A comparison of quartile calculation methods to measure audit report lag

C.W.H. Boshoff
N. Wesson*
Stellenbosch University Business School

ABSTRACT

The importance of prompt financial reporting is well-established, and can be measured by audit report lag (ARL). Audit report lag is the time that elapses between the end of the financial year and the signing of the audit report. Dividing ARL data into quarters eases interpretation.

The literature identifies 11 quartile calculation methods without recommending any one of them. In addition, researchers dividing data into quarters often find overlapping values. In the case of ARL, for example, this means that companies with the same ARL may be categorised into different quarters.

This study applied the quartile calculation methods to four discrete ARL data sets to test whether the results differed and to recommend the most appropriate quartile calculation method. It was found that the results differed between methods, and that all methods resulted in companies with the same ARL being categorised in different quarters.

Out of the 11 methods tested, the 'Freund and Perles' method was recommended because of its ease of use. A secondary operation for treating overlapping values was also developed.

The standardisation of quartile calculation methods, incorporating the treatment of overlapping values, will improve interpretation of data divided in quarters. The consistency created will benefit both researchers and other users of data. The study recommends that the standardisation should also be extended to all quantile calculations, and that the suggested treatment should be automated as an Excel or another software application.

INTRODUCTION

Audit report lag (ARL) may be defined as the length of the delay, represented by the number of calendar days that have elapsed, between the financial year end and the date on which the audit report is signed (Knechel and Payne, 2001). Research has shown that a long ARL diminishes the value of the audit report to investors (Knechel and Payne, 2001) and increases the probability of faulty information (Blankley, Hurtt and MacGregor, 2015). The diminished value associated with a long ARL indicates that ARL could be an additional source of information for stakeholders. However, for the ARL to be relevant, stakeholders must be able to interpret the ARL in relation to those of other companies. Courtis (1976) was the first author to use quartiles to identify the relationship between the ARL and corporate profitability, corporate attributes and industry groups. He ranked all ARLs and compared the first quartile (Q1), which he called the 'fast reporters', and the fourth quartile (Q4), which he called the 'slow reporters'.

According to Altman and Bland (1994), data are often divided into thirds (terciles), quarters (quartiles), tenths (deciles) and hundredths (percentiles). The advantage of using quartiles is that this division allows the identification of the exact median (which is equal to the second quartile) and is not overly complicated (as the data are only divided into four parts). Cangur, Pasin and Ankarali (2015) support the use of quartiles, especially to investigate the basic distribution of the data and to identify outliers.

A quartile is one of three points that divide a set of data into four parts of equal size (Cangur *et al.*, 2015). The term 'quartile' should not be confused with 'quarter'. The quartile is the dividing point, while the quarter is the resultant data segment. The division is not always possible with a given data set, for which the quartiles must often be estimated by interpolation (Freund and Perles, 1987).

The same authors call these estimated quartiles 'hinges', which may be interpreted as quartiles. This study uses the term 'quartiles' to mean both quartiles and hinges.

Several methods of calculating quartiles are found in the literature. Cangur *et al*. (2015), for example, identify 16 different calculation methods, whereas Langford (2006) found 15. In the present study these different methods are applied to the data set of ARLs of JSE-listed companies. The expectation is that some companies with the same ARL will fall into different quarters (i.e., they will be divided by a quartile). Such an arbitrary division will diminish the value of the information for stakeholders, as the quarter in which a company falls will depend on the way in which data points with the same ARL were sorted. Also, having companies with the same ARL falling into different categories (e.g., 'slow ARL' and 'average ARL') makes drawing conclusions from the categories almost impossible.

The aim of this study is to alert stakeholders to the existence of the different calculation methods, and to compare them with one another. This comparison may then be used to recommend a quartile calculation method for classifying companies according to ARL.

AUDIT REPORT LAG

The concept of audit report lag

Audit report lag is a measure of how long the auditing process of an entity lasts. It is calculated by counting the duration (in calendar days) between the fiscal year end and the signing of the audit report (Abernathy, Barnes, Stefaniak and Weisbarth, 2017; Blankley *et al.*, 2015; Cohen and Leventis, 2013; Knechel and Payne, 2001). Some researchers refer to ARL as 'audit delay' (Oladipupo, 2013; Abbott, Parker and Peters, 2012; Ashton, Willingham and Elliott, 1987) or 'financial reporting delay' (Amari and Jarboui, 2013) without modifying the definition.

Several other measures of the timeousness of financial reporting exist. The duration represented by calendar days that have elapsed between the auditor's report and the annual general meeting is called 'management delay' (Oladipupo, 2013). Brown, Dobbie and Jackson (2011) called this 'the total delay reporting lag'. Brown *et al*. (2011) also describe three measures of timeousness related to the number of calendar days it takes for reported financial information to be absorbed into the share price. The three measures are: stock price reflecting income information prior to release, price discovery time regardless of the information to be priced, and the extent to which the year's share price approximates the share price two weeks after publication of financial statements.

In accordance with most of the available literature, this study used ARL (measured in calendar days) as the measurement for the timeousness of financial reporting.

Determinants of audit report lag

Several quantitative studies have been conducted to determine the factors that influence ARL. Dao and Pham (2014) categorise these factors into firm- and auditor-related factors.

Company size has been found to be negatively correlated to ARL (Abernathy *et al.*, 2017; Abidin and Ahmad-Zaluki, 2012; Karim, Ahmed and Islam, 2006).

Another way in which a firm influences the ARL is the composition of its audit committee. Sultana, Singh and Van der Zahn (2015) found that the financial expertise, experience and independence of its members led to a reduced ARL, while gender composition, committee size and meeting frequency were not found to have a significant influence on ARL.

Bad financial news tends to increase ARL (Abernathy *et al.*, 2017). It should therefore not be surprising that a qualified audit leads to significant delays in reporting (Abidin and Ahmad-Zaluki, 2012; Soltani, 2002; Whittred, 1980) Higher leverage was also found to increase ARL (Abidin and Ahmad-Zaluki, 2012; Schwartz and Soo, 1996).

Good internal control over reporting has been found to have a link to a shorter ARL (Abernathy *et al.*, 2017). However, Karim *et al*. (2006) found in Bangladesh that ARL increased after introducing internal control legislation. A factor related to internal control is external control, or regulations. Most countries have regulatory bodies, and most stock exchanges have regulations related to financial reporting (Karim *et al.*, 2006). For example, the listing requirements of the Johannesburg Stock Exchange (JSE) (2017) specify that an interim report must be filed within three months of the financial half-year-end, while audited financial statements must be distributed to shareholders within three months after year-end – or, alternatively, a provisional report can be published within three months, followed by audited financial statements within four months after year-end. This means that the maximum ARL expected for a company listed on the JSE is four months. Karim *et al*. (2006) found that regulation had no effect on ARL in Bangladesh, while other studies (Cohen and Leventis, 2013; Ettredge, Li and Sun, 2006; McLelland and Giroux, 2000) reported an increase in ARL as regulation increased.

