an analysis plan: The analysis plan describes how to use sample data to evaluate the null hypothesis. The evaluation often focuses around a single test statistic. The analysis plan includes decision rules for rejecting the null hypothesis. In practice, statisticians describe these decision rules in two ways - with reference to a P-value or with reference to a region of acceptance.

1.16.3 Probability Value (P-Value): The strength of evidence in support of a null hypothesis is measured by the P-value. Suppose the test statistic is equal to *S*. The P-value is the probability of observing a test statistic as extreme as *S*, assuming the null hypothesis is true. If the P-value is less than the significance level, ($p<α$) we reject the null hypothesis.

1.16.4 Region of acceptance: The region of acceptance is a range of values. If the test statistic falls within the region of acceptance, the null hypothesis is not rejected. The region of acceptance is defined so that the chance of making a Type I error is equal to the significance level.

1.16.5 Region of Rejection: The set of values outside the region of acceptance is called the region of rejection. If the test statistic falls within the region of rejection, the null hypothesis is rejected. In such cases, we say that the hypothesis has been rejected at α level of significance. These approaches are equivalent. Some statistics texts use the P-value approach**;** others use the region of acceptance approach.

2. Measures of Central Tendency and Dispersion (Variability) from Practical Perspectives
2.1 Mean: It is also called the arithmetic mean which is defined as the sum of all scores divided by the total number of scores. The mean is the most popular and significant of all numerical descriptive measurements (Mumba, 2008) [7]. The mean possesses many properties which make it significant.

General Formula for Mean

Thus $\bar{X} = \dfrac{Sum\ of\ all\ scores}{Total\ number\ of\ scores} = \dfrac{\sum x}{n}$

Where $\bar{X}$ = mean,
$\sum$ = Summation,
X = variable,
n = number of scores

Note: $\bar{X}$ is used to denote the sample mean and μ to denote the population mean.

2.1.1 Mean from Ungrouped or Raw Data
Example 1:
Given $X_1 = 4$, $X_2 = 5$, $X_3 = 9$, $X_4 = 18$, $X_5 = 14$.
Compute the mean.

Solution

$$\bar{X} = \frac{\sum x}{n} = \frac{4+5+9+18+14}{5} = \frac{50}{5} = 10$$

2.1.2 Mean from Grouped or aggregated Data

$$\bar{X} = \frac{\sum fx}{\sum f}$$

Example 2
An ICT teacher at a named Zambian rural school recorded her daily use (in minutes) of a computer aided tablet in the teaching of Mathematics on 50 randomly selected working days.

5	32	7	31	26	35	35	45	47	10
14	10	15	25	8	18	28	23	33	23
22	16	24	7	26	28	16	20	24	8
15	18	40	34	17	9	34	37	21	22
6	23	11	17	42	16	14	9	24	25

Group the data into classes 4 – 12, 13 – 21, 22 – 30 etc, hence calculate the mean minutes of the teacher using the computer tablets in the teaching of mathematics.

Solution

Class Interval	Frequency
4-12	11
13-21	13
22-30	14
31-39	8
40-48	4

$$\bar{X} = \frac{\Sigma f(x)}{\Sigma f} = \frac{1129}{50} = 22.58$$

2.1.3 Assumed Mean: The long method of calculating mean is an accurate and straightforward method. However, it is very often involves the handling of large numbers and requires tedious calculations. Hence, to overcome these difficulties, the 'Assumed Mean' method or simply the short method has been devised for the computation of mean from the frequency distribution.

2.1.4 Strengths of Mean
- All data is used to find the answer
- It is a more accurate estimate as compared to the median and mode.
- It can easily be calculated
- It is a well known and commonly used statistic
- It is more efficient estimate as compared to median and mode.
- It can easily be manipulated in order to compute other useful statistical measures.
- It is subject to less error

2.1.5 Weaknesses of the Mean
- It is discouraged to use the mean when data is clustered at the very low or very high end.
- It cannot be read from a graph
- A few extreme scores can cause distortion which makes the mean not representative of the given data set.
- Lastly but not the least, when the data is discrete, it may produce a value which appears to be unrealistic.

2.2 Median: The median of a data set is the middle value of the data set. It is that value such that at least 50 percent of the observations are less than or equal to it and at least 50 percent of the values are greater than or equal to it (Chimedza, Chipoyera and Mupambireyi, 2014) [3]. It is further observed by Mumba (2008) [7] that median of a set of scores is the middle value when the scores are arranged in either ascending or descending order. Putting it simply, the median is the value of the item which lays exactly half way along the data set arranged in order of size. After first arranging the original data set in order of size, the median will be either of the following:
- If the number of observations is odd, the median is the number that is exactly in the middle of the list.