The tenure of an auditing firm at a company is an auditor-related factor. Dao and Pham (2014) found that a short tenure led to an increase in ARL. A long tenure, however, does not equate to a shorter ARL. Lee, Mande and Son (2009) found an association between short tenure and a long ARL, but did in fact find a shortening of ARL as auditor tenure increased.

Lee *et al*. (2009) found a negative association between non-audit services provided to the company by the auditing firm, and the ARL. This, in combination with the negative association between tenure and ARL, leads to the assumption that a deeper relationship between the company and its auditor results in a shorter ARL.

The specialisation of the auditing firm, another auditor-related factor, influences the level of understanding of the company's operations. The previously mentioned negative impact on ARL experienced when employing a new auditing firm can be negated by employing an auditor specialising in the particular industry (Dao and Pham, 2014). Abernathy *et al*. (2017) found that the complexity of the company's operations increased ARL, whereas Abidin and Ahmad-Zaluki (2012) found no statistically significant shortening of the ARL when employing an industry specialist such as an auditing firm.

There seems to be evidence that the type of industry in which a company operates also influences ARL. In France, financial, hotel and mining companies were found to report much more promptly than the average, while property had a long ARL (Abidin and Ahmad-Zaluki, 2012). Ashton *et al*. (1987) found industrial companies to have a longer ARL than financial companies in the United States. In Spain, Bonson-Ponte, Escobar-Rodriguez and Borrero-Dominguez (2008) found energy and communications to be the sectors with the shortest ARL, while construction and investment goods intermediates were the sectors that took the longest on average to publish audit reports.

Implication of audit report lag

The promptness of financial reporting, which implies a short ARL, has been established to be of great concern to an organisation's stakeholders (Dao and Pham, 2014; Amari and Jarboui, 2013; Cohen and Leventis, 2013; Abbott *et al*., 2012) Karim *et al*. (2006) state that timeousness is very important in emerging economies. This may be due to the underdevelopment of other ways of dispersing organisational information. The JSE (2015) characterises South African ARLs as long by international standards. Reporting deadlines differ between countries: listed companies are required to lodge their audited annual reports within 90 days in Australia, and within 120 days in the United Kingdom, while the United States apply a tiered system – with small companies having to lodge their statements within 90 days, large companies within 60 days and all others within 75 days (Brown *et al*., 2011). The average ARL reported in recent studies is 80.43 days in the United States (Dao and Pham, 2014), while developing countries generally report higher average ARLs (e.g., an average ARL of 148.7 days in Russia; 181 days in China; and 122 days in Nigeria) (Efobi and Okougbo, 2015).

The timeous release of audit reports has been shown to be associated with audit quality (Cohen and Leventis, 2013). By the same token, ARL has been shown to be a proxy for audit effort (Blankley, Hurtt and MacGregor, 2014). Because of this, Blankley *et al.* (2015) caution that an abnormally short ARL may be an indication of an audit of low quality.

Cohen and Leventis (2013) stress the importance of financial reports as a medium of accountability and a source to assist with decision-making. Given this, it is not surprising that a longer ARL has been associated with a perception of bad news in the market, leading to lower returns (Blankley *et al*., 2015; 2014; Dao and Pham, 2014; Abbott *et al*., 2012; Ashton *et al*., 1987). Karim *et al*. (2006) concur that market value could be affected by the promptness of financial reporting.

Other effects of undue delays in issuing an audit report are a decrease in the value of the financial information (Knechel and Payne, 2001), information asymmetry (Abbott *et al*., 2012) and a higher probability of future restatements (Blankley *et al*., 2015; 2014).

QUARTILES

The concept of quartiles

An exploratory data analysis (EDA) may be used to find (Williamson, Parker and Kendrick, 1989) and understand (Behrens and Yu, 2012) patterns in data. According to Behrens and Yu (2012), this stands in contrast to hypothesis testing and confirmatory data analysis. Tools commonly used in EDA include data summaries and graphic representations (Behrens and Yu, 2012; Tukey, 1974).

One of the techniques used in EDA is the box plot, which is a graphic representation of the median, the quartiles and the minimum and maximum values of a data set (Williamson *et al*., 1989; McGill, Tukey and Larsen, 1978). Sometimes outliers are included on the box plot, and then excluded for the purpose of calculating the three quartiles, the minimum and the maximum (MINITAB, 2016). The box plot is useful in understanding distribution and in comparing different distributions with one another (MINITAB, 2016). Tukey (1974), however, states that, when data can be summarised by five numbers (the minimum, the maximum and the three quartiles), no picture is needed to describe the data.

A quartile is one of the three points (the middle one of which is called the median) that divide data into four roughly equal parts (Altman and Bland, 1994; Freund and Perles, 1987). Similarly, two tertiles split the data into three pieces; four quintiles split the data into five pieces; nine deciles split the data into ten pieces; and 99 percentiles split the data into 100 pieces (Altman and Bland, 1994).

While quartiles are used to determine the distribution shape of data and to detect outliers (Cangur *et al*., 2015), Freund and Perles (1987) associate quartiles primarily with grouped data. The same authors remarked at the time that the use had started to shift towards ungrouped data owing to increased computational power. The difference between the third and first quartiles is also a useful measure of skewness, and is called the 'interquartile range' (Cangur *et al.*, 2015).

Some authors (National Institute of Standards and Technology, 1997; Freund and Perles, 1987) refer to 'hinges', which may be defined as pseudo-quartiles (National Institute of Standards and Technology, 1997). Freund and Perles (1987) describe the lower hinge as the median of the set of values smaller than or equal to the total median, while the upper hinge is the median of the set of values larger than, or equal to, the total median. This study proceeds by calling all points intended to divide data into four equal parts 'quartiles', instead of making a distinction between hinges and quartiles. This will avoid any confusion of the general term with a calculation method known as 'Tukey's hinges' (Cangur *et al.*, 2015; Peltier, 2013; Knechel and Payne, 2001). McGill *et al.* (1978) also treat hinges and quartiles as synonyms.

For grouped data, quartiles are calculated by retrospectively calculating a discrete data set (Montgomery and Runger, 2007). This implies that information is lost, and no simplicity is gained, by grouping data. It is therefore preferable to work with ungrouped, original observations when calculating quartiles.

Langford (2006) mentions four types of data sets that can be divided into quartiles. The variables used to define the four types are explained in Equation 1:

$$\begin{gathered} n \triangleq \text{The number of data points} \\ k \triangleq \text{A variable such that } k \text{ can be any positive integer} \\ r \triangleq \text{the remainder value, after dividing the number of data points by 4} \end{gathered} \tag{1}$$

From these variables, the four types of data sets can be defined as in Equation 2:

$$\begin{aligned} n &= 4k, \text{ making } r = 0 \\ n &= 4k+1, \text{ making } r = 1 \\ n &= 4k+2, \text{ making } r = 2 \\ n &= 4k+3, \text{ making } r = 3 \end{aligned} \tag{2}$$

Calculation methods

It is important for the researcher to understand how a quartile is calculated, even if it is not calculated by hand. This is because, inevitably, the process of categorising data eliminates information (Altman and Bland, 1994). The user of the categorised data should therefore be very aware of what information was lost. McGill *et al.* (1978: 13) stress this point when they state that misinterpretation is a result of the user of the data "[attempting] to gain more information from the display than it contains".