- If the number of observations is even, the median is found by computing the mean of the two middle numbers.

2.2.1 Median from Raw Data
Steps:
- Arrange the data in array
- Determine the position of the median using the formula
$$\text{Median Position} = \frac{n+1}{2}$$
- Read the value of the medium from the number list.

Example 3
Given $X_1 = 4$, $X_2 = 5$, $X_3 = 9$, $X_4 = 18$, $X_5 = 14$, $X_6 = 5$. Calculate the median.

Solution
- Data is even
- Rearranging in ascending order gives 4,5,5,9,18,14
- Median position $= \frac{n+1}{2} = = \frac{6+1}{2} = = \frac{7}{2} = 3.5^{th}$ position
- Median is the average of the 3rd and 4th values $= \frac{5+9}{2}$
$= \frac{14}{2} = 7$

2.2.2 Median from Grouped Data
Mumba (2008) [7] furthermore postulates that the median can be determined either graphically or by calculation. As regards to grouped data, we are unable to determine where the true middle value falls, but we can estimate the median by using a formula and assuming that the median value will be $\frac{n}{2}$ value.

MEDIAN = L + [n/2 – fm-1] x C
                    Fm

Where
   L = lower boundary of the median class
   Fm-1 = sum of the frequencies up to but not including the median class i.e cumulative frequency immediately prior to median class
   Fm = frequency of the median class
   C = class width or interval
   N = Σf

From example 2,
Calculate the median score
The median class contains the

$$\frac{n^{th}}{2}, \text{ i.e. } \frac{50}{2} = 25^{th} \text{ observation}$$

$\Rightarrow$ Median class is 21.5 – 30.5

MEDIAN = L + [n/2 – fm-1] x C
                    Fm
$= 21.5 + \dfrac{9(25 - 24)}{14}$
$= 21.5 + 0.643$
$= 22.14$

2.2.3 Advantages of Median
- The value of the median could be one of the data values.
- Unlike the mean, the median is not distorted by extreme values.
- All the observations are used to order the data even though only the middle one or two observations are used in the calculation.
- Another notable merit of a median is that it can be illustrated graphically in a very simple way.

2.2.4 Disadvantages of Median
- In an aggregated/grouped frequency distribution, the value of the median within the median class can only be an estimate, whether it is calculated or read from a graph
- It is difficult to manipulate the median arithmetically.
- It is of little use in calculating other statistical measures.
- Takes a long time to calculate for a very large set of data.

2.3 Mode
The mode of a dataset is the most frequently occurring data value in the data set. It is possible to have two or more values occurring with a higher frequency than the rest. When dealing with data grouped into class intervals, the class with the highest frequency is called the modal class. In other words and according to Mumba (2008) [7], the following should be considered:
- If the variable is discrete, the mode is that value of the variable which occurs most frequently. This value can be found by ordering the observations or inspecting the sample frequency distribution or its histogram.
- If the variable is continuous, the mode is located in the class with the largest frequency, and its value must intervals to have the same frequency, a set of data may have the modes.

- A set of data with one mode is called unimodal.
- A set of data with two models is called bimodal.
- A set of data with more than two modes is called multimodal.
- A set of data with no observation repeated indicates no mode.

2.3.1 Mode from Raw/Ungrouped Data
From example 3,
There is no mode as each value occurs only once.
Example 4
> Given $X_1 = 7$, $X_2 = 8$, $X_3 = 7$, $X_4 = 12$, $X_5 = 8$. Calculate the mode.

Solution
- Ordered list is 7,7,8,8,12
- The mode is 7 and 8 as they both occur twice hence a bimodal data set.

2.3.2 Mode from Grouped/Aggregated Data
It is not possible to calculate the exact value of the mode of the original data in a grouped frequency distribution, since information is lost when data are grouped. However, it is possible to make an estimation of the mode. The class with the highest frequency is the modal class. The formula below illustrates:

$$\text{MODE} = L + \left[\frac{D_1}{D_1 + D_2}\right] \times C$$

Where L= lower boundary of the modal class
D_1 = Frequency of modal class minus frequency of previous class
D_2 = Frequency of modal class minus frequency of next class i.e. below modal class
C = Class width of the modal class

Example 5
The times the teacher does not use ICTs to teach Mathematics as a result of none-availability of internet services at a named rural school are recorded for a sample of 100 Mathematics lessons as summarised below:

Time x (Minutes)	$0 \leq x < 10$	$10 \leq x < 20$	$20 \leq x < 30$	$30 \leq x < 40$	$40 \leq x < 50$	$50 \leq x < 60$	$60 \leq x < 70$	$70 \leq x < 80$	$80 \leq x < 90$
Frequency	3	13	30	25	14	8	4	2	1

Compute the modal none-usage (in minutes) of ICTs in the teaching of Mathematics due to non-availability of internet by the teacher.

Solution
- Choose class with highest frequency which is $20 \leq x < 30$
- Use the formula $\text{MODE} = L + \left[\frac{D_1}{D_1 + D_2}\right] \times C$

$$= 20 + \left[\frac{30\text{-}13}{(30\text{-}13) + (30\text{-}25)}\right] \times C$$

$$= 20 + \left[\frac{17}{17+5}\right] \times 10$$

$$= 20 + \frac{17}{22} \times 10$$

$$= 20 + 7.72727272$$

$$= 27.73$$

Thus, modal number of the teacher not using ICTs due to non-availability of internet to teach Mathematics is 27.73 minutes.

2.3.3 Merits of Mode
- It is the only average we can use when the data is not numerical
- Occasionally used as an alternative to the mean and median when the situation calls for the most popular value to represent some data (Andre, 2004) [1].
- Easy to understand
- It usually isolates extreme score or not affected by extreme scores.
- It can easily be estimated from the graph.

2.3.4 Downside of mode
- It is a measure of very limited practical value.

- It may have more than one value (although this feature helps to show of the distribution) as observed by Mumba (2008) [7].
- It does not lend itself readily to further algebraic manipulations.
- It does not in general, enter into the calculations of further statistical measure.
- It may acquire meaning if the number of measurements under consideration is fairly very large.
- It may not accurately represent the data.

2.3.5 Relationship between mean, median and mode: It should be noted that each of the relationships to be tabulated below is only an approximate rule for moderately skewed distributions:

$$\text{Median} = \frac{2\,(\text{mean}) + \text{mode}}{3}$$

$$\text{Mode} = 3(\text{median}) - 2\,(\text{mean})$$

$$\text{Mean} = \frac{3\,(\text{median}) - \text{mode}}{2}$$

2.4 Practical guidelines for the use of various measures of central tendency
My experience of teaching and lecturing mathematics and statistics for the past 15 years, I have received and dealt with many questions from pupils and students about when and where to apply the above statistical measures. Perhaps, Aggarwal (2004: 46-47) [2] provides very useful rules which I also subscribe to as follows:

2.4.1 Mean is useful
- When scores are symmetrically or nearly symmetrically distributed around a central point.
- When the situation warrants a measure of central tendency, having the greatest stability.
- When the researcher wishes to compare SD, coefficient of correlation and other statistics which are based upon the mean.

2.4.2 Median is useful
- When the exact mid-point separating the distribution into two equal halves is wanted.
- When extreme scores are present in the distribution. Extreme scores affect the mean more than the median.
- When the number of scores above or below the central tendency is known but not their exact values.
- When the upper exact limit of the top class interval (or the lower exact limit of the lowest class interval) is not known, i.e., the frequency distribution is not complete.

2.4.3 Mode is useful
- When a quick measure is all what is wanted.
- When an approximate measure of central tendency would do.
- When only the most typical value is required. For example, the most typical size of the shirt or shoes worn by an average man.

2.5 Comments on the Best Measure of Central Tendency (Mean Median and Mode)
Borrowing an analysis from Mumba (2008) [7], it has come to light that for a data set that is approximately symmetric with one mode, the mean, the mode and the median tend to have the same value. For the data set that is obviously asymmetric, it would be good to report both the mean and the median. The mean is relatively reliable, that is, when samples are drawn from the same population, the sample means tend to be more consistent than any of the averages. It should be noted however, that a more comparison of mean and the median can reveal information about skewness. Therefore, if asked to offer a critique on a result, a better answer can be given if several measures are compared. It should be born in our mind that all these measures of location replace the details of the raw data by one number, which summarises the details to tell us that one characteristic of the distribution. This observation agrees with the position taken by many statisticians whose point of departure on measures of central tendency is that in order to obtain a good summary of the data, there is need to though fully consider other statistics or estimates as well (Chimedza, *et al*, 2014[3]; Andre, 2004[1]; Aggarwal, 2004[2]).