As will be seen when discussing the various calculation methods, many such methods exist. Cangur *et al.* (2015) identify 16 methods, although some of them are mathematically equivalent. Freund and Perles (1987) point out that only a quarter of positive integers yield an integer when divided by four, which explains the absence of one final method. When different methods arrive at essentially the same result, but use different inputs and require different technologies for calculation, or yield results to differing levels of accuracy, the presence of different methods can be said to assist researchers in calculation. However, as different calculation methods for quartiles achieve different results, the presence of these methods only serve to complicate the calculation and interpretation of quartiles. These differences may be affected by whether a data set has an even or odd number of data points, the distribution, and the overall set size (Cangur *et al.*, 2015). All the methods calculate a position in the data set for the quartile. When a data set contains duplicate values, it is therefore possible that data points with equal values could end up in different quarters. Hyndman and Fan (1996) have proposed a standard calculation method to avoid confusion.

Most software packages can calculate more quartiles in more than one way. However, the default option is rarely the same between two packages. The available options also do not necessarily overlap. It is therefore imperative that the user of the quartile data also knows (and specifies) which method is used. In fact, Freund and Perles (1987: 200) consider this knowledge to be "of critical concern". This will ensure the reproducibility of any results that may be achieved. For example, Blankley *et al.* (2015) report ARL in terms of sextiles without defining or describing the calculation method. Other researchers who seek, for example, to reproduce the Blankley *et al.* (2015) results, may cast doubt on the reported results simply because Blankley *et al.* (2015) failed to disclose their calculation method.

Multiple ways of calculating quartiles are used. However, for all of these methods, the first step is to order the data (Altman and Bland, 1994).

Authors and software packages choose to describe their algorithms in different ways, ranging from using natural language (Microsoft Office, 2016) to describing the probability distribution in mathematical terms (Hyndman and Fan, 1996). The present study aims to describe these methods using a constant notation, and combining algorithms that are stated differently but that will yield the same results. The general notation and definitions given in Equation 3 are used for all methods:

$$\begin{gathered} y \triangleq \text{Quartile value} \\ n \triangleq \text{Number of data points} \\ x_i \triangleq \text{Datapoint with } i \in \{1; 2; \ldots; n\} \\ p \triangleq \begin{cases} 0.25 & \textit{for the first quartile} \\ 0.50 & \textit{for the second quartile} \\ 0.75 & \textit{for the third quartile} \end{cases} \\ j \triangleq \lfloor n \cdot p \rfloor \text{ (That is, the integer part of } n \cdot p\text{)} \\ g \triangleq (n \cdot p) - j \text{ (That is, the fraction part of } n \cdot p\text{)} \end{gathered} \tag{3}$$

Method 1

This method is associated with Method 1 (or PCTLDEF = 1) of the SAS (previously 'Statistical Analysis System') statistical package (SAS Institute Inc., 2014). It is also used with Type 4, typed as *quantile (..., type = 4, ...),* of the R Statistical package (R Core Team, 2014). Hyndman and Fan (1996) refer to this as Definition 4, or $\hat{Q}_4(p)$. The method is described by Equation 4:

$$y = (1 - g)x_j + gx_{j+1} \tag{4}$$

Method 1 gives a weighted average of the two observed values between which $n \cdot p$ falls. Of the six methods described by Hyndman and Fan (1996) that yield continuous values for x_i (discussed here as Methods 1, 4, 7, 8, 9 and 10), Method 1 is the only method that does not find the sample median when the second quartile is calculated (i.e., when p is chosen as 0.5).

Method 2

This method is associated with Method 2 (or PCTLDEF = 2) of the SAS statistical package (SAS Institute Inc., 2014). It is also used with Type 1, typed as *quantile (..., type = 1, ...)*, of the R Statistical package (R Core Team, 2014). Hyndman and Fan (1996) refer to this as Definition 3, or $\hat{Q}_3(p)$. The method is described by Equation 5:

$$\begin{cases} x_j \text{ if } g < \frac{1}{2} \\ x_j \text{ if } g = \frac{1}{2} \text{ and } j \text{ is even} \\ x_{j+1} \text{ if } g = \frac{1}{2} \text{ and } j \text{ is odd} \\ x_{j+1} \text{ if } g > \frac{1}{2} \end{cases} \tag{5}$$

Method 2 implies that the observed value closest to $n \cdot p$ will be chosen as the quartile.

Method 3

This method is associated with Method 3 (or PCTLDEF = *3*) of the SAS statistical package (SAS Institute Inc., 2014). It is also used with Type 3, typed as *quantile (..., type = 3, ...)*, of the R Statistical package (R Core Team, 2014). Hyndman and Fan (1996) call this the most studied definition of quartiles, and refer to this as Definition 1, or $\hat{Q}_1(p)$. The method is described by Equation 6:

$$y = \begin{cases} x_j \text{ if } g = 0 \\ x_{j+1} \text{ if } g \neq 0 \end{cases} \tag{6}$$

According to Method 3, if $n \cdot p$ does not return an exact observable value, the next observable value will be used as the quartile.

Method 4

This method is associated with Method 4 (or PCTLDEF = 4) of the SAS statistical package (SAS Institute Inc., 2014). It is also used with Type 6, typed as *quantile (..., type = 6, ...),* of the R statistical package (R Core Team, 2014). Hyndman and Fan (1996) refer to this as Definition 6, or $\hat{Q}_6(p)$. Furthermore, this method is executed when the QUARTILE.EXC function is used in Microsoft Excel (Datapig Technologies, 2013). This is also the method used by the statistical package MINITAB (Cangur *et al.*, 2015).

This method of calculation was pioneered by Freund and Perles (1987), leading other authors (Cangur *et al.*, 2015) to refer to this as the 'Freund and Perles' method.

Method 4 is similar to Method 1 (Equation 4), except for a differing definition of j and g. The method is explained by Equation 7:

$$\begin{aligned} j &\triangleq \lfloor (n + 1) \cdot p \rfloor \\ g &\triangleq [(n + 1) \cdot p] - j \\ y &\triangleq (1 - g)x_j + gx_{j+1} \end{aligned} \tag{7}$$

Method 4 is therefore also a weighted average (like Method 1), but gives a weighted average of the two values between which $(n + 1) \cdot p$ falls. Although this method is biased, it has the advantage of being independent of the distribution of the data points (Hyndman and Fan, 1996).