3. Measures of Position
A measure of position basically is the value that indicates the position of other values in relation to itself. Quartile values are good examples of measures of position. The median Q_2 divides the data set into two equal parts. It also divides the area of the histogram into two halves. The lower quartile value Q_1 in turn divide the first half of the entire data set while the upper quartile value Q_3 would divide the other half of the entire data set. Thus approximately, a quarter of the values will be less than Q_1 and a quarter of the values will be greater than Q_3. Therefore, the interval Q_1 to Q_3 would then contain the middle 50% of the data values. In a nutshell, the quartiles partition the data into four equal portions (Chimedza, *et al*, 2014) [3].
The other measures of position worth mentioning are percentiles and Box-and-Whisker plot. Basically the percentiles divide a ranked data set into 100 equal parts. In this regard, the lower quartile Q_1 is the 25th percentile, Q_2 account for the 50th percentile and Q_3 is the 75th percentile. On the other hand, a box-and whisker plot of a data set is a diagram which shows the position of the quartiles and the maximum and minimum values of a data set.

4. Measures of Dispersion or Variability
It must be acknowledged that measures of central tendency and position alone are not sufficient to fully describe the distribution of a given data set. In this regard, measures of dispersion come into play because they tell us the spreading structure of a data set. Although there are several, the most applied are as follows:

4.1 Range
Range = maximum value – minimum value
- Maximum expenditure = 20
- Minimum expenditure = 4
- Range = 20 – 4 = 16

4.1.2 Range from Grouped Data
Range = Highest Class (upper Limit) – Lowest Class (Lower Limit)

4.1.3 Merits of Range
- It is a simple concept to understand
- It is easy to calculate
- One of its common practical uses is for quality control purposes
- It is a good measure for comparison as it spans the whole distribution.

4.1.4 Demerits of Range
- It has no natural partner in a measure of location
- It is not used in further advanced statistical work.
- It takes only two values into account
- It is affected by extreme scores

4.2 Variance
The variance of a population data is the mean of the squares of the deviations of the observations of all the values from the mean. In other words, the variance is the square of the standard deviation. It takes every observation into account and it is based on an average deviation from the central value.

4.2.1 Strengths of Variance
- Variance makes the measure independent of the length of the list of numbers.
- As long as the mean and standard deviation are provided, it is easy to calculate.
- It is not affected by changes in absolute values or units of measurements.
- Although it cannot be applied to other statistical analysis, it is however part of other calculations such as standard deviation and ANOVA.

4.2.2 Weaknesses of Variance
- It is measured in square units as compared with the number in the original list.
- It is not used in further statistical analysis apart from a few like standard deviation and ANOVA.
- It does not make sense for a single data set.

4.3 Standard Deviation
A standard deviation is a statistical measure which expresses the average deviation about the mean in the original units of the random variable (Mumba, 2008) [7]. In other words, standard deviation is the root of the mean of the squares of deviations from the common mean of set values. The standard deviation is by far the most widely used measures of dispersion largely due to the fact that it has statistical properties that are useful in Mathematical statistics (Chimedza, *et al*, 2014)[3]. A high standard deviation would indicate that data is very much spread out over a large area (random), whereas a low standard deviation would indicate that all the data is very similar.

4.3.1 Practical uses of Standard Deviation
- It reflects the variability of the scores.
- The standard deviation is the positive square root of variance and hence will always be positive.

- It is closely related to the most important and commonly used measure of location, i.e. the mean.
- If there is no variability among scores, that is, all the scores in the distribution is identical the value of standard deviation will be zero.
- It is easy to manipulate arithmetically.
- The standard deviation is more sensitive to variability in a group of scores and less variable in itself.
- It is frequently used in other statistical analysis and manipulation and hence more important than the other measures of variability.
- The standard deviation is used when statistic of the greatest stability is sought.
- It is less susceptible to outliers.

4.3.2 Downside of Standard Deviation
- Extreme scores can distort its value.
- Its calculation is not only tedious but also complicated.
- Particular problems can occur when calculating the standard deviation if the mean is not a whole number.