Method 5

This method is associated with Method 5 (or PCTLDEF = 5) of the SAS statistical package (SAS Institute Inc., 2014). It is also used with Type 2, typed as *quantile (..., type = 2, ...)*, of the R Statistical package (R Core Team, 2014). Hyndman and Fan (1996) refer to this as Definition 2, or $\hat{Q}_2(p)$. The method is described by Equation 8:

$$y = \begin{cases} \frac{1}{2}(x_j + x_{j+1}) \text{ if } g = 0 \\ x_{j+1} \text{ if } g > 0 \end{cases} \tag{8}$$

In Method 5, if an integer is found by $n \cdot p$, the value halfway between that observable value and the next observable value will be used as the quartile. If the value calculated by $n \cdot p$ is not an integer, the next observable value will be the quartile.

Method 6

This method is executed when the QUARTILE. INC function is used in Microsoft Excel (Datapig Technologies, 2013). In previous versions of Excel, this function was called QUARTILE (Datapig Technologies, 2013; Langford, 2006). This function has been retained by later versions to ensure backward capability.

The calculation formulae behind Excel functions are not published by Microsoft. Like Method 4 (Equation 7), this method is similar to Method 1 (Equation 4), but with differing definitions for j and g. Method 6 is defined by Equation 9:

$$
\begin{aligned}
j &\triangleq \lfloor [(n-1)\cdot p] + 1 \rfloor \\
g &\triangleq \Big[[(n-1)\cdot p] + 1 \Big] \\
y &= (1-g)x_j + gx_{j+1}
\end{aligned} \tag{9}
$$

Methods 7 through 9 (respectively, Equations 10, 11 and 12) are interpolation methods, like Methods 1 (Equation 4) and 4 (Equation 7).

Method 7

This method is used in R as Type 5, typed as *quantile (..., type = 5, ...)* (R Core Team, 2014) and is called Definition 5, or $\hat{Q}_5(p)$ by Hyndman and Fan (1996). The method is described by Equation 10:

$$
\begin{aligned}
j &\triangleq \left\lfloor (n\cdot p) + \frac{1}{2} \right\rfloor \\
g &\triangleq \left[(n\cdot p) + \frac{1}{2} \right] - j \\
y &= (1-g)x_j + gx_{j+1}
\end{aligned} \tag{10}
$$

While Method 1 interpolates between two values of i, Method 7 interpolates between two knots. These knots are halfway between two values of i. Method 7 represents an old calculation method, and is usually used by hydrologists (R Core Team, 2014; Hyndman and Fan, 1996).

Method 8

This method is used in R as Type 7, typed as *quantile (..., type = 7, ...)* (R Core Team, 2014) and is called Definition 7, or $\hat{Q}_7(p)$ by Hyndman and Fan (1996). The method is described by Equation 11:

$$
\begin{aligned}
j &\triangleq \lfloor (n\cdot p) + 1 - p \rfloor \\
g &\triangleq [(n\cdot p) + 1 - p] - j \\
y &= (1-g)x_j + gx_{j+1}
\end{aligned} \tag{11}
$$

According to Hyndman and Fan (1996), Method 8 is distribution-free, but biased. In R versions prior to Version 2, this was the default calculation method for quartiles (R Core Team, 2014).

Method 9

This method is used in R as Type 8, typed as *quantile (..., type = 8, ...)* (R Core Team, 2014) and is called Definition 8, or $\hat{Q}_8(p)$ by Hyndman and Fan (1996). The method is described by Equation 12:

$$
\begin{aligned}
j &\triangleq \left\lfloor (n\cdot p) + \frac{p+1}{3} \right\rfloor \\
g &\triangleq \left[(n\cdot p) + \frac{p+1}{3} \right] - j \\
y &= (1-g)x_j + gx_{j+1}
\end{aligned} \tag{12}
$$

Method 9 is recommended by Hyndman and Fan (1996), as it is approximately median-unbiased and distribution-free.

Method 10

This method is used in R as Type 9, typed as *quantile (..., type = 9, ...)* (R Core Team, 2014) and is called Definition 9,or $\hat{Q}_9(p)$ by Hyndman and Fan (1996). The method is described by Equation 13:

$$
\begin{aligned}
j &\triangleq \left\lfloor (n\cdot p) + \frac{p}{4} + \frac{3}{8} \right\rfloor \\
g &\triangleq \left[(n\cdot p) + \frac{p}{4} + \frac{3}{8} \right] - j \\
y &= (1-g)x_j + gx_{j+1}
\end{aligned} \tag{13}
$$

Although Method 10 is distribution-free, it is only median-unbiased for normally distributed data (R Core Team, 2014; Hyndman and Fan, 1996).

Method 11

The final method discussed in this study is known as 'Tukey's hinges' (Cangur *et al.*, 2015; Peltier, 2013; Langford, 2006). It is possible to calculate this in MINITAB using the letter values option (Langford, 2006). The method is described by Equation 14:

$$
\begin{aligned}
m &\triangleq \begin{cases} \frac{1}{2}\, n \text{ if } n \textit{ is even} \\ \frac{1}{2}\,(n+1) \text{ if } n \textit{ is } \text{odd} \end{cases} \\
j &\triangleq \left\lfloor \frac{1}{2}\,(n+1) \right\rfloor \\
g &\triangleq \frac{1}{2}\,(n+1) - j \\
k &\triangleq \left\lfloor \frac{1}{2}\,(m+1) \right\rfloor \\
h &\triangleq \frac{1}{2}\,(m+1) - k \\
l &\triangleq \left\lfloor n - \frac{m+1}{2} \right\rfloor \\
o &\triangleq n - \frac{m+1}{2} - l \\
y &= \begin{cases} (1-h)x_k + hx_{k+1} \text{ if } p = 0.25 \\ (1-g)x_j + gx_{j+1} \text{ if } p = 0.5 \\ (1-o)x_l + ox_{l+1} \text{ if } p = 0.75 \end{cases}
\end{aligned} \tag{14}
$$

Because y only has a defined value if p is equal to 0.25, 0.5 or 0.75, Method 11 as defined here can only be used to calculate quartiles (i.e., no other quantiles).

In practice, Method 11 works well if calculated by hand, which is done as described in the next paragraph (Peltier, 2013).

The ordered data set is divided in two by calculating the median. If the number of data points n is even, the resulting halves are taken as new data sets. If n is odd, the median is included in both halves, which makes the median the maximum value of the lower half and the minimum value of the upper half. The first quartile is the median of the lower half, while the third quartile is found by determining the median of the upper half.

Method 11 uses a cumbersome algorithm compared with most other methods. However, the algorithm is extremely easy to calculate by hand. These characteristics made Method 11, or Tukey's hinges, very practical before computers became commonplace, but they also explain why so few software packages use this method.

The study found and described 11 quartile calculation methods. A summary of the software packages in which the various methods are used is contained in Table 1.

RESEARCH PROBLEM AND OBJECTIVES

The importance of the timeousness of financial reporting is well-established, and has been thoroughly investigated. ARL is a commonly used measure of timeousness. One way in which to extract useful information from ARL data is to divide the data into four groups by calculating the three quartiles.

As mentioned earlier, 11 calculation methods for calculating quartiles have been identified. There is no literature recommending any one of the methods.

The primary objective of this study was to compare different quartile calculation methods. This comparison should determine whether, and to what extent, the different calculation methods result in different quartiles. If a difference is indeed found between methods, the secondary objective is to recommend one of the calculation methods, or a variation on one of the methods. Although the application in this study was on ARL, the accomplishment of the research objectives is applicable to any data set that is to be divided into four parts. The tertiary objective is to develop a procedure for the recommended method to address overlapping values. Overlapping values occur when two data points with the same value fall into different segments. In most data sets (including the sets used in this study), data points of equal value are sorted alphabetically – which means that a company's name affects their position in the data set.

RESEARCH METHODOLOGY

The research approach followed in this study is quantitative, using secondary data obtained from published annual reports. Financial databases (e.g., IRESS) do not publish audit report lag, the date of the audit report, nor the actual end date of the published annual report (e.g., a company with a 31 March year-end may have an actual end of the financial year of 1 or 2 April, etc.). The ARL data were therefore obtained from the Stellenbosch University Business School ARL databank. The compilation of the ARL databank entailed a rather labour-intensive method of capturing the relevant dates from the published annual reports and calculating the ARL as the difference, in calendar days, between the end of the financial year of the company and the date that the annual report was signed by the auditor. The ARL data of companies listed on the main board of the Johannesburg Stock Exchange (JSE) for the period 2002 to 2014 were included in the data set. Because the ARL of each company was calculated in calendar days and recorded individually, the data set contained discrete (as opposed to continuous) and ungrouped data.

The years indicated in Table 2 (2002, 2009, 2011 and 2013) were then chosen to comply with Langford's (2006) statement that only one data set for each $r \in \{0; 1; 2; 3\}$ is required to test the effect of the different calculation methods for quartiles. The four years were therefore chosen deliberately, and can be said to be a non-probability purposive sample (Welman, Kruger and Mitchell, 2012).

TABLE 1
SUMMARY OF CALCULATION METHODS
USED IN FOUR COMMON SOFTWARE PACKAGES

	SAS	R	Excel	MINITAB
Method 1	PCTLDEF = 1	quantile(..., type = 4, ...)		
Method 2	PCTLDEF = 2	quantile(..., type = 1, ...)		
Method 3	PCTLDEF = 3	quantile(..., type = 3, ...)		
Method 4	PCTLDEF = 4	quantile(..., type = 6, ...)	QUARTILE.EXC	Default
Method 5	PCTLDEF = 5	quantile(..., type = 2, ...)		
Method 6			QUARTILE.INC	
Method 7		quantile(..., type = 5, ...)		
Method 8		quantile(..., type = 7, ...)		
Method 9		quantile(..., type = 8, ...)		
Method 10		quantile(..., type = 9, ...)		
Method 11				Letter values

TABLE 2
YEARS SAMPLED TO GENERATE TEST DATA

Sample years and their properties			
Year	***n* =**	***r* =**	***n* given by**
2002	306	2	$4k + 2$
2009	243	3	$4k + 3$
2011	265	1	$4k + 1$
2013	264	0	$4k$

The quartiles for each year were calculated using, in turn, each of the 11 identified methods. The main objective was to ascertain whether the results of the 11 methods differed. The methods were also evaluated and compared for their applicability to the problem of ARL evaluation, by looking at the following three criteria:

- Are companies with the same ARL values placed into different quarters?
- Are the resultant quarters of equal or similar size?
- Can researchers use the method easily and repetitively?

Three possible approaches to handling overlapping values were also tested, in order to satisfy the tertiary objective.

RESULTS

The ARL for the four years tested (2002, 2009, 2011 and 2013) showed an average ARL of 80.43 days. The average ARL resembles the averages generally reported in developed countries such as the United States, but is lower than the averages reported in most developing countries. Minimum values of 26 days (for 2002 and 2013); 27 days (for 2009); and 29 days (for 2011) were observed, while maximum values of 260 days, 333 days, 188 days and 212 days were observed for 2002, 2009, 2011 and 2013 respectively. Some audit reports, therefore, were signed within one month after year-end, while the maximum values reflected periods ranging from more than six months to almost one year after year-end. Dividing the data into quarters will therefore allow a better interpretation of ARL to be able to distinguish between 'fast' and 'slow' reporters.

The primary objective of this study was to show the difference between the 11 identified methods of calculating quartiles. As shown in Table 3, the different methods did indeed yield different positions in the data set for the quartiles. Table 4 shows that, because of the prevalence of duplicate values, the ARL values (in calendar days) could be the same between calculation methods even while the position in the data set differed. Therefore, if the results of quartile calculations were given for the ARL (instead of as a position in the data set), the differences between methods were masked. This masking effect made the primary objective of showing the differing results even more important, as researchers who are not aware of the different methods will not notice the differing results obtained when using different software packages.

TABLE 3
POSITIONS OF QUARTILES FOR THE 11 CALCULATION METHODS

	2002			2009			2011			2013		
	Q1	Q2	Q3	Q1	Q2	Q3	Q1	Q2	Q3	Q1	Q2	Q3
M1	76.50	153.00	229.50	60.75	121.50	182.25	66.25	132.50	198.75	66.00	132.00	198.00
M2	77.00	153.00	230.00	61.00	122.00	182.00	66.00	133.00	199.00	66.00	132.00	198.00
M3	77.00	153.00	230.00	61.00	122.00	183.00	67.00	133.00	199.00	66.00	132.00	198.00
M4	76.75	153.50	230.25	61.00	122.00	183.00	66.50	133.00	199.50	66.25	132.50	198.75
M5	77.00	153.50	230.00	61.00	122.00	183.00	67.00	133.00	199.00	66.50	132.50	198.50
M6	77.25	153.50	229.75	61.50	122.00	182.50	67.00	133.00	199.00	66.75	132.50	198.25
M7	77.00	153.50	230.00	61.25	122.00	182.75	66.75	133.00	199.25	66.50	132.50	198.50
M8	77.25	153.50	229.75	61.50	122.00	182.50	67.00	133.00	199.00	66.75	132.50	198.25
M9	76.92	153.50	230.08	61.17	122.00	182.83	66.67	133.00	199.33	66.42	132.50	198.58
M10	76.94	153.50	230.06	61.19	122.00	182.81	66.69	133.00	199.31	66.44	132.50	198.56
M11	77.00	153.50	229.00	61.50	122.00	181.50	67.00	133.00	198.00	66.50	132.50	197.50