From example 2, the standard deviation would be calculated as follows:

$$\text{Standard deviation} = \sqrt{\left[\frac{1}{n-1}\left(\sum f_i x_i^2 - \frac{\left(\sum f_i x_i\right)^2}{\sum f_i}\right)\right]}$$

$$= \sqrt{\left[\frac{1}{49}\left(31469 - \frac{(1129)^2}{50}\right)\right]}$$

$$= \sqrt{\left[\frac{1}{49}(5976,18)\right]}$$

$$= \sqrt{121,96}$$

$$= 11.04$$

Example 6
A computer manufacturing company claims that the mean execution time for its computers is not more than 45 minutes. A random sample of 20 computers selected from the production line of this company showed that the mean execution time for this sample is 40 minutes with a standard deviation of 3 minutes. Construct a 99% confidence interval of the mean execution time of the computers. Is the claim by the company true?

Solution
$$[\bar{X} - t_{\alpha/2}(n-1), \bar{X} + t_{\alpha/2}(n-1)]$$
$$= [40 - t_{0.005}(19)\ 3/\sqrt{20}, 40 + t_{0.005}(19)\ 3/\sqrt{20}]$$
$$= [40 - 2.86 \times 3/\sqrt{20}, 40 + 2.86 \times 3/\sqrt{20}]$$
$$= [38.08; 41.92]$$

Interpretation of the confidence limits
- There is 99% chance that the true mean execution time of the computer lays between 38.08 minutes and 41.92 minutes.

- The width of the interval (41.92 – 38.08) is called the precision of the estimate

5. Conclusion

The aim of the paper was not necessarily to bring out new statistical formulas or concepts but to present the usual statistical calculations from a more practical and real life perspective. As argued in the abstract that most students and citizens do not find statistics real in their school academics, college education, university statistics and daily lives in general. The article started by presenting concepts and definitions of key statistical terms. These terms have carefully been chosen in order to create a cohesion and clear link and understanding between and among these statistical terms. Thereafter, measures of central tendency such as the mean mode and median have been highlighted including their practical uses, advantages and disadvantages as a way of helping the would be readers of this article to intelligently apply these statistical measures in appropriate situations. Measures of position and dispersion have equally received an extensive discussion, analysis, evaluation and critique. Thereafter, practical application of descriptive statistics occupied some considerable amount of space in this paper in order to concretise statistical concepts and theories with real life situations. It is hoped that this paper would go along way helping pupils, students, researchers and practitioners appreciate, apply and draw meaningful conclusions from collected data. Another notable contribution this paper brings to the fore is that academicians and practitioners would be linked to work together thereby collectively resolving many challenges our communities experience. The two should not work in isolation. Perhaps borrowing the words of Kwame Nkrumah would help best summarise this paper as he argued that "*theory without practice is hollow; and practice without theory is blind"*. Clearly, academicians need the practitioners while on the other hand, practitioners are thirst of academician's research findings thereby fulfilling an axiom that states that "*theory illuminates practice and practice illuminates theory*" (Omari, 2011) [9].

6. References

1. Andre F. Business Mathematics and Statistics. Mexico: South-Western Cengage Learning 2004.
2. Aggarwal YP. Statistical Methods Concepts, Application and Computation, New Delhi: Sterling Publishers 2004.
3. Chimedza C, Chipoyera HW, Mupambireyi F. Statistics for Managers, Harare: ZOU Press 2014.
4. Kasonde-Ng'andu S. Writing a Research Proposal in Educational Research. Lusaka: University of Zambia Press 2013.
5. Leave M. Research Review. London: Rowntree Foundation 2005.
6. Mann PS. Introductory Statistics. New York: John Wiley and Sons 1992.
7. Mumba C. Understanding Statistical Analysis, Lusaka: Premier Co 2008.
8. Neter J, Wasserman W, Whitmore GA. Applied Statistics. Boston: Allyn and Bacon 1988.
9. Omari, IM Educational Psychology for Teachers (2nd Ed.). Dar es Salaam: Oxford University Press 2011.
10. Phiri W. Pupils Perceptions Towards Grade Retention and its Effects on Academic Performance in Selected Grant-Aided Secondary Schools in Central Province, Zambia: A Masters Dissertation, University of Zambia, Lusaka: UNZA 2015.
11. Stattreck. AP Statistics: Online Video, Leon Schlesinger 2016; 13:41.
12. Tackman, BW. Racial Identity Beliefs and Academic Achievement: Does Being Black Hold Students Back, Social Psychology of Education 2006; 9:381-403.
13. Wegner T. Applied Business Statistics: Methods and Applications. South Africa: Creda Press 1993.