TABLE 4
AUDIT REPORT LAG VALUES OF QUARTILES FOR THE 11 CALCULATION METHODS

	2002			2009			2011			2013		
	Q1	Q2	Q3	Q1	Q2	Q3	Q1	Q2	Q3	Q1	Q2	Q3
M1	56.00	68.00	90.50	57.00	72.00	87.25	59.00	74.00	88.75	58.00	73.00	88.00
M2	56.00	68.00	91.00	57.00	72.00	87.00	59.00	74.00	89.00	58.00	73.00	88.00
M3	56.00	68.00	91.00	57.00	72.00	88.00	59.00	74.00	89.00	58.00	73.00	88.00
M4	56.00	68.50	91.25	57.00	72.00	88.00	59.00	74.00	89.00	58.00	73.00	88.00
M5	56.00	68.50	91.00	57.00	72.00	88.00	59.00	74.00	89.00	58.00	73.00	88.00
M6	56.00	68.50	90.75	57.00	72.00	87.50	59.00	74.00	89.00	58.00	73.00	88.00
M7	56.00	68.50	91.00	57.00	72.00	87.75	59.00	74.00	89.00	58.00	73.00	88.00
M8	56.00	68.50	90.75	57.00	72.00	87.50	59.00	74.00	89.00	58.00	73.00	88.00
M9	56.00	68.50	91.08	57.00	72.00	87.83	59.00	74.00	89.00	58.00	73.00	88.00
M10	56.00	68.50	91.06	57.00	72.00	87.81	59.00	74.00	89.00	58.00	73.00	88.00
M11	56.00	68.50	90.00	57.00	72.00	87.00	59.00	74.00	88.00	58.00	73.00	88.00

It is also evident from Table 4, based on the median (or Q2), that most companies' auditors signed off their audit reports after two months but before three months in the relevant years; the 'fast' reporters (or quarter 1) resembled companies where auditors signed off the audit report before two months after year-end; whereas the 'slow' reporters (or quarter 4) resembled companies where the auditors signed off the annual reports after three months (until as late as six months or a year subsequent to the financial year-end). Although it may be argued that other methods of dividing the data (e.g., deciles or quintiles) might provide an even better interpretation of the results than quartiles, the aim of this study was to test the method (quartiles) that is widely used in all statistical packages. The conclusion about the application of quartiles can then be used to evaluate the application on other methods of dividing data.

The secondary objective of the study was to determine whether one of the methods could be recommended as superior to the other methods.

The first of the three criteria for a superior calculation method, as stated earlier, is that companies with the same ARL should not be classified into different quarters. The 11 methods described earlier all determine the quartiles by calculating a position in the data set, as opposed to an ARL value. It logically follows that all the methods can divide companies with the same ARL into different quarters.

Second, to determine whether the resultant quarters are of a similar size, the number of data points in the largest quarter is subtracted from the number in the smallest quarter for each of the calculation methods. The results of this operation are shown in Table 5, where it is demonstrated that, despite the slightly lower average size difference of methods 4 to 10, performance between the methods is remarkably similar.

The final criterion is ease of use. Table 1 indicates that Method 4 may be used in any of the four studied software packages, making this the easiest of the methods to use. Methods 6 and 11 can only be used in one of the four packages; but Method 11 is extremely easy to do by hand, making Method 6 the most difficult of the methods to use. As the other two criteria (overlapping and size) do not yield a compelling reason to choose one method over another, the prevalence of Method 4 should make it the recommended calculation method.

In order to satisfy the tertiary objective, three approaches to dealing with overlapping values were tested. As previously indicated, these modified quarters will be called 'segments' to avoid confusion with the calculated quarters.

Approach 1 entailed using the ARL value calculated (and given in Table 4) as the exclusive limit of each segment. For example, as displayed in Table 6, the calculated ARL value for Q3 in 2013 using Method 1 was 88. All companies with an ARL of less than 88 will therefore be included in the first segment when using Approach 1.

Approach 2 is similar to Approach 1, but uses the calculated value as the inclusive limit of each segment. In the example used earlier (as displayed in Table 6), the first segment will include all companies with an ARL of less than or equal to 88.

TABLE 5
THE DIFFERENCE IN SIZE
BETWEEN THE LARGEST AND SMALLEST QUARTERS

	2002	2009	2011	2013	Average	Standard deviation
M1	1	1	1	2	1.25	0.50
M2	1	2	2	2	1.75	0.50
M3	1	1	1	2	1.25	0.50
M4	1	1	1	0	0.75	0.50
M5	1	1	1	0	0.75	0.50
M6	1	1	1	0	0.75	0.50
M7	1	1	1	0	0.75	0.50
M8	1	1	1	0	0.75	0.50
M9	1	1	1	0	0.75	0.50
M10	1	1	1	0	0.75	0.50
M11	3	2	1	0	1.50	1.29

TABLE 6
ILLUSTRATION OF APPROACH 3

Excerpt from the 2013 data set, illustrating Approach 3						
ARL in days	87	88	88	88	88	89
Position in data set	196	197	198	199	200	201
Quarter as per calculation Method 1	3	3	4	4	4	4
Segment after using Approach 1	3	4	4	4	4	4
Segment after using Approach 2	3	3	3	3	3	4
Segment after using Approach 3	3	4	4	4	4	4

TABLE 7
DIFFERENCE BETWEEN
LARGEST AND SMALLEST SEGMENTS AFTER APPROACH 1

	2002	2009	2011	2013	Average	Standard deviation
M1	9	11	11	5	9.00	2.83
M2	9	7	10	5	7.75	2.22
M3	9	6	8	5	7.00	1.83
M4	5	7	8	5	6.25	1.50
M5	5	7	8	5	6.25	1.50
M6	5	7	8	5	6.25	1.50
M7	5	7	8	5	6.25	1.50
M8	5	7	8	5	6.25	1.50
M9	5	7	8	5	6.25	1.50
M10	5	7	8	5	6.25	1.50
M11	5	7	8	5	6.25	1.50
Average:					6.70	1.72

TABLE 8
DIFFERENCE BETWEEN
LARGEST AND SMALLEST SEGMENTS AFTER APPROACH 2

	2002	2009	2011	2013	Average	Standard deviation
M1	5	2	5	12	6.00	4.24
M2	5	2	5	12	6.00	4.24
M3	5	2	5	12	6.00	4.24
M4	13	2	5	12	8.00	5.35
M5	6	2	5	12	6.25	4.19
M6	6	2	5	12	6.25	4.19
M7	6	2	5	12	6.25	4.19
M8	6	2	5	12	6.25	4.19
M9	13	2	5	12	8.00	5.35
M10	13	2	4	12	7.75	5.56
M11	7	2	5	12	6.50	4.20
Average:					6.66	4.54

TABLE 9
DIFFERENCE BETWEEN
LARGEST AND SMALLEST SEGMENTS AFTER APPROACH 3

	2002	2009	2011	2013	Average	Standard deviation
M1	5	3	5	5	4.50	1.00
M2	5	3	5	5	4.50	1.00
M3	5	3	5	5	4.50	1.00
M4	5	3	5	5	4.50	1.00
M5	5	3	5	5	4.50	1.00
M6	5	3	5	5	4.50	1.00
M7	5	3	5	5	4.50	1.00
M8	5	3	5	5	4.50	1.00
M9	5	3	5	5	4.50	1.00
M10	5	3	5	5	4.50	1.00
M11	5	3	5	5	4.50	1.00
Average:					4.50	1.00

Approach 3 was developed in an effort to balance Approaches 1 and 2. The procedure used is somewhat more cumbersome than either of the preceding approaches. If the ARL value overlaps – i.e., if there are data points with the same value – the segment that contains the greatest number of data points for this value should contain all data points of this value. For Approach 3 (as displayed in Table 6), the ARL value of 88 does overlap, as there are four instances, one of which falls into quarter 3, and three of which fall into quarter 4. After using Approach 3, all the data points with a value of 88 should be in the fourth segment, as this entails moving only one data point from its original quarter to its new resultant segment.

Table 5 was replicated for all the calculation methods after the three approaches for overlapping values were performed. The results of this operation are found in Tables 7, 8 and 9.

As with the unchanged quartile calculations, the three approaches should be judged according to the three criteria set out earlier. Because of the operations that were followed, none of the three approaches yielded segments with overlapping values. The results show that Approach 3 yielded the segments with the most similar sizes, as shown by the low average difference.

CONCLUSION AND RECOMMENDATIONS

Using quartiles

In this study, 11 methods of calculating quartile values were identified and tested, of which none show a superiority in performance. This does not mean, however, that the choice of method is inconsequential. Because none of the calculation methods can be recommended based on superior calculation properties, this study recommends one method based on prevalence and, therefore, ease of use. As the method found most commonly among the statistical software that was investigated, future researchers should use Method 4 (explained in Equation 7), also known as the Freund and Perles method.

Even after the quartile has been calculated, the overlapping values still make interpretation ambiguous. In order to eliminate these overlapping values, Approach 3 (followed in this study) should be used. In this approach, overlapping values are assigned to the segment already containing the highest number of that value.

To facilitate any large-scale research, this suggested treatment could be automated by writing an implementation in Excel or another software application.

Managerial implications

When interpreting quartile data, cognisance should be taken of the difference in results when using the available methods to do quartile calculations – and especially the effect of overlapping values. The ARL results showed that companies with the same ARL can easily be categorised into different quarters, thus complicating the distinction between 'fast', 'average' and 'slow' reporters. In the managerial context, however, this research can easily be extended beyond the scope of dividing ARL data. Examples where a consistent, transparent and non-overlapping division of data sets into segments may be needed include (among countless others) grading of employees, payment of bonuses (for example) to the top quarter of salespeople, and classifying products in terms of revenue generated.

FUTURE RESEARCH

Of the 11 calculation methods, only Tukey's hinges (Method 11, Equation 14) are presented in an equation that cannot be readily used to calculate any quantiles. This study can therefore be replicated in the calculation of deciles or quintiles, for example, by using the first 10 calculation methods. By using the same procedure and the approaches for overlapping values, the methods can be evaluated to discover whether Method 4 should also be preferred when quantiles rather than quartiles are calculated.

REFERENCES

Abbott, L.J., Parker, S. and Peters, G.F. 2012. Internal audit assistance and external audit timeliness. *Auditing: A Journal of Practice and Theory*, 31(4): 3-20.

Abernathy, J.L., Barnes, M., Stefaniak, C. and Weisbarth, A. 2017. An international perspective on audit report lag: A synthesis of the literature and opportunities for future research. *International Journal of Auditing*, 21(1): 100-127.

Abidin, S. and Ahmad-Zaluki, N.A. 2012. *Auditor industry specialism and reporting timeliness*. Paper published in Jakarta International Congress on Interdisciplinary Business and Social Science proceedings, 65: 873-878.

Altman, D.G. and Bland, J.M. 1994. Quartiles, quintiles, centiles, and other quantiles. *British Medical Journal*, 39: 996. Available: http://pubmedcentralcanada.ca/pmcc/articles/PMC2541298/pdf/bmj00461-0034.pdf [Accessed: 15 February 2017].

Amari, M. and Jarboui, A. 2013. Financial reporting delay and corporate governance from Tunisia. *International Journal of Information, Business and Management*, 5(4): 32-46.

Ashton, R.H., Willingham, J.J. and Elliott, R.K. 1987. An empirical analysis of audit delay. *Journal of Accounting Research*, Autumn 25(2): 275-292. Available: http://raw.rutgers.edu/docs/Elliott/22An Empirical Analysis of Audit Delay.pdf [Accessed: 3 April 2017].

Behrens, J.T. and Yu, C.-H. 2012. Exploratory data analysis. In Weiner, I.B. (Ed.). *Handbook of psychology*. Hoboken: John Wiley & Sons, Inc. 2: 33-64. Available: https://twu.seanho.com/08fall/cpsy501/articles/Behrens03-Exploratory_Analysis_Overview.pdf [Accessed: 15 May 2017].

Blankley, A.I., Hurtt, D.N. and MacGregor, J.E. 2014. The relationship between audit report lags and future restatements. *Auditing: A Journal of Practice and Theory*, 33(2): 27-57.

Blankley, A.I., Hurtt, D.N. and MacGregor, J.E. 2015. Are lengthy audit report lags a warning signal? *Current Issues in Auditing*, 9(2): 19-28.

Bonson-Ponte, E., Escobar-Rodriguez, T. and Borrero-Dominguez, C. 2008. Empirical analysis of delays in the signing of audit reports in Spain. *International Journal of Accounting*, 12(2): 129-140. Available: http://onlinelibrary.wiley.com.ez.sun.ac.za/doi/10.1111/j.1099-1123.2008.00375.x/epdf [Accessed: 15 August 2017].

Brown, P., Dobbie, G.W. and Jackson, A.B. 2011. Measures of the timeliness of earnings. *Australian Accounting Review*, 21(3): 222-234.

Cangur, S., Pasin, O. and Ankarali, H. 2015. Comparison of sampling distributions and performances of MINITAB and Freund and Perles quartiles. *Pakistan Journal of Statistics*, 31(1): 1-20.

Cohen, S. and Leventis, S. 2013. Effects of municipal, auditing, and political factors on audit delay. *Accounting Forum*, 37(1): 40-53. Available: https://www.researchgate.net/profile/Sandra_Cohen2/publication/256979451_Effects_of_municipal_auditing_and_political_factors_on_audit_delay/links/00b7d5360a6dca1e99000000.pdf [Accessed: 15 May 2017].

Courtis, J.K. 1976. Relationships between timeliness in corporate reporting and corporate attributes. *Accounting and Business Research*, 7(25): 45-46. Available: http://www.tandfonline.com/doi/abs/10.1080/00014788.1976.9729085?journalCode=rabr20. [Accessed: 15 Augustus 2017].

Dao, M. and Pham, T. 2014. Audit tenure, auditor specialization and audit report lag. *Managerial Auditing Journal*, 29(6): 490-512.

Datapig Technologies. 2013. *Why Excel has multiple quartile functions and how to replicate the quartiles from R and other statistical packages*. Available: http://datapigtechnologies.com/blog/index.php/why-excel-has-multiple-quartile-functions-and-how-to-replicate-the-quartiles-from-r-and-other-statistical-packages/ [Accessed: 3 February 2017].

Efobi, U. and Okougbo, P. 2012. Timeliness of financial reporting in Nigeria. *South African Journal of Accounting Research*, 28(1): 65-77.

Ettredge, M.L., Li, C. and Sun, L. 2006. The impact of SOX Section 404 internal control quality assessment on audit. *Auditing: A Journal of Practice and Theory*, 25(2): 1-23. Available: http://web.a.ebscohost.com.ez.sun.ac.za/ehost/pdfviewer/pdfviewer?sid=f040a0c9-052f-491e-b892-541b94273ad8%40sessionmgr4008&vid=1&hid=4104 [Accessed: 15 May 2017].

Freund, J.E. and Perles, B.M. 1987. A new look at quartiles of ungrouped data. *The American Statistician*, 41(3): 200-203. Available: http://www.jstor.org/stable/2685103. [Accessed: 1 February 2017].

Hyndman, R.J. and Fan, Y. 1996. Sample quantiles in statistical packages. *The American Statistician*, 50(4): 361-365. Available: http://links.jstor.org/sici?sici=0003-1305%28199611%2950%3A4%3C361%3ASQISP%3E2.0.CO%3B2-G [Accessed: 20 April 2017].

Johannesburg Stock Exchange. 2015. *Amendments to the JSE listings requirements, Part 2 of 2015*. Johannesburg. Available: https://www.jse.co.za/content/JSEAnnouncementItems/ExplanatoryMemorandum Part 2 of 2015.pdf [Accessed: 3 April 2017].

Johannesburg Stock Exchange. 2017. *JSE Limited Listings Requirements*. Johannesburg. Available: https://www.jse.co.za/content/JSERulesPoliciesandRegulationItems/JSE Listings Requirements.pdf [Accessed: 13 November 2017].

Karim, W., Ahmed, K. and Islam, A. 2006. The effect of regulation on timeliness of corporate financial reporting: Evidence from Bangladesh. *Journal of Administration and Governance*, 1(1): 15-35. Available:https: //www.researchgate.net/profile/Kamran_Ahmed7/publication/228469690_The_effect_of_regulation_on_timeliness_of_corporate_financial_reporting_Evidence_from_Bangladesh/links/53f472a60cf22be01c3eba95.pdf [Accessed: 15 May 2017].

Knechel, W.R. and Payne, J.L. 2001. Audit report lag. *Journal of Accountancy*, 192(2): 88-89. Available: http://search.proquest.com/docview/206756327/abstract/28823D14C07A4322PQ/1?accountid=14049. [Accessed: 3 February 2017].

Langford, E. 2006. Quartiles in elementary statistics. *Journal of Statistics Education*, 14(3): 1-18. Available: https://ww2.amstat.org/publications/jse/v14n3/langford.html [Accessed: 2 February 2017].

Lee, H.-Y., Mande, V. and Son, M. 2009. Do lengthy auditor tenure and the provision of non-audit services by the external auditor reduce audit report lags? *International Journal of Auditing,* 13(2): 87-104.

McGill, R., Tukey, J.W. and Larsen, W.A. 1978. Variations of box plots. *The American Statistician,* 32(1): 12- 16. Available: http://www.jstor.org/stable/2683468. [Accessed: 1 February 2017].

McLelland, A.J. and Giroux, G. 2000. An empirical analysis of auditor report timing by large municipalities. *Journal of Accounting and Public Policy*, (19): 263-281. Available: http://harbert.auburn.edu/~mclelaj/research_files/JAPP00.pdf [Accessed: 15 May 2017].

Microsoft Office. 2016. *Quartile functions support*. Available: https://support.office.com/en-us/article/QUARTILE-function-93cf8f62-60cd-4fdb-8a92-8451041e1a2a [Accessed: 21 April 2017].

MINITAB. 2016. *What is a boxplot?* Available: http://support.minitab.com/en-us/minitab/17/topic-library/basic-statistics-and-graphs/graphs/graphs-that-compare-groups/boxplots/boxplot/ [Accessed: 15 May 2017].

Montgomery, D.C. and Runger, G.C. 2007. *Applied statistics and probability for engineers* (4^{th} ed.). Hoboken: John Wiley & Sons, Inc.

National Institute of Standards and Technology. 1997. *Statistics LET subcommands*. Available: http://www.itl.nist.gov/div898/software/dataplot/refman2/ch2/upperhin.pdf [Accessed: 20 April 2017].

Oladipupo, A.O. 2013. Global demand for timely financial reporting: How prepared are Nigerian companies? *Research Journal of Finance and Accounting*, 4(8): 2222-2847. Available: http://www.iiste.org/Journals/index.php/RJFA/article/viewFile/6695/7105 [Accessed: 14 May 2017].

Peltier, J. 2013. *Hinge techniques for determining quartiles*. Available: http://peltiertech.com/hinges/ [Accessed: 3 February 2017].

R Core Team. 2014. *Sample quantiles*. Available: https://stat.ethz.ch/R-manual/R-devel/library/stats/html/quantile.html [Accessed: 20 April 2017].

SAS Institute Inc. 2014. *Base SAS 9.4 procedures guide: Statistical procedures* (4th ed.). Cary, N.C. Available: https://support.sas.com/documentation/cdl/en/procstat/67528/PDF/default/procstat.pdf [Accessed: 9 April 2017].

Schwartz, K.B. and Soo, B.S. 1996. The association between auditor changes and reporting lags. *Contemporary Accounting Research*, 13(1): 353-370.

Soltani, B. 2002. Timeliness of corporate and audit reports: Some empirical evidence in the French context. *International Journal of Accounting,* 37(2): 215-246. Available: http://ac.els-cdn.com/S0020706302001528/1-s2.0-S0020706302001528-main.pdf?_tid=08fc8712-81ec-11e7-9f78-00000aacb362&acdnat=1502823822_13e0c314f9a3bb9cb21df70d0e323787 [Accessed: 15 August 2017].

Sultana, N., Singh, H. and Van der Zahn, J.L.W.M. 2015. Audit committee characteristics and audit report lag. *International Journal of Auditing,* 19(2): 72-87.

Tukey, J.W. 1974. *Mathematics and the picturing of data*. Paper presented at the International Congress of Mathematicians. Vancouver (August).

Welman, C., Kruger, F. and Mitchell, B. 2012. *Research methodology* (3rd ed.). Cape Town: Oxford University Press Southern Africa.

Whittred, G.P. 1980. Audit qualification and the timeliness of corporate annual reports. *The Accounting Review,* 55(4): 563-577. Available: http://www.jstor.org.ez.sun.ac.za/stable/pdf/245775.pdf?refreqid=excelsior:8e79da4a075606f545a5d3c367672453 [Accessed: 15 August 2017].

Williamson, D.F., Parker, R.A. and Kendrick, J.S. 1989. The box plot: A simple visual method to interpret data. *Annals of Internal Medicine*, 110(11): 916-921. Available: https://www.researchgate.net/profile/Juliette_Kendrick/publication/20437659_The_box_plot_A_simple_visual_method_to_interpret_data/links/09e4150ad19424089b000000/The-box-plot-A-simple-visual-method-to-interpret-data.pdf [Accessed: 15 May 2017].

*To whom all correspondence should be addressed: Prof Nicolene Wesson, University of Stellenbosch Business School, nwesson@sun.ac.